P1.4 NONLINEAR APPROACH TO WINTER NORTH ATLANTIC SLP VARIABILITY WITH THE HELP OF AUTOASSOCIATIVE NEURAL NETWORK

Ladislav Metelka *
Czech Hydrometeorological Institute, Hradec Kralove, Czech Republic

1. INTRODUCTION

North Atlantic SLP variability is in the focus of climatological research for many years. The main reason is that the North Atlantic Oscillation (NAO), which is present here, is the most pronounced atmospheric oscillatory system in the northern hemisphere. It strongly affects the occurrence of temperature climatological anomalies in Europe, especially in winter. NAO is usually described by the differences of observed sea level pressure (SLP) anomalies between one station in the region of Iceland (Reykjavik, Stykkisholmur) and the other in the region of Canary Islands (Tenerife), Azores (Ponta Delgada) or SW Europe (Lisbon, Gibraltar). More sophisticated approach to NAO is statistical, with the help of rotated principal component analysis (RPCA), originally at 700 hPa level, see Barnston (1987). Unfortunately, both these approaches provide us with linear description of NAO only.

Recently, some new definitions of NAO and NAO-index emerged, as for example so called „mobile NAO-index“, which is not based on SLP differences between fixed points (stations) but between the positions of seasonal centers of low and high pressure regions in northern Atlantic.

The difficulties in the attempts of simple description of North Atlantic SLP variability may originate from the fact that climate system is non-linear in its origin. For the same reason we may suppose that NAO itself is non-linear too (or, at least, that its variability may contain some non-linear components).

Our study may be regarded as an attempt to non-linear description of North Atlantic SLP variability. From the methodological point of view it is based on application of autoassociative neural network.

2. DATA AND DATA PRE-PROCESSING

As a data source we used the database of reanalyzed monthly mean SLP values from NCAR (geographical grid 2.5 x 2.5 degrees, period from January 1948 to December 1998). The region of interest was restricted to 90ºW - 50ºE and 10ºN - 80ºN, so it covered the North Atlantic region and great part of Europe.

Geographical grid is not homogenous from the point of view of grid points „density“ because zonal distances between neighboring grid points decrease poleward and correspond to D.cosφ (D is the zonal distance at the equator, φ is geographical latitude). For this reason the data set was transformed into new, more „regular“ grid with the distances between neighboring grid points corresponding to 5º at the equator.

At each grid point of the regular grid the SLP anomalies were calculated month by month and averaged into mean seasonal anomalies (overlapping 3-month seasons JFM, FMA, MAM, … , DJF. Finally, only „winter“ seasons (NDJ, DJF and JFM) were selected from the data set.

As the final step of data pre-processing, winter seasonal SLP anomalies in regular grid (151 cases from JFM 1948 to JFM 1998) were processed by linear non-rotated principal component analysis (LPCA) – see Monahan ((2001), Hsieh (2001a). It is worth to stress that LPCA was not used here for pattern recognition but for dimensionality reduction only. The first LPC explains about 45% of the original SLP variability, the second one about 14% and the third one about 12% (LPC1-LPC3 together explain about 71%). LPC1 – LPC3 were selected for the consecutive processing. Each of following LPCA components explains less than 10%.

LPC1-LPC3 loadings are shown in Fig.1. It is evident from the figure that the 1st LPC mode corresponds well the NAO, the 2nd mode is similar to a combination of East Atlantic and Scandinavian oscillations and the 3rd mode resembles to the East Atlantic/West Russian oscillation. Exact and unambiguous identification of the „prototype“ oscillation patterns in LPC1-LPC3 is not possible because the oscillation „prototypes“ - see Barnston (1987) - were defined with the help of rotated PCA (in our study the non-rotated PCA was used) at 700 hPa level (here the SLP values were processed) and in hemispheric scale (in our study data from North Atlantic region were processed). Moreover, the definition of the „regular“ grid in our study differs from the „regular“ grid in Barnston (1987).

3. AUTOASSOCIATIVE NEURAL NETWORK

The scores of LPC1 – LPC3 were non-linearly combined by the means of 5-layers autoassociative neural network – see Hsieh (2001a). Linear activation functions were used in the 1st (input), the 3rd („bottleneck“) and the 5th (output) layers, hyperbolic tangens activation functions in the 2nd („encoding“) and the 4th („decoding“) layers. It is evident that the

Corresponding author address: Ladislav Metelka, Czech Hydrometeorological Institute, Dvorska 410, 503 11 Hradec Kralove, Czech Republic; e-mail: metelka@chmi.cz

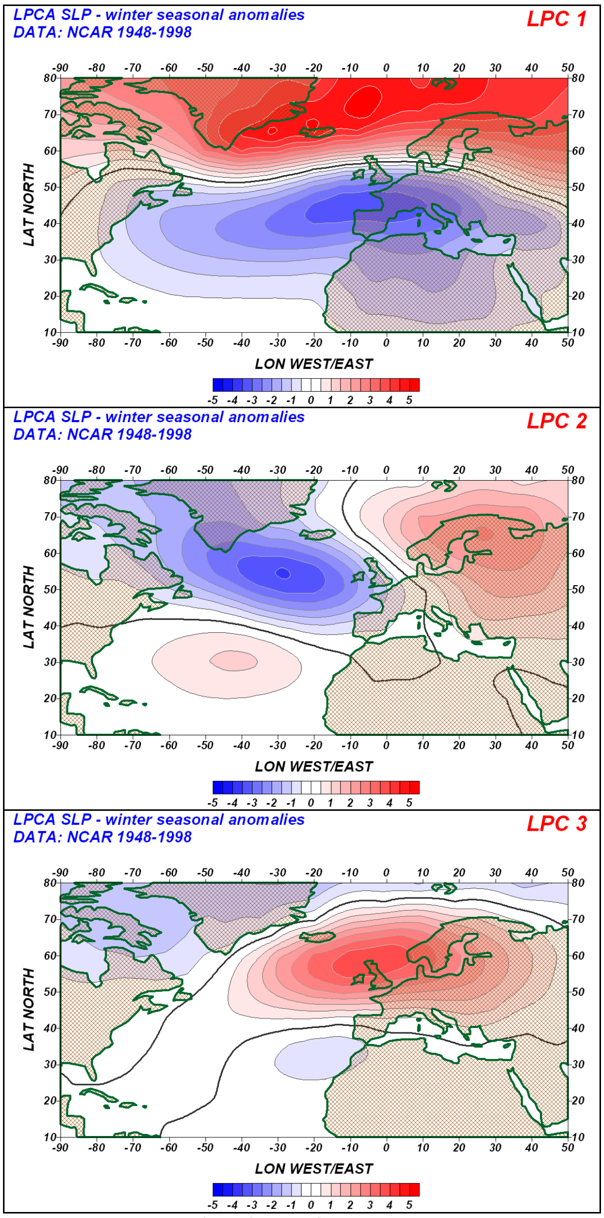

Fig.1: Loadings of LPC1-LPC3.

first non-linear transformation of the signal is present in encoding layer (signal transformation from the input layer to the encoding layer is linear) and the input data (LPCA scores) are linearly independent. So it is not possible to combine the LPC1-LPC3 scores linearly into lower number of variables without significant loss of information. For this reason the minimal possible number of neurons in encoding layer and decoding layer is the same as the number of input (output) variables. On the other hand we wanted to use as simple network as possible. So the network architecture 3-3-1-3-3 was selected (see Fig.2)

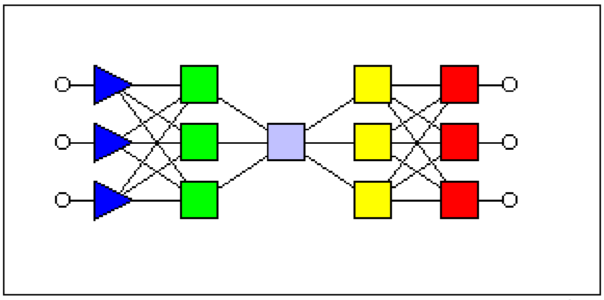

Fig.2: Autoassociative neural network 3-3-1-3-3 for non-linear principal component analysis. Blue – input layer, green – encoding layer, gray – „bottleneck“ layer, yellow – decoding layer, red – output layer.

Scores of LPC1 – LPC3 were linearly transformed to the interval from –1 to +1 and used as both input and output data. The back error propagation (BEP) network training was cross-validated with 101 randomly selected cases in training subset, 25 cases in verification subset and 25 cases in the independent testing subset. The random division to training/verification/testing subsets was repeated 20 times and for each division 10 trainings from random initial conditions were carried out. The best network was then selected with respect to training, verification and testing errors (networks with low errors of similar magnitude were preferred) and predictor/predictand correlations for all 3 LPC scores (the ability of the network to decode the LPC scores correctly after their encoding was tested). Moreover, weak Weigend regularization (Weigend λ=0.0005, scale factor=1) was used.

Best network from 200 networks trained was selected subjectively, with respect to the criteria mentioned above. Selected network was then split into encoding (from the input layer to the „bottleneck“ layer) and decoding (from the „bottleneck“ layer to the output layer) parts. The „scores“ of the first non-linear PCA (NLPCA) were calculated from the input data (LPCA scores 1-3 of the seasonal SLP anomalies) with the help of encoding part of the network. Then the quantiles (from 5% to 95% with the step of 5%) of these NLPC scores were calculated and consecutively these quantiles were processed by decoding part of the network. Scores of LPCA were reconstructed in this way for the empirical quantiles of NLPCA „scores“ mentioned above and then the corresponding spatial patterns were calculated.

4. RESULTS

As for the ability of the 1st linear principal component to reproduce the real field of SLP anomalies, correlations between winter SLP anomalies and the 1st linear PC loading were calculated (Fig.3).

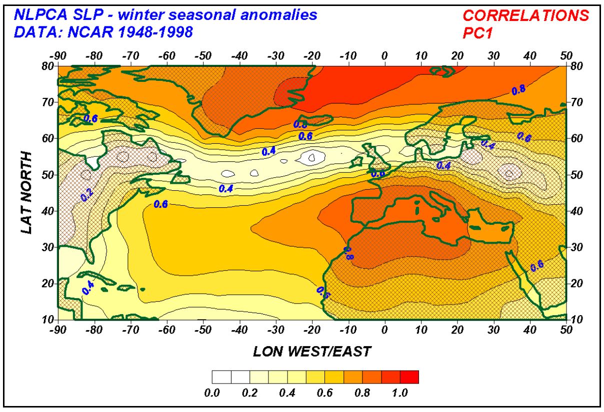

Fig.3: Correlations between winter SLP anomalies and the 1st linear PC loading

It is evident that the 1st LPC reproduces the real field of SLP anomalies well, especially near the regions of the individual NAO centers. On the other hand, correlations are weak (and even statistically insignificant at 5% significancy level) in the strip between the individual NAO centers. It may indicate the presence of SLP variability, which is not linearly linked to NAO, in this area.

For the comparison the correlations between winter SLP anomalies and non-linear principal component loading were calculated (see Fig.4)

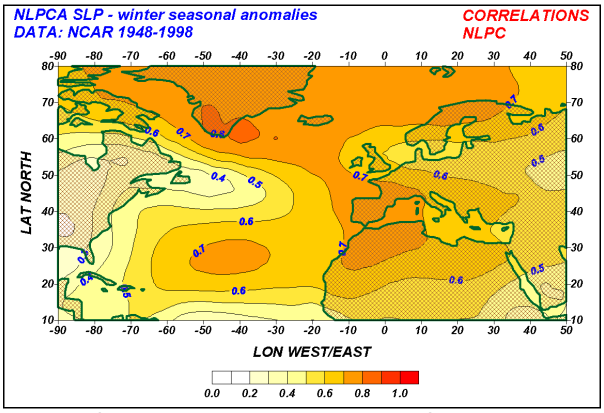

Fig.4: Correlations between winter SLP anomalies and the non-linear PC loading

In this case the strip of low correlations is weaker. It indicates that the non-linear principal component loading reproduces the real SLP variability better than the linear one. As for the explained variability, the NLPC explains about 46% of the original SLP variability, which is similar to the portion of variability explained by the 1st LPC (45%). Lower correlations near the East and West border of the region of interest are affected by the presence of other teleconnections (PNA, East Atlantic/West Russian) which are not fully covered by data in the region of interest and so they are not present in leading LPC modes.

The trained autoassociative neural network (AANN) as a whole maps the 3-dimensional signal (LPCA scores 1-3 in the input layer) onto 1-dimensional one (NLPCA scores in the „bottleneck" layer) and then „reconstructs" the original 3-dimensional signal („reconstructed" LPCA scores 1-3 in the output layer) as well as possible. The AANN performance may be seen in the graphs showing the relations between the original LPCA scores, NLPCA scores and „reconstructed" LPCA scores. These graphs are in fig.5.

