STARDEX

STAtistical and Regional dynamical Downscaling of EXtremes for European regions

EVK2-CT-2001-00115

Deliverable D10

Recommendations on the best predictor variables for extreme events

D10 AUTHORS

Authors:

Carlo Cacciamani, ARPA-SMR
Valentina Pavan, ARPA-SMR
Rodica Tomozeiu, ARPA-SMR

Date: 5 May 2003

CONTENTS

1. Introduction

This study investigates the link between atmospheric circulation and extreme events of winter precipitation in order to identify some potential predictors of extremes. This connection could be studied in different ways. One way is to find statistical relationship between regional climate variables (extreme precipitation) and some ready defined circulation indices (NAO, blocking index,…) or circulation types. Another way is to study the connection between the objectively defined patterns of the large-scale circulation and regional climate variability using multivariate techniques like the canonical correlation analysis (CCA). The CCA method has two important advantages: it supplies the optimum correlated couples patterns of predictor and predictand and the time series associated to them allow the analysis of the temporal variability of these configurations.
Further, the results of the CCA analysis was used in a "downscaling" model in order to estimate the local climatic anomalies from the large scale ones (Werner and von Storch,1993, Busuioc et al,1998). The model has been constructed for the period 1958-1980 and validated for the period 1981-2000. The skill of the model is defined by the correlation coefficient between the observed and estimated values as well as by the variance explained by the estimated values from the total observed variance. The model is considered skilful if the correlation coefficient is statistically significant.

2. Data base and methodology

The data set used in this work is referred only at the winter (DJF) season and it consists of:

- the time series of the number of events greater than 90th percentile of raindays (*692 index-predictand*) for each year. The index was computed for 44 stations uniformly distributed over Emilia-Romagna, using observed daily precipitation from the interval 1958-2000;
- monthly sea level pressure (*SLP- "possible" 'predictor*) which covers the area between 30°W-30°E and 30°N-60°N, at resolution 2.5°x2.5° (NCEP reanalysis) ;
- monthly mean of specific humidity at the levels 1000mb, 950mb, 850mb, 750mb(*"possible predictor"*) for the area 45°N, 7.5°E-12.5°E, NCEP reanalysis, at resolution 2.5°x2.5°(3 grid points). The area was selected such as to cover our test region, Emilia- Romagna. An average at all levels of specific humidity was computed.

The main physical mechanisms which control the climate variability of the frequency of events greater than 90th percentile (*692 index*) are identified through the CCA analysis. CCA selects pairs of spatial patterns of two space-time –dependent variables (large scale predictors and regional scale predictand), so that their time components are optimally correlated. The first CCA pair gives the maximum correlation between the two parameters, followed by the second CCA pair, and so on. Prior to the CCA, the original data are projected onto their EOFs to eliminate noise and to reduce the dimensionality of the data set .
Thus, the steps of our analysis are:

1. *EOF analysis* of the observed station data – was performed for each parameter, separately, in order to filter the data. The result of this analysis provides information about the main characteristics of spatial and temporal variability. A *cluster analysis*

technique is also applied to the predictand from Emilia-Romagna (winter 692 index) in order to divide the region into 5 homogeneous areas (figure 1). For each area it was then computed a mean of 692 winter index, obtaining in this way 5 time series which represents our predictand from Emilia- Romagna.

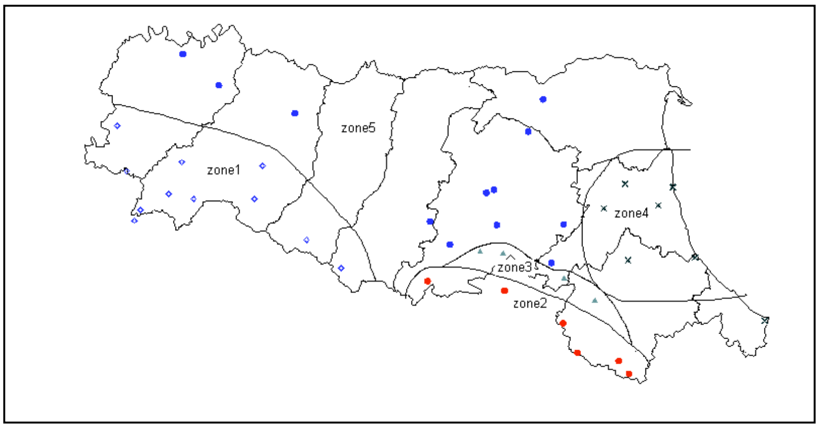


Figure1 Clusters of 692 winter index

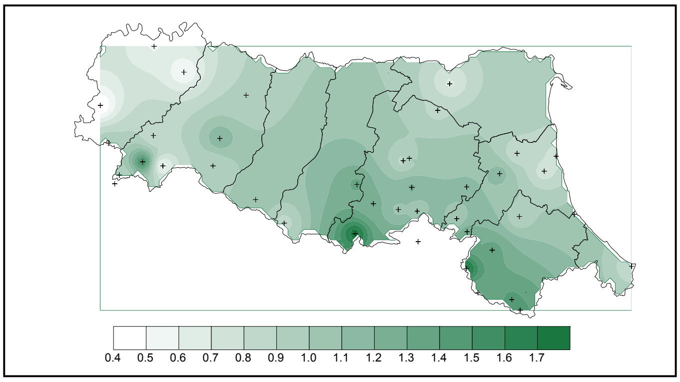
The EOF analysis for 692 index was performed firstly for 1958-2000, and the time series of the principal components associated with the most important patterns (first 3 PCs) were analysed from the trend point of view. Taking into account that the downscaling model was constructed for the period 1958-1980, the EOF analