U. S. Department of Commerce

National Oceanic and Atmospheric Administration

National Weather Service

National Centers for Environmental Prediction

Office Note 449

PREDICTION SKILL OF MAJOR EXTRATROPICAL TELECONNECTION PATTERNS ON DAILY TIME SCALES CORRECTED BY AN ANALOG APPROACH

Åke Johansson[1]

SAIC/Environmental Modeling Center,
National Centers for Environmental Prediction, NWS/NOAA/DOC
5200 Auth Road, Camp Springs, MD 20746

6 March 2006

[1] Corresponding author address: NCEP/Environmental Modeling Center, Camp Springs, MD 20746
E-mail: ake.johansson@noaa.gov; Telephone: 301-763-8000, Ext. 7260; Telefax: 301-763-8545

1. Introduction

The North Atlantic Oscillation (NAO) and the Pacific North American (PNA) teleconnection patterns are the two most important modes of variability in the Northern Hemisphere extratropical atmosphere. The overall prediction skill of the boreal mid and high latitude atmosphere is therefore intimately related to the skill of predicting these modes. The two modes account for increasingly more of the variability as the averaging period is increased, and are thus more important in monthly and seasonal forecasting than in daily forecasting. Their importance is furthermore most pronounced during winter when the overall variability is highest.

The skill of predicting the NAO and PNA indices on daily time scales out to day 15 is considered in this study. The forecast skill of the two indices by state-of-the-art general circulation models used operationally for numerical weather prediction and short-term climate prediction on these as well as longer time scales has recently been reported in Johansson (2005). The aim of the present study is to investigate if the obtained skill levels by these models can be improved upon by using a post-processing method. The motivation is the considerable skill improvement in probabilistic quantitative precipitation forecasts obtained when utilizing a post-processing method based on an analog approach and a large reforecast data set [Hamill and Whitaker (2005)]. The recent availability of long and homogeneous reforecast data sets, like the one used in this study, opens up unprecedented possibilities to improve the accuracy of weather prediction (e.g., Hamill et al. 2005).

The raw forecast data and the model used to produce it is described in section 2. The procedure for calculating the indices is described in section 3. The rationale for the post-processing procedure and the methodology based on utilizing analogs of the NAO and PNA indices is described in section 3. Results are given in section 4 together with a discussion.

2. Model and data

Forecast data produced by the new Climate Forecast System (CFS) is used in this study. The CFS is a state-of-the-art coupled atmosphere-land-ocean forecast model used for monthly and seasonal forecasting which has been run operationally at NCEP since August 2004. A comprehensive description of the CFS and its properties as evaluated from a large set of retrospective forecasts is provided by Saha et al. (2005). The atmospheric component of the CFS is the NCEP Global Forecast System (GFS) model, as of February 2003. For details see Moorthi et al. (2001) and NCEP Office Note 442 (2003). The GFS model uses spherical harmonics on a global horizontal domain as basis functions and in the operational CFS a spectral triangular truncation at 62 waves (T62) is employed. A finite differencing is used in the vertical with 64 sigma layers and a model top at 0.2 hPa. The oceanic component is the GFDL Modular Ocean Model V.3 (MOM3) (Pacanowski and Griffies 1998). The horizontal domain is quasi-global and extends from 74°S to 64°N with a resolution of 1°×1° in the extratropics and a higher meridional resolution equatorwards of 30°. There are 40 layers in the vertical. The atmospheric and oceanic components are coupled with no flux adjustment or correction with exchange of daily averaged quantities once a day.

The retrospective ("hindcast") data set covers the period from 1981 to 2004, i.e., a 24 year period, and consists of 4320 retrospective forecasts. The forecasts are integrated from fifteen 00 UTC initial conditions that span the month immediately preceding each of the 12 calendar months in the 24 year period (a total of 288 months). In this study we only consider those forecasts that verify at 00 UTC on any date in the boreal winter season, i.e., December to February. This is because both the NAO and PNA are strongest in winter and because the existence of the PNA is without a doubt during winter. The total number of forecasts considered is thus $15 \cdot 3 \cdot 24 = 1080$.

3. Calculation of indices

There is no unique definition of the NAO and PNA. The most common methodologies based on field data include one-point teleconnection patterns (Wallace and Gutzler 1981), empirical orthogonal functions (EOFs) (Thompson and Wallace 1998), regionalized EOFs, rotated EOFs (REOFs) (Barnston and Livezey 1987) and empirical orthogonal teleconnections (EOTs) (van den Dool et al. 2000). Furthermore, the different techniques can be based on covariances or correlations, applied to different variables, horizontal and vertical domains, and data sets. The correlation between different NAO/PNA indices derived from different techniques (EOFs, REOFs, EOTs, traditional station based techniques) are very high, see e.g., Hurrell et al. (2003).

In this study, we use the definitions used operationally[2] at the Climate Prediction Center (CPC) of NCEP/NWS/NOAA. The methodology is based on the REOF technique as reported originally in Barnston and Livezey (1987) but adjusted later on. The REOF technique uses 10 rotated varimax spatial patterns derived from monthly mean 500 hPa geopotential anomalies in the Northern Hemsiphere between the equator and 87.5°N. Monthly mean anomaly data from all 12 calendar months are pooled together. The spatial patterns, or loading patterns, that define the NAO and PNA are subjectively determined by inspecting the 10 rotated modes. Since the NAO and PNA have the largest variability during the winter season, the loading patterns largely reflect characteristics of the cold season patterns. The patterns for the NAO and PNA as defined in this study are displayed in Fig. 1. Even though these patterns have been derived from monthly mean data they are here used for daily data. This may be justified since the spatial patterns derived from daily, monthly and seasonal data are quite similar (not shown).

The indices are calculated by projecting the daily 500 hPa geopotential anomaly fields onto the patterns of Fig. 1. The anomaly fields are obtained by removing a daily observed climatology at 00 UTC based on 17 year of data, 1979-1995 (Schemm et al. 1997).

[2] At the time of this study. Revised as of June 1, 2005

4. Preliminary considerations regarding systematic error correction

The most straightforward way to take into account a model's systematic forecast error for the purpose of improving upon the model's raw forecasts is to remove the ***flow-independent*** part of the systematic error. Let a t_F-day lead forecast of either the NAO or the PNA index valid at time t_O be denoted by $INX(t_O, t_F)$. The flow-independent systematic error corrected forecast can then be written

$$\begin{aligned} INX^{C}_{SEFI}(t_O, t_F) &= INX(t_O, t_F) - \frac{1}{T}\sum_{\tau_O=1}^{T}\left[INX(\tau_O, t_F) - INX(\tau_O,0)\right] = \\ &= INX(t_O, t_F) - SE_{FI}(t_F) \end{aligned} \tag{1}$$

where T is the number of verifying days available for determining the systematic error. Note that the systematic error is a function of forecast lead time, i.e., $SE_{FI}(t_F)$. The requirement that has to be satisfied for the systematic error to be successfully corrected, is that the true flow-independent part of the systematic error is large enough compared to the uncertainty in its estimation from the available sample. The difficulties in fulfilling this requirement in an operational numerical weather prediction setting with limited amount of homogeneous forecast data is highlighted by Saha (1992). The present study has on the contrary access to an almost unprecedented large homogeneous data set. However, this does not necessarily imply that it is large enough.

A natural further step is to also take into account that the systematic error may be ***flow-dependent***. The corresponding corrected forecast can formally be written

$$\begin{aligned} INX^{C}_{SEFD}(t_O, t_F) &= INX(t_O, t_F) - \frac{1}{\hat{T}}\sum_{\tau_O}\left[INX(\tau_O, t_F) - INX(\tau_O,0)\right] = \\ &= INX(t_O, t_F) - SE_{FD}(t_O, t_F) \end{aligned} \tag{2}$$

where the summation now is only over those verifying days τ_O when the forecast index values are in a certain category. The number of such days are denoted by $\hat{T}$. Note that the systematic error is now a function of both the initial time and forecast lead time, i.e., $SE_{FD}(t_O, t_F)$

5. Analog correction procedure

Consider a t_F-day lead forecast of either the NAO or the PNA index valid at time t_O^B, denoted $INX(t_O^B, t_F)$, here referred to as the base forecast. For the sake of clarity we consider a specific case, namely a 5 day forecast, t_F = 5 days, from an initial condition of 1 Jan1991 and a verification date, t_O^B = 6 Jan 1991. For each of the 1080 base forecasts of the NAO and PNA indices, the following post-processing procedure is performed :

1. A search for the closest forecast analog, $\mathrm{INX}(t_O^{A_1}, t_F)$, to the base forecast in all 5 day forecasts that verify at 00 UTC in the boreal winter months of December through February, in all years from 1981-2004, except for the year 1991, is performed. The exclusion of the verification year of the base forecast is a form of cross-validation. The criterion for closeness, or goodness, is the absolute value of the difference in NAO/PNA forecast indices. Suppose the closest analog in this example is the 5 day forecast verifying at 4 Feb 1984 = $t_O^{A_1}$.

2. The corrected forecast is defined to be the observed index for 4 Feb 1984, $\mathrm{INX}(t_O^{A_1},0)$.

3. Steps 1 and 2 are repeated to search for the next closest forecast analogs $\mathrm{INX}(t_O^{A_i}, t_F)$ and their corresponding corrected analog forecasts, $\mathrm{INX}(t_O^{A_i},0)$, i = 2 , 3 , … , N.

4. An ensemble analog corrected forecast, INX_A^C, is defined as the average of the first N corrected forecasts:

$$\mathrm{INX}_A^C(t_O^B, t_F) = \frac{1}{N}\sum_{i=1}^{N}\mathrm{INX}(t_O^{A_i},0) \tag{3}$$

   The reason for performing the ensemble averaging is that it acts as a filter so that only the systematic part of the error is distilled and corrected for.

In the following the steps (1)-(4) are denoted the ***analog correction procedure***.

The analog correction procedure is of the flow-dependent systematic error correction type (2) discussed above since the correction is dependent on the phase and amplitude of the NAO/PNA.

6. Results

The skill of the uncorrected as well as corrected forecasts are here measured by the following definition of correlation coefficient

$$\mathrm{COR}(t_F) = \frac{\frac{1}{T}\sum_{t_O=1}^{T}\mathrm{INX}(t_O, t_F)\cdot\mathrm{INX}(t_O, 0)}{\sqrt{\frac{1}{T}\sum_{t_O=1}^{T}\mathrm{INX}(t_O, 0)\cdot\mathrm{INX}(t_O, 0)}\sqrt{\frac{1}{T}\sum_{t_O=1}^{T}\mathrm{INX}(t_O, t_F)\cdot\mathrm{INX}(t_O, t_F)}}$$

where T is the number of verifying days. Since the indices are calculated from anomaly fields with respect to a long term climatology, a subtraction of the temporal mean is not performed when calculating the correlation coefficient. The skill measure is thus designed to give credit to correct predictions relative to a longer term climatology.

In the previous sections it is emphasized that a correction procedure will only give improvements over the raw forecasts if the systematic error is large enough and known with sufficient accuracy. With the availability of a very large homogeneous forecast data set the possibility to fulfill these requirements has increased.

The flow-independent systematic error is first calculated and examined. It is calculated in a cross-validated (CV) manner, where the systematic error to be applied in a specific winter is calculated over all the other winters, thus leaving the one considered out. In Fig. 2 is displayed the systematic error of the NAO and PNA indices as a function of consecutive winter and forecast lead time. The PNA has considerably larger systematic errors at shorter leads than the NAO, but for both indices they are relatively small compared to the standard deviation of the indices themselves, less than ~15%. The variation from winter to winter is rather modest compared to the magnitude of the error itself, which is an indication of usefulness in the present context. However, it is illustrative to compare the estimated CV systematic error for a specific year with the time mean error for that year. In Fig. 3 is therefore shown the difference between the CV systematic error for a specific winter (year) and forecast lead time and the time mean error for the same winter (year) and forecast lead time divided by the former. It is evident that the fluctuations in the winter mean error from winter to winter is large compared to the CV systematic error. The only exception is for PNA at short leads. It is therefore not to be expected that the corrections will be particularly effective (except for PNA at short leads) since the relatively modest CV systematic error (Fig. 2) is often acting in the opposite direction to the actual time mean error for the particular winter.

From the discussion above it is expected that the correction procedures (1)-(3) will give better results for the PNA than for the NAO. This is indeed the case at short leads as can be seen in Fig. 4 for procedures (1) and (2). Black curves are for the raw forecasts, blue for the flow-independent systematic error corrected forecasts, while red is for the flow-dependent systematic error corrected forecasts. However, the systematic error for the NAO is apparently too small to give any improvement at all. At longer leads the flow-independent systematic error corrected forecasts give almost identical skill as the raw forecasts, while the flow-dependent systematic error corrected forecasts cause a deterioration of skill.

The flow-dependent correction procedure (2) requires estimates of the systematic error as a function of the flow itself. In Fig. 5, left panel, is displayed how the systematic error of the NAO is distributed as a function of the value of the index itself as well as of consecutive winters for a forecast lead time of 2 days. The estimates seem to be rather consistent from winter to winter and therefore indicate potential usefulness. However, there is a caveat in the use of this kind of flow-dependent quantities. This is illustrated in Fig. 5, right panel, which shows that the error for large forecast lead times is just a linear function of the index value. (Note the difference in magnitude of the error between the two graphs.) This feature is due to the general loss of predictability for large leads, and thus not a reflection of the model's systematic error. For large forecast lead times it is therefore likely that the technique will be detrimental, since it will always result in a correction towards zero anomaly, which is broadly similar to what a regression technique would yield. This is an example of the so called "regression towards the mean" phenomenon.

The analog correction procedure is now investigated. How the skill of the analog corrected NAO/PNA index forecasts depends on the ensemble size as well as the forecast lead time is shown in Fig. 6, upper panels. As anticipated the skill increases with ensemble size. For short leads most of the improvements is achieved by just adding a few ensemble members, while for longer leads skill continues to improve when adding more and more members. In the forecast lead range 15-20 days the maximum skill is obtained when 50-100 analogs are averaged.

The improvements obtained by the analog correction procedure compared to the raw, uncorrected forecasts are shown in Fig. 6, lower panels. It is seen that the procedure in general gives a deterioration of the skill and the deterioration increases with increasing forecast lead time and decreasing number of analogs. Note that the values plotted are absolute skill differences, which leads to smaller values at longer leads due to the fact that the skill itself is then very small. The only occurrences of skill improvements are for the PNA index at short leads, less than 5 days. This is consistent with the fact that the PNA index has larger systematic error (Fig. 2) and the discussion above.

The improvements obtained by the analog correction procedure compared to the systematic error corrected forecasts (1) and (2) are shown in Fig. 7. The analog correction procedure is generally less effective than the other two correction methods.

7. Analysis and discussion of results

The NAO/PNA indices are just one-dimensional quantities that are quasi-normally distributed (Johansson 2005). With a sample size of over 1000 it is to be expected that there will be many close analogs. This is indeed the case with the closest analog virtually identical to the base index and with a decrease in the goodness of the analog that is quite slow. The fact that two index values are almost identical does not necessarily imply that the two corresponding flow fields are identical. The index value is just one number that summarizes (in a clever way) important characteristics of the large scale flow in a particular region of the globe. However, the large scale flow in the particular region has many more degrees of freedom than just one. The analog correction procedure used here focuses on only one of these degrees of freedom. A schematic illustration is given in Fig. 8 where the abscissa represents the value of the NAO index while the ordinate represents the value of another relevant, unspecified, large scale index value in the NAO region. The two filled points represent the base forecast and the closest analog forecast. In this case the analog is perfect since both points project on the same NAO index value. The open points represent the observations at the initial time of the forecast and at the time corresponding to the forecast points. The solid curves represent the trajectories of the real atmosphere, while the dashed curves represent the trajectories of the model forecasts. By replacing the forecast NAO index with the corresponding verifying observed NAO index of another forecast that has a similar projection onto the NAO does thus not necessarily imply an improved NAO index forecast, as is exemplified in Fig. 8. A way to reduce the effect of the non-systematic component of the error is to not only consider just the closest analog, but instead consider a collection, or ensemble, of the N best analogs as in (3).

Suppose that the model has no systematic error. The analog correction procedure with only one analog (the best) will then actually give an inferior forecast compared to the raw forecast. This can be deduced from the following chain of arguments:

a. The correlation between the base forecast and the best analog is very close to one.
b. The correlation between the base forecast and the verifying observation is less than one, since the forecast is not perfect.
c. The correlation between the best analog forecast and its verifying observation is less than one for the same reason as in b.
d. (a)-(c) implies that the correlation between the analog corrected forecast and the verifying observation is even more less than one than either the base or the best analog forecast.

If we denote the correlation by CORR (a)–(d) can formally be expressed as

$$\begin{cases} \mathrm{CORR}[\mathrm{INX}(t_O^B, t_F)\,, \mathrm{INX}(t_O^A, t_F\,] \approx 1 & \text{(a)} \\ \mathrm{CORR}[\,\mathrm{INX}(t_O^B, t_F)\,, \mathrm{INX}(t_O^B, 0)] = \alpha < 1 & \text{(b)} \\ \mathrm{CORR}[\,\mathrm{INX}(t_O^A, t_F)\,, \mathrm{INX}(t_O^A, 0)] = \alpha < 1 & \text{(c)} \end{cases}$$

$$\Downarrow$$

$$\mathrm{CORR}[\mathrm{INX}_A^C(t_O^B, t_F)\,, \mathrm{INX}(t_O^B, 0\,] =$$

$$\mathrm{CORR}[\mathrm{INX}(t_O^A, 0)\,, \mathrm{INX}(t_O^B, 0] \approx \alpha \cdot 1 \cdot \alpha = \alpha^2 < \alpha \qquad \text{(d)}$$

The analog correction procedure with only one analog (the best) can thus only improve upon a raw model forecast if, and only if, the model has a large enough systematic error that is known with sufficient accuracy. This criterion is also true for the two formal prototype systematic error correction methods discussed in section 4. The analog correction procedure ought to give better results than flow-independent systematic error correction when the systematic error is flow-dependent. How it compares with other flow-dependent systematic error correction methods is an open question.

While this topic has not been exhaustively investigated, the first impression is that the analog correction method does not offer any special advantages.

Acknowledgments

Hua-Lu Pan (EMC) suggested and initiated this study and I thank him for his encouragement and advice. The substantial support and advice from Suranjana Saha (EMC) and Huug van den Dool (CPC) are greatly appreciated. I thank the NCEP internal reviewers, Shrinivas Moorthi and James Purser, for their thoughtful comments which helped to improve and clarify the manuscript. I thank Stephen Lord (EMC) and Mike Pecnick (SAIC) for making this visit possible and for providing a stimulating work environment.

REFERENCES

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

Global Climate and Weather Modeling Branch, 2003: The GFS Atmospheric Model. *NCEP Office Note*. **442**, 14 pp.

Hamill, T. M., and J. S. Whitaker, 2005: Probabilistic quantitative precipitation forecasts based on reforecast analogs: Theory and application. Submitted to *Mon. Wea. Rev.*, 57 pp.

Hamill, T. M., J. S. Whitaker, and S. L. Mullen, 2006: Reforecasts: An important data set for improving weather predictions. *Bull. Amer. Meteor. Soc.*, **87**, 33-46.

Hurrell, J. W., Y. Kushnir, G. Ottersen, and M. Visbeck, 2003: An overview of the North Atlantic Oscillation. *The North Atlantic Oscillation: Climate Significance and Environmental Impact*. J. W. Hurrell et al., Eds., Amer. Geophys. Union. 1-35.

Johansson, Å, 2005: Prediction Skill of the NAO and PNA from Daily to Seasonal Time-scales. Submitted to the J. Climate.

Moorthi, S., H.-L. Pan, and P. Caplan, 2001: Changes to the 2001 NCEP operational MRF/AVN global analysis/forecast system. *NWS Technical Procedures Bulletin*, **484**, pp14. [Available at http://www.nws.noaa.gov/om/tpb/484.htm].

Pacanowski, R. C., and S. M. Griffies, 1998: *MOM 3.0 Manual*, NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, USA.

Saha, S., 1992: Response of the NMC MRF model to systematic-error correction within integration. *Mon. Wea. Rev.*, **120**, 345-360.

Saha, S., S. Nadiga, C. Thiaw, J. Wang, W. Wang, Q. Zhang, H. M. van den Dool, H.-L. Pan, S. Moorthi, D. Behringer, D. Stokes, M. Peña, S. Lord, G. White, W. Ebisuzaki, P. Peng, P. Xie, 2005: The NCEP Climate Forecast System. Accepted for publication in *J. Climate*.

Schemm, J-K. E., H. M. van den Dool, J. Huang, and S. Saha, 1997: Construction of daily climatology based on the 17-year NCEP-NCAR reanalysis. *Proceedings of the First WCRP International Conference on Reanalyses*. Silver Spring, Maryland, USA. 290-293.

Thompson, D. W. J., and J. M. Wallace, 1998: The Arctic Oscillation signature in the wintertime geopotential height and temperature fields. *Geophys. Res. Lett.*, **25**, 1297-1300.

van den Dool, H. M., S. Saha, and Å. Johansson, 2000: Empirical Orthogonal Teleconnections. *J. Climate*, **13**, 1421–1435.

Wallace, J. M., and D. S. Gutzler, 1981: Teleconnections in the geopotential height field during the northern hemisphere winter. *Mon. Wea. Rev.*, **109**, 784–812.

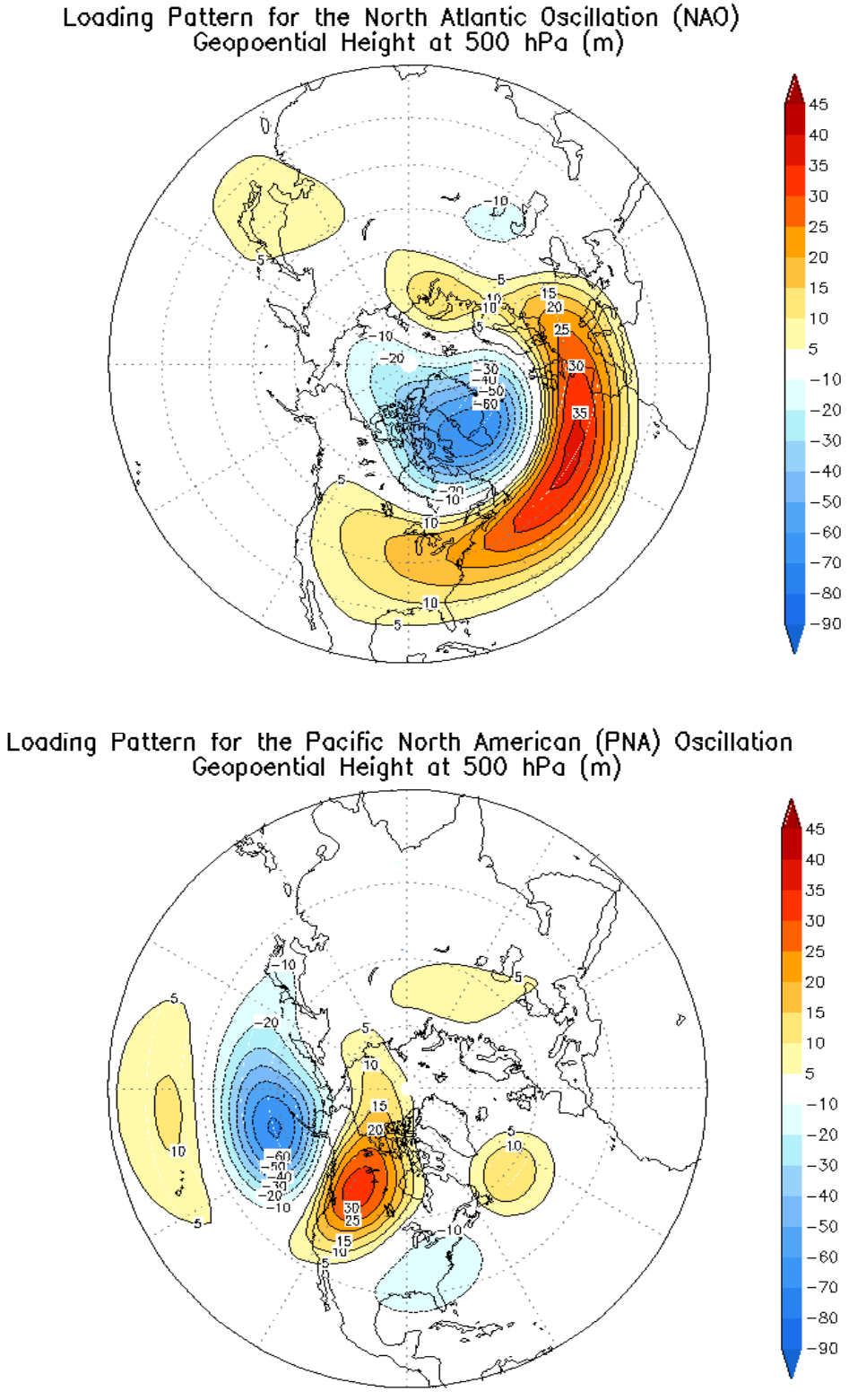

Fig. 1. The NAO and PNA patterns used operationally at CPC.

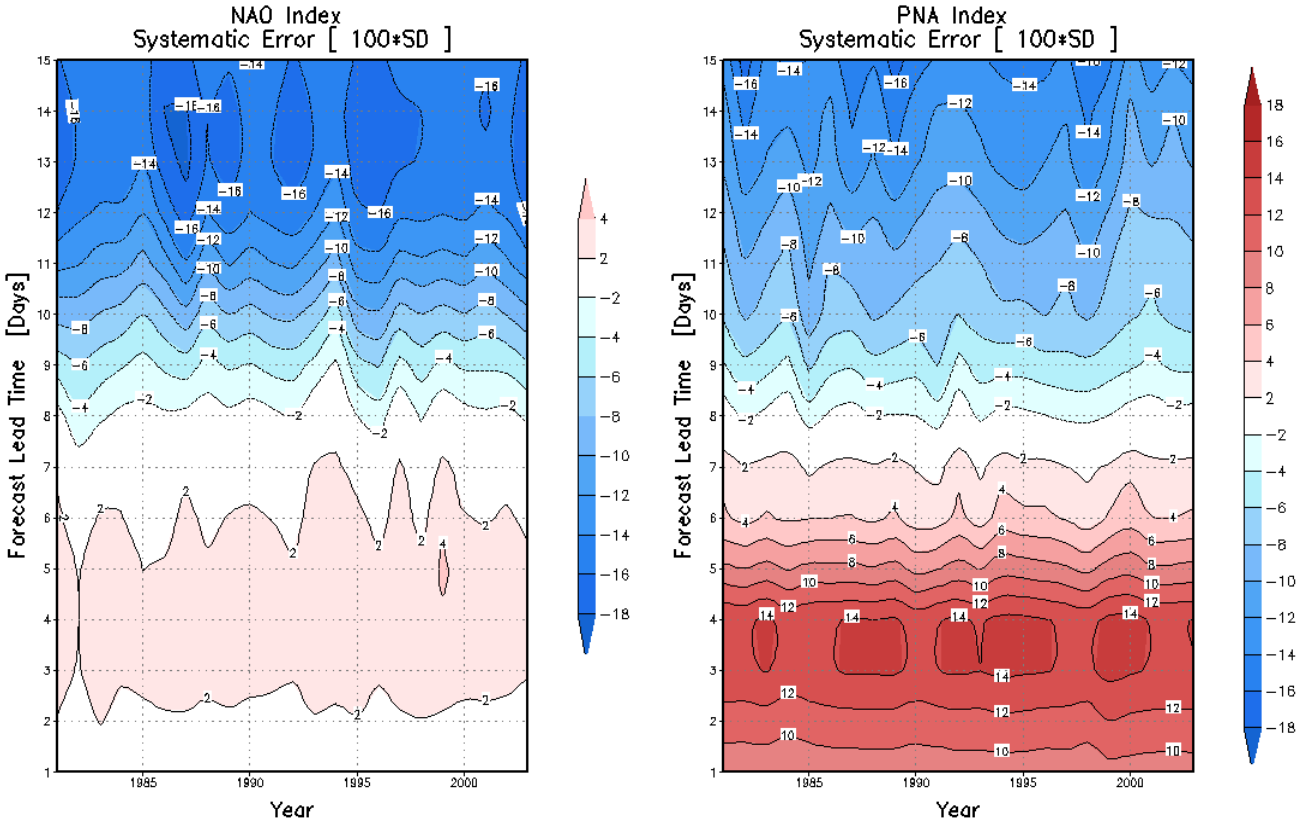

Fig. 2. The flow-independent systematic error of the NAO index [Left panel] and PNA index [Right panel] as a function of consecutive winter (denoted by the year in which the month of January occur) and forecast lead time. The values are dimensionless by division by the observed standard deviation and are furthermore multiplied by 100.