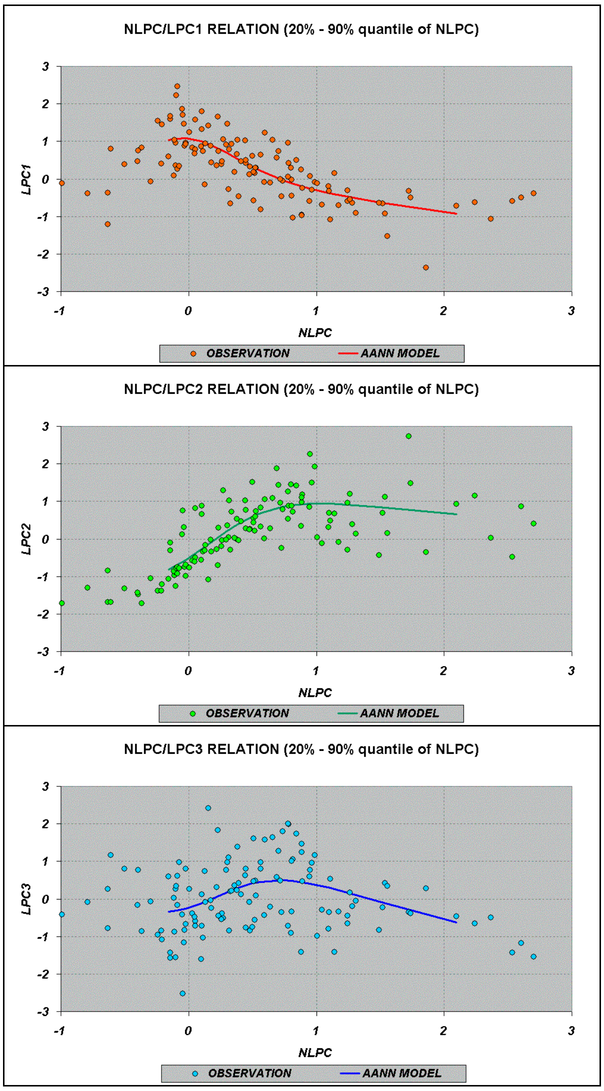

Fig.5: Relation between AANN input (original LPCA scores - vertical axis), „bottleneck" (NLPCA scores - horizontal axis) and output („reconstructed" LPCA scores - vertical axis) values. Original LPCA scores=dots, „reconstructed" LPCA scores=lines.

It is evident from the fig.5 that the relations between LPCA scores (both original and „reconstructed") and NLPCA score are non-linear Moreover, the „reconstructed" LPCA scores correspond well to the original LPCA scores which indicates that the decoding part of the network is able to reconstruct well the LPCA scores 1-3 from single characteristics (NLPCA score).

Low values of NLPCA correspond to the negative NAO phase (positive 1st LPCA mode indicates negative NAO). In this case both the 2nd and the 3rd LPCA modes are negative. Increasing NLPCA scores indicate weakening of the negative NAO and both 2nd and 3rd LPCA scores become neutral. At the levels of NLPCA score, corresponding to the neutral NAO phase, both 2nd and the 3rd LPCA modes are positive. Further increasing of the NLPCA score indicates strengthening positive NAO phase and in this case the 2nd LPCA mode remains positive while the 3rd one changes back to neutral and later to the negative one.

The evolution of the non-linear relations between LPCA modes 1-3 is evident from the fig.6. With respect to the fact that the 1st LPCA mode, corresponding well to the NAO, explains almost half of the original SLP variability, the non-linear evolution of SLP anomalies resembles the non-linear NAO (NAO non-linearly influenced by other oscillatory systems, which are present in the 2nd and the 3rd LPCA modes). The spatial patterns were calculated (with the help of „reconstructed“ LPC) for selected empirical quantiles of NLPCA score.

5. CONCLUSIONS

It was demonstrated in the paper that the 1st linear mode (the „best possible“ linear mode) of North Atlantic SLP variability does not describe the real SLP variability well, especially in the strip between New Founland and the Northern Sea. Better description of the real SLP variability is possible by the appropriate non-linear combination of 3 leading LPCA modes (combination found with the help of autoassociative neural network) into single non-linear principal component (NLPC) mode. This approach improves especially the simulation of SLP variability in the strip between the individual NAO centers where the linear description does not correspond well to the observed variability. The non-linear mode of North Atlantic SLP variability resembles the non-linear North Atlantic Oscillation and it may be described with the help of single non-linear PCA mode.

It may be seen from the results that positions of the NAO centers in positive and negative phase slightly differ. In comparison to the positive NAO phase, in negative one the northern center is shifted westward and the southern center is more pronounced over Europe. Moreover, in transition (neutral) NAO phase the region of positive SLP anomalies develops over Scandinavia (it may be connected with Scandinavian oscillation). These phenomena could not be revealed with the help of linear methods only.

6. 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.

Hsieh, W. W., and B. Tang, 1998: Applying Neural Network Models to Prediction and Data Analysis in Meteorology and Oceanography. *Bull. Amer. Meteor. Soc.,* **79,** 1856-1870.

Hsieh, W. W., 2001a: Nonlinear Principal Component Analysis by Neural Networks, *Tellus,* **53A,** 599-615.

Hsieh, W. W., 2001b: Nonlinear Canonical Correlation Analysis of the Tropical Climate Variability Using a Neural Networks Approach. *J. Climate,* **14,** 2528-2539.

Monahan, A. H., 2000: Nonlinear Principal Component Analysis by Neural Networks: Theory and Application to the Lorentz System. *J. Climate,* **13,** 821-835.

Monahan, A. H., 2001: Nonlinear Principal Component Analysis: Tropical Indo-Pacific Sea Surface Temperature and Sea Level Pressure. *J. Climate,* **14,** 219-233.

STATISTICA - Neural networks, 1998: StatSoft Inc., Tulsa, OK.

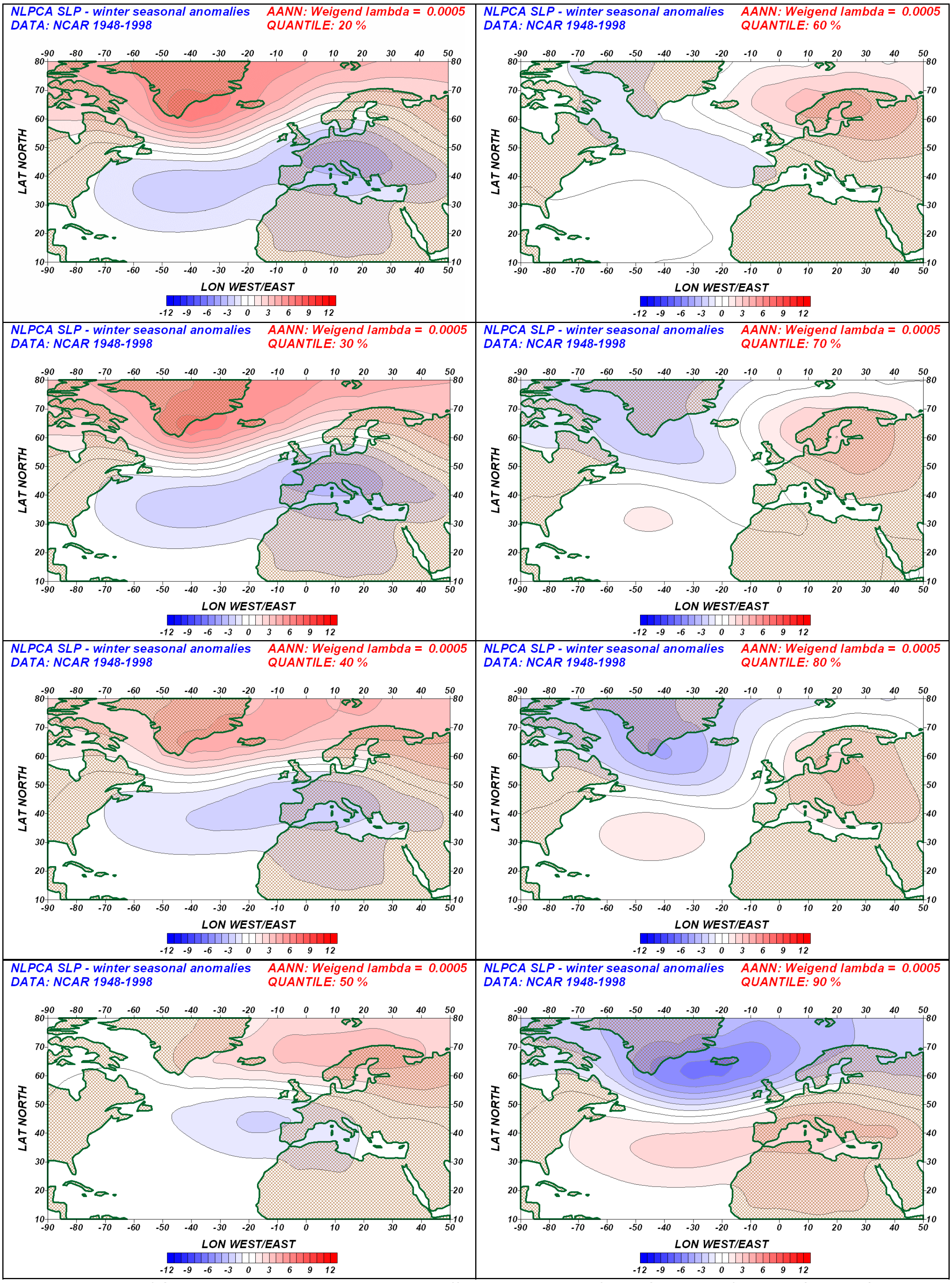

Fig.6: Patterns of SLP anomalies corresponding to different quantiles of NLPC scores (from 20% to 90% with the step of 10%)