J2.3

MAJOR INFLUENCES OF CIRCULATION PATTERNS ON TEMPERATURES IN THE ITALIAN SIDE OF THE GREATER ALPINE REGION: AN INVESTIGATION VIA NEURAL NETWORK MODELING

Antonello Pasini *, Rocco Langone

CNR, Institute of Atmospheric Pollution, Rome, Italy

1. INTRODUCTION

Attribution is a well-founded field of studies in climate science and reliable results of this activity have been obtained by means of both dynamical estimations (see IPCC 2007, for a recent review) and neural network (NN) modeling (Pasini et al. 2006). These studies aim at understanding which are the major (natural or anthropogenic) external forcings influencing the mean values of some meteo-climatic parameters, such as temperature or precipitation, at global or continental scale.

When passing at regional or local scale, however, climate natural variability can mask any direct link between global forcings and temperatures at these scales, so that the influence of circulation patterns is the key element for "attributing" climate at this spatial resolution. Of course, an important field of research is represented by studies which aim at understanding how these circulation patterns are affected by changes in global external forcings.

In any case, a regional/local correct climate reconstruction is possible only if one is able to understand how the several circulation patterns influence the main climatic variables at these scales (first step) and if Global Climate Models (GCMs) show their ability in simulating the behavior of these patterns during the last decades (second step).

In this paper we deal with the first of these two steps. In particular, by limiting ourselves to consider data about the Italian Alpine region, we try to understand by which circulation patterns we are able to reconstruct mean annual and seasonal temperatures in this region. In doing so, we apply neural network modeling as a tool which permits to achieve fully nonlinear relationships between circulation patterns and the temperature itself.

In what follows we will briefly describe the data sets (section 2) and a NN tool developed during the last years (section 3). Then, after some preliminary hints obtained in terms of bivariate linear and nonlinear analyses, the NN modeling will be applied in order to assess which combination of patterns leads to the best reconstruction results for temperatures (section 4). Finally, brief conclusions will be drawn and prospects of further study will be envisaged in the last section.

* *Corresponding author address:* Antonello Pasini, CNR, Institute of Atmospheric Pollution, via Salaria Km 29.300, I-00016 Monterotondo Stazione (Rome), Italy; e-mail: pasini@iia.cnr.it

2. DATA

Recently, a homogeneized data-base, characterized by very long time-series (≈ 200 years), has been completed and refers to the so-called "Greater Alpine Region" (Auer et al. 2007). In what follows we use the freely available data about the SW region of this Alpine data-base (see Figure 1).

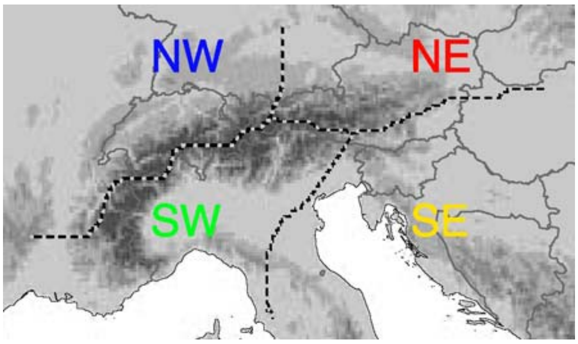

Figure 1. A division in climatic regions of the Greater Alpine Region. Data about the SW region shall be considered in the present paper.

It is worthwhile to stress that attempts at assessing the influence of large-scale atmospheric circulation on climate variability in this region have been performed by Efthymiadis et al. (2007).

In the present paper a more accurate (and fully nonlinear) analysis will be performed by consideration of data about 8 circulation patterns in the last 50 years:

- North Atlantic Oscillation (NAO),
- East Atlantic pattern (EA),
- Arctic Oscillation (AO),
- Scandinavian pattern (SCAN),
- East Atlantic/West Russian pattern (EAWR),
- Atlantic Blocking Index (ABI),
- European Blocking Index (EBI),
- El Niño Southern Oscillation (ENSO).

These data about patterns are usually synthetized by indices which can be easily used for studies of influence.

The first 5 indices are freely downloadable by www.cpc.noaa.gov, ABI and EBI data have been courtesly supplied by ARPA-SMR, Bologna, Italy, and data about ENSO have been obtained by www.cru.uea.ac.uk and then transformed in monthly anomalies in order to render them homogeneous with other data.