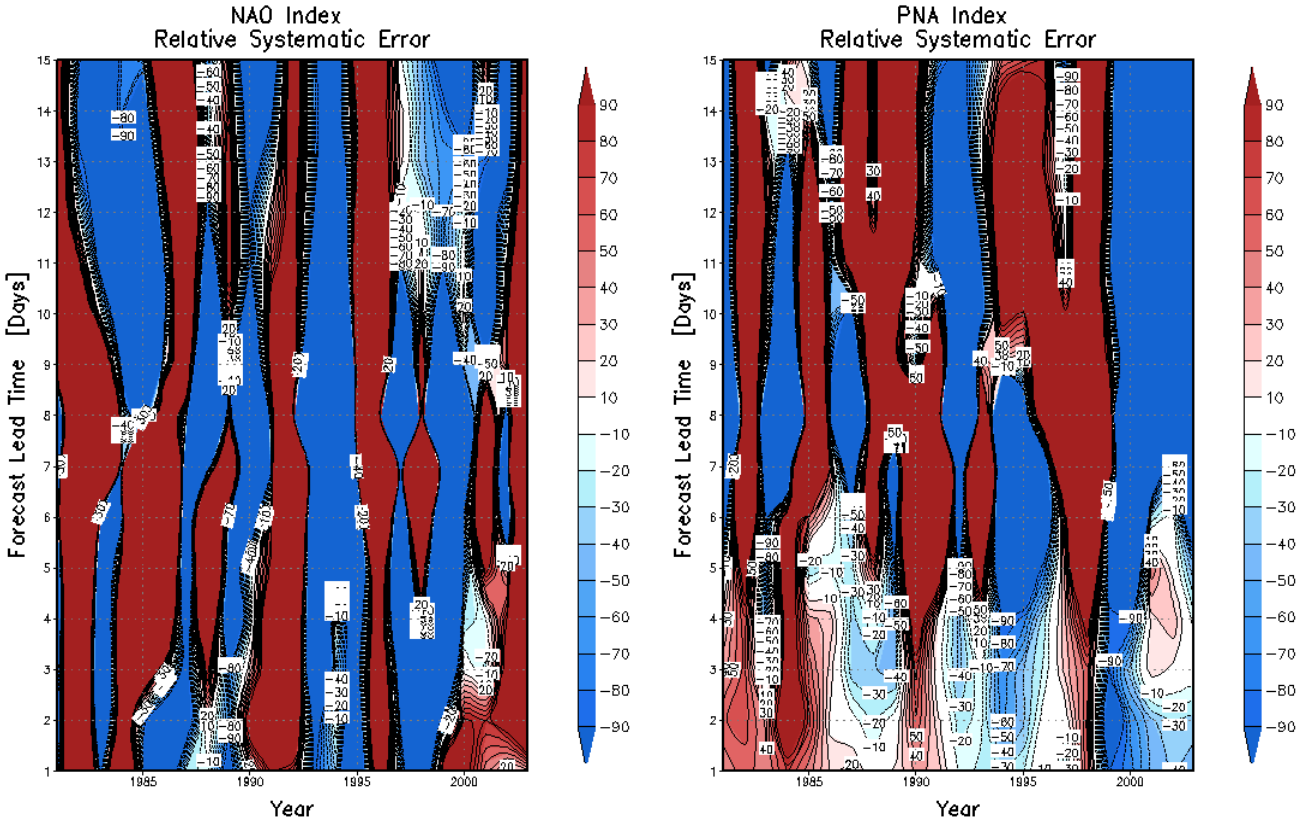

Fig. 3. The difference between the CV flow-independent systematic error for a specific winter (year) and forecast lead time and the time mean error for the same winter (year) divided by the former. The left panel is for the NAO index and the right panel is for the PNA index. Values are multiplied by 100.

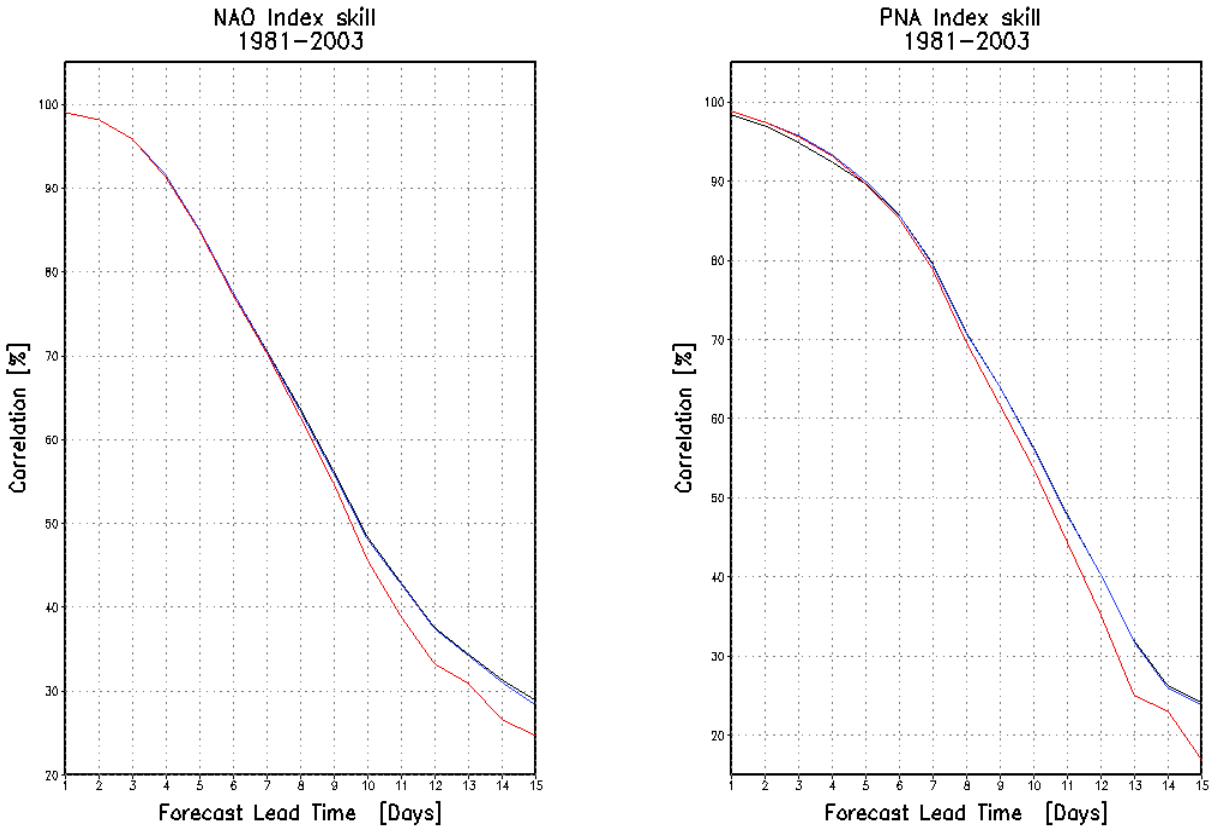

Fig. 4. The skill of predicting the NAO index [left panel] and PNA index [right panel] as measured by the correlation coefficient as a function of forecast lead time. Black curves are for the raw forecasts, blue for the flow-independent systematic error corrected forecasts, while red is for the flow-dependent systematic error corrected forecasts.

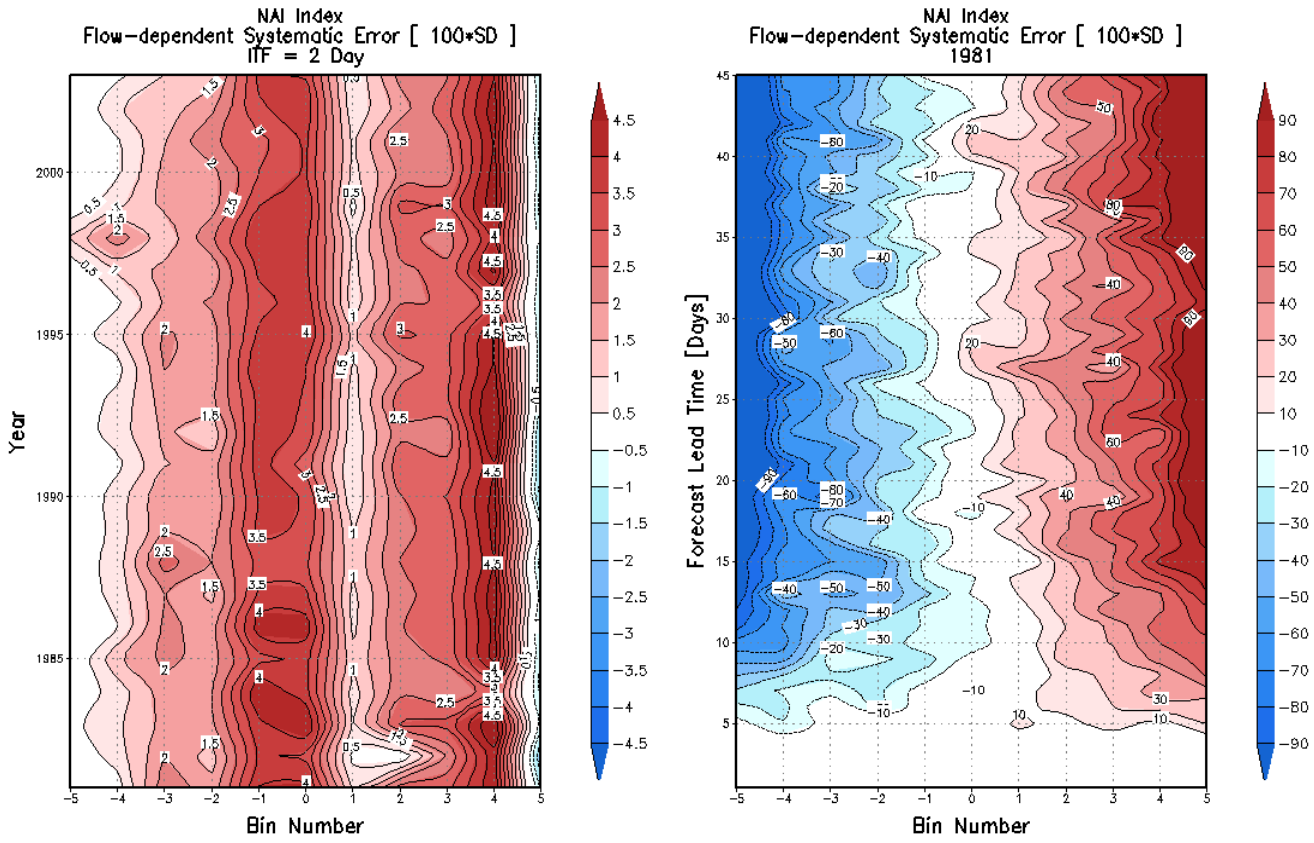

Fig. 5. Left panel: The systematic error of the NAO index as a function of the index itself and consecutive winter at a forecast lead time of 2 days. Right panel: The systematic error of the NAO index as a function of the index itself and forecast lead time in the winter of 1981-1982. The values are dimensionless by division by the observed standard deviation and are furthermore multiplied by 100. The bins are defined as follows:

Bin #	INX-interval
-5	[- ∞ , -1]
-4	[-1 , -0.7]
-3	[-0.7 , -0.5]
-2	[-0.5 , -0.3]
-1	[-0.3 , -0.1]
0	[0.1 , 0.1]
1	[0.1 , 0.3]
2	[0.3 , 0.5]
3	[0.5 , 0.7]
4	[0.7 , 1]
5	[1 , ∞]

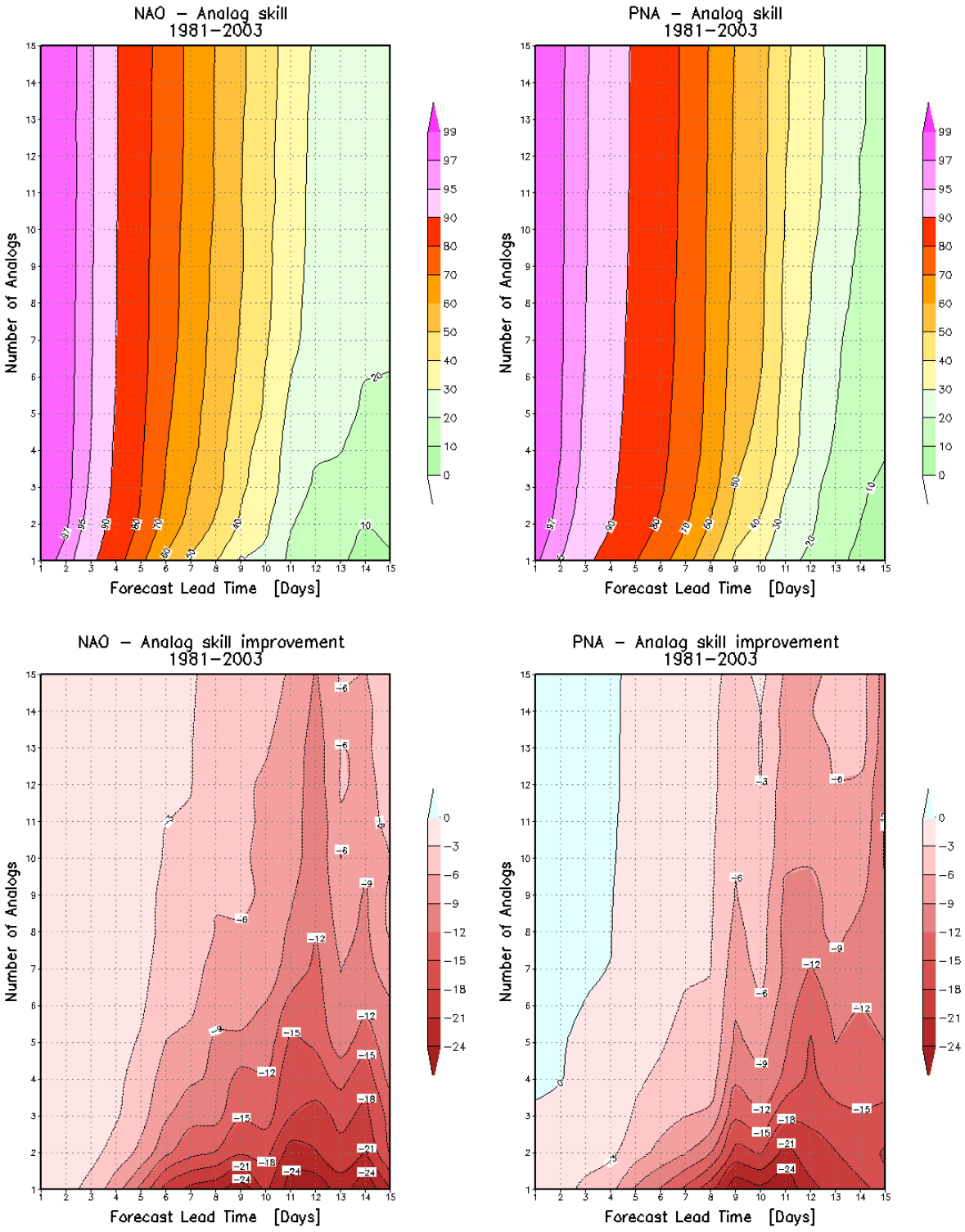

Fig. 6. Upper panels: Skill of analog corrected NAO index forecasts [Left panel] and PNA index forecasts [Right panel] as a function of forecast lead time and the number of analogs. The measure of skill is the correlation coefficient which is here multiplied by 100.
Lower panels: The same as above but skill values are relative to the skill of the raw model forecasts. Values are multiplied by 100.

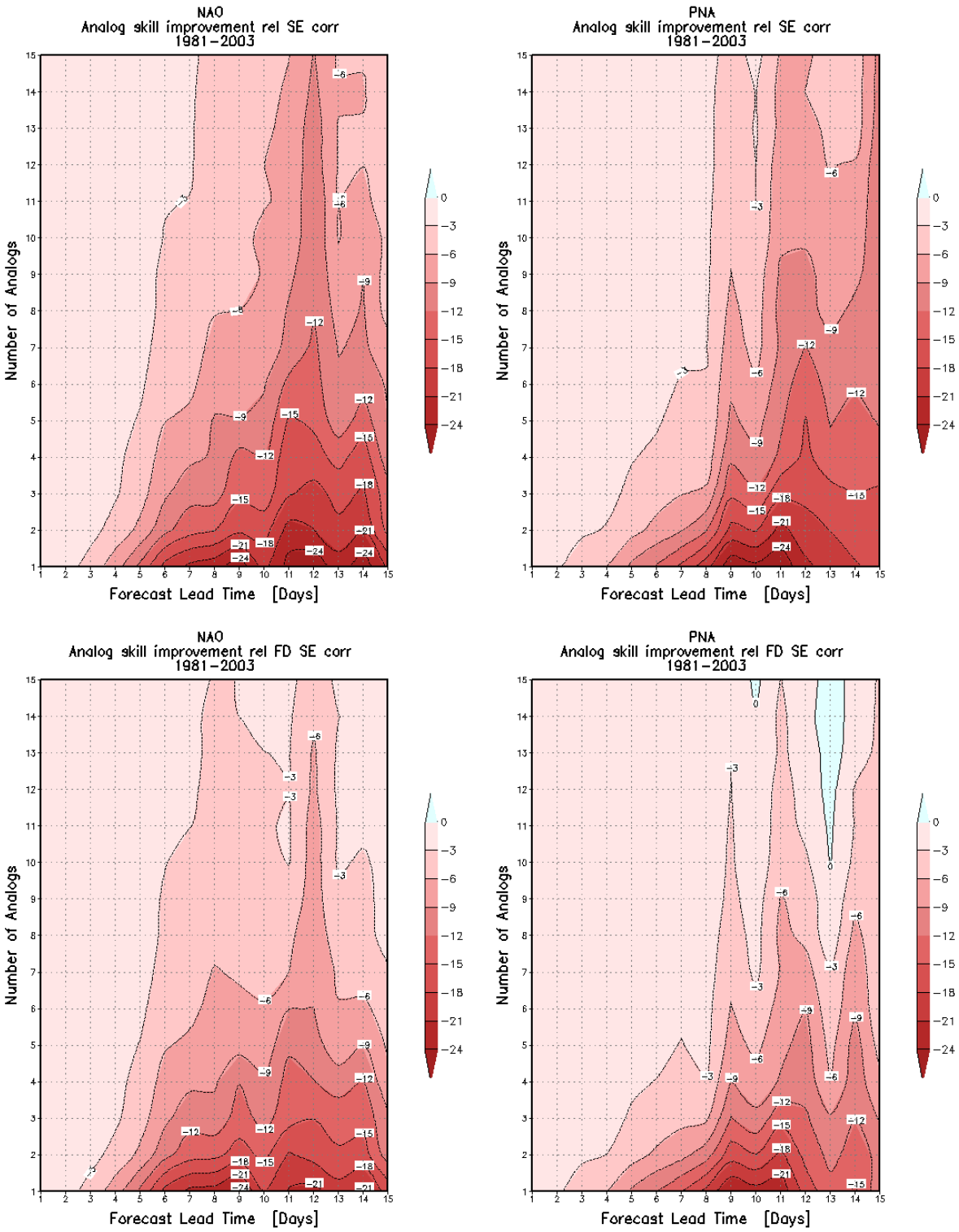

Fig. 7. Upper panels: The same as Fig. 6, lower panels, but now relative to flow-independent systematic error corrected forecasts. Lower panels: The same as Fig. 6, lower panels, but now relative to flow-dependent systematic error corrected forecasts.

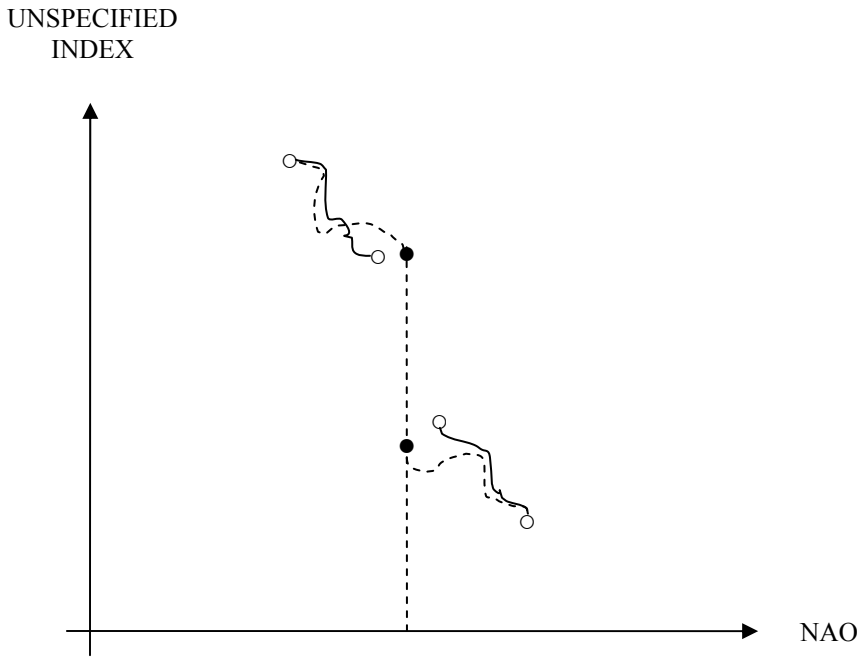

Fig. 8. Schematic picture of the analog correction procedure.