3. THE NN TOOL

A NN tool for both diagnostic characterization and forecast in complex systems has been developed some years ago (Pasini and Potestà 1995). Since that date it has been applied to diagnostic and prognostic problems in the boundary layer (Pasini and Potestà 1995, Pasini et al. 2001, 2003a,b, Pasini and Ameli 2003) and recently, as cited above, also to the analysis of climatic data (Pasini et al. 2006, Pasini 2009).

As far as the kernel of this NN tool is concerned, it has been extensively described elsewhere (see, for instance, Pasini et al. 2003a). Here, it is sufficient to note that the NNs adopted are feedforward and characterized by a backpropagation training endowed with gradient descent and momentum terms in the rules for weights updating. Furthermore, an early stopping method is also available. In the present study we adopt networks endowed with sigmoids in the unique hidden layer and sigmoids or linear functions in the output layer.

Together with these quite standard features (see Hertz et al. (1991) and Bishop (1995) for two reviews on these topics), this tool provides us many training facilities, useful for handling historical data from complex systems.

In particular, we try to estimate mean temperature values from indices data, but, due to the quite short time series available for circulation indices (50 years), each temperature value is estimated at a time after the exclusion of the correspondent inputs-target pattern from the training set used for fixing the connection weights. Here we use a facility of our tool, the so called "all-frame" or "leave-one-out" cross-validation procedure: it is simply sketched in Figure 2, where our total set of patterns is divided in two subsets. The white squares represent the elements (patterns) of our training set, while the gray square (one single element) represents the validation set. The relative compositions of training and validation sets change at each step of an iterative procedure of training + validation cycles. A "hole" in the complete set represents our validation set and moves across this total set of patterns, thus permitting the estimation of all temperature values at the end of the procedure.

In this paper we adopt the "all-frame" procedure just described, for both the neural model and the multi-linear regression.

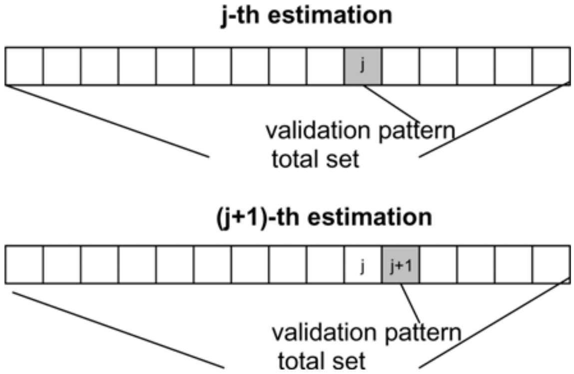

Figure 2. A sketch of the "all-frame" or "leave-one-out" cross-validation procedure.

4. MAIN RESULTS

The first preliminary step in our investigation is to assess which circulation patterns may be the most influencing on temperature, in a bivariate manner. Thus, we perform a bivariate analysis in linear and nonlinear terms, *via* calculation of the Pearson coefficient R and of its nonlinear analogue Rnl, the so-called Correlation Ratio (see Marzban et al. 1999, for its formula and calculation).

Rnl allows us to understand if some indices, which do not show high linear correlation with temperature, can however be influent on it by means of more complex and nonlinear relationships.

PERIOD	CORRELATION T vs INDEX	NAO	EA	SCAN	EAWR	ABI	EBI	AO	ENSO
WINTER	R	**0.345**	0.240	0.025	**0.398**	-0.269	0.101	**0.339**	0.020
	Rnl	[0.357, 0.490]	[0.409, 0.506]	0.372	[0.441, 0.457]	-0.460	0.559	[0.357,0.470]	0.482
SPRING	R	0.236	**0.514**	-0.275	0.231	-0.336	-0.134	**0.418**	-0.082
	Rnl	0.441	[0.481,0.648]	-0.324	0.549	-0.359	[-0.377,-0.587]	0.477	[-0.514, -0.534]
SUMMER	R	0.112	**0.470**	**-0.477**	-0.069	-0.042	**-0.409**	0.269	0.021
	Rnl	[0.372, 0.551]	0.558	-0.547	[-0.619, -0.620]	[-0.359, -0.397]	[-0.459, -0.498]	0.482	0.323
AUTUMN	R	0.029	**0.539**	-0.246	0.054	-0.071	-0.045	**0.340**	-0.283
	Rnl	0.393	[0.742, 0.801]	[-0.495, -0.507]	[0.479, 0.683]	[-0.369, -0.389]	[-0.558, -0.633]	[0.432, 0.628]	[0.407, 0.461]
EXTENDED WINTER	R	**0.600**	**0.556**	-0.188	0.108	**-0.513**	**-0.401**	**0.494**	-0.097
	Rnl	[0.715, 0.774]	[0.703, 0.765]	[-0.330, -0.449]	[0.420, 0.542]	[-0.694, -0.597]	[-0.562, -0.634]	[0.607, 0.635]	[-0.282, -0.398]
HOT SEMESTER	R	0.036	**0.671**	**-0.386**	-0.047	-0.072	**-0.475**	**0.323**	0.012
	Rnl	[0.259, 0.453]	0.741	-0.425	-0.301	[-0.534, - 0.564]	-0.503	0.539	0.350
COLD SEMESTER	R	**0.379**	**0.332**	-0.112	0.260	-0.217	-0.004	**0.318**	-0.008
	Rnl	0.538	0.564	-0.350	0.334	-0.366	[-0.2612, -0.38]	0.358	-0.152
MEAN OF THE YEAR	R	**0.366**	**0.695**	**-0.337**	0.166	**-0.362**	**-0.366**	**0.472**	0.201
	Rnl	[0.497, 0.544]	0.742	[-0.447, -0.548]	[0.342, 0.510]	[-0.391, -0.534]	[-0.502, -0.586]	[0.554, 0.567]	[0.451, 0.628]

Table 1. Calculation of R and Rnl (indices vs. T) for several periods. Extended winter = December to March. Bolded values indicates linear correlations which are founded significant with a two-tails Student test. Rnl often shows an interval of values for including several histograms which are compatible with its calulation method.

PERIOD	CORRELATION T vs ENSO	SPRING	SUMMER	AUTUMN	WINTER
SPRING	R	-0.082	-0.106	-0.099	-0.083
	Rnl	[-0.514, -0.534]	[-0.353, -0.458]	[-0.247, -0.309]	[-0.3321, -0.4385]
SUMMER	R	-0.091	0.021	**-0.359**	-0.191
	Rnl	[-0.473, -0.60]	0.323	-0.413	[-0.224, -0.230]
AUTUMN	R	**-0.395**	-0.283	-0.283	-0.179
	Rnl	[-0.509, -0.618]	[-0.618, -0.711]	[0.407, 0.461]	[-0.460, -0.517]
WINTER	R	-0.165	-0.011	0.087	0.020
	Rnl	[-0.584, -0.693]	[-0.475, -0.575]	[0.420, 0.452]	0.482

Table 2. Linear and nonlinear bivariate cross-correlations (ENSO vs. T).

Some considerations from Tables 1 and 2:

- both AO and EA are correlated significantly in almost all seasons, while NAO seems to be important just in winter and extended winter;
- sometimes, indices which show low linear correlation are endowed with a higher correlation in the calculation of Rnl: this may induce us to consider these variables as important in a fully nonlinear analysis of influence on T;
- the blocking indices show interesting anticorrelations during the extended winter;
- ENSO shows negligible correlations in the same season, but significant anti-cross-correlations.

After this preliminary bivariate analysis, our aim is to try to reconstruct the behavior of T in various periods by means of combinations of these indices.

In doing so, due to the limited number of data at disposition, we cannot build a network including all the indices as inputs, because we should fall into overfitting conditions. Thus, just few indices are considered as inputs (exactly, combinations of 3 or 4 indices) and a maximum of 4 neurons are inserted in a single hidden layer.

In what follows we show just the best results obtained in this attempt. These results (calculated on the validation set) are summarized in Tables 3 and 4, where the error bars related to the NN performance come from ensemble runs of the networks starting from different random intial weights, so that the network itself is able to widely explore the landscape of the cost function in this local backpropagation training method. The interval indicates $\pm$ 2 standard deviations.

PERIOD	INPUT (GENERAL ATMOSPHERIC CIRCULATION PATTERNS)	NEURAL NETWORK PERFORMANCES	MULTIVARIATE LEAVE-ONE-OUT LINEAR CORRELATION
SPRING	EA, AO, SCAN	**0.503 ± 0.005**	0.466
SUMMER	EA, AO, EBI	**0.579 ± 0.042**	0.513
AUTUMN	EA, AO, ENSO (SPRING MEAN)	0.657 ± 0.015	0.671
WINTER	EBI, AO, EAWR	**0.424 ± 0.026**	0.313
EXTENDED WINTER	EA, AO, EBI	0.711 ± 0.023	0.693
YEAR	EA, AO, ENSO	**0.749 ± 0.029**	0.696

Table 3. Results of neural and linear reconstructions starting from combinations of 3 circulation patterns (bolded values indicate significant improvements of NN performance if compared with linear one).

PERIOD	INPUT (GENERAL ATMOSPHERIC CIRCULATION PATTERNS)	NEURAL NETWORK PERFORMANCES	MULTIVARIATE LEAVE-ONE-OUT LINEAR CORRELATION
SPRING	EA,AO,SCAN,EAWR	**0.473 ± 0.012**	0.460
SUMMER	EA, EBI, AO, SCAN	**0.566 ± 0.065**	0.471
AUTUMN	EA, AO, SCAN, ENSO (SPRING MEAN)	0.651 ± 0.028	0.653
WINTER	EBI, AO, EAWR, ABI	**0.407± 0.076**	0.273
EXTENDED WINTER	AO, EA, EBI, SCAN	**0.729 ± 0.022**	0.683
YEAR	EA, AO, ENSO, EBI	**0.750 ± 0.023**	0.708

Table 4. As in Table 3, but for combinations of 4 circulation patterns.

As just shown from bivariate analysis, even the NN reconstruction shows that the influence of AO and EA are very important almost in all seasons. Now, the nonlinear reconstructions from 4 indices appear better than linear ones in all cases, except autumn. In any case a major percentage of variance is explained by our NN model.

It is worthwhile to note that, in order to achieve the best results, networks with sigmoids in the output layer must be chosen for intermediate seasons (spring and autumn), while a linear transfer function must be considered for other periods. This fact suggests the existence of stronger nonlinear relationships between indices and temperature in intermediate - more complex - seasons, if compared with other periods.

Finally, Figure 3 shows the time series of T on extended winters and its NN good reconstruction.

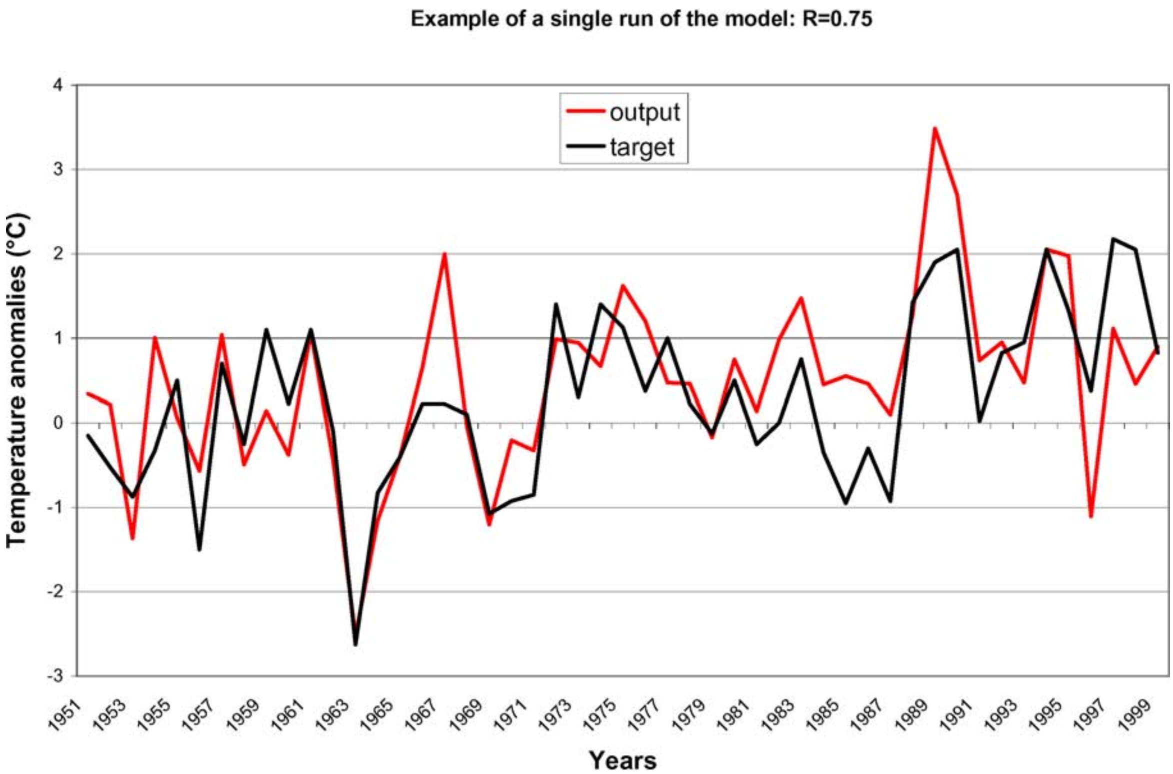

Figure 3. Perfomance of temperature reconstruction on extended winters by a single run of NN modeling.

From this Figure, we can appreciate how much our analysis is able to catch the interannual variability of T. Furthermore, it is evident the lack of a consistent bias in the reconstruction, while, even in cases of similar values for Pearson coefficient in Tables 3 and 4, linear reconstructions often show a large value of bias.

Further analysis may be performed in terms of some indices of performance built on thresholds and contingency tables, such as POD, FAR, HSS. This analysis, not shown here, confirms the superiority of our NN method (with respect to multilinear regression), as far as the reconstruction of time series of T is concerned.

5. CONCLUSIONS AND PROSPECTS

In short, our research, even if quite preliminary, contributes to understand how circulation patterns influence temperature on the SW portion of the Greater Alpine Region. Of course, this investigation should be extended to other variables, such as precipitation, if we would like to achieve a more complete characterization of climate variability in this region: this is a concrete prospect of further study.

Thus, this kind of studies is very useful for establishing major influences on regional temperatures and, therefore, makes available a crucial information for a downscaling activity, too. In fact, in doing so, we have to choose those GCMs which are able to reconstruct the behavior of these particular circulation patterns. Only in this manner we could achieve a correct downscaling for successfully reconstructing climate at regional level and possibly supplying reliable future scenarios at this scale.

REFERENCES

Auer, I. et al., 2007: HISTALP – historical instrumental climatological surface time series of the Greater Alpine Region, *Int. J. Clim.*, **27**, 17-46.

Bishop, C.M., 1995: *Neural networks for pattern recognition*, Oxford University Press, 482 pp.

Efthymiadis, D., P.D. Jones, K.R. Briffa, R. Böhm, and M. Maugeri, 2007: Influence of large-scale atmospheric circulation on climate variability in the Greater Alpine Region of Europe, *J. Geophys. Res.*, **112**, D12104.

Hertz, J., A. Krogh, and R.G. Palmer, 1991: *Introduction to the theory of neural computation*, Addison-Wesley, 327 pp.

IPCC, 2007: *Climate Change 2007: The Physical Science Basis*, Cambridge University Press, 996 pp.

Marzban, C., E.D. Mitchell, and G.J. Stumpf, 1999: On the notion of "best predictors": An application to tornado prediction, *Weather For.*, **14**, 1007-1016.

Pasini, A., 2009: Neural network modeling in climate change studies, in *Artificial intelligence methods in the environmental sciences* (Haupt, S.E., A. Pasini, and C. Marzban eds.), Springer, pp. 235-254.

Pasini, A., and F. Ameli, 2003: Radon short range forecasting through time series preprocessing and neural network modeling, *Geophys. Res. Lett.*, **30 (7)**, 1386, doi:10.1029/2002GL016726.

Pasini, A., F. Ameli, and M. Lorè, 2003a: Mixing height short range forecasting through neural network modeling applied to radon and meteorological data, *Proc. 83rd AMS annual meeting (Third Conference on artificial intelligence applications to the environmental science)*, Long Beach, CA, AMS, CD-ROM, paper 3.5.

Pasini, A., M. Lorè, and F. Ameli, 2006: Neural network modelling for the analysis of forcings/temperatures relationships at different scales in the climate system, *Ecol. Model.*, **191**, 58-67.

Pasini, A., V. Pelino, and S. Potestà, 2001: A neural network model for visibility nowcasting from surface observations: Results and sensitivity to physical input variables, *J. Geophys. Res.*, **106 (D14)**, 14,951-14,959.

Pasini, A., C. Perrino, and A. Žujić, 2003b: Non-linear atmospheric stability indices by neural-network modelling, *Nuovo Cimento*, **26C**, 633-638.

Pasini, A., and S. Potestà, 1995: Short-range visibility forecast by means of neural-network modelling: a case study, *Nuovo Cimento*, **18C**, 505-516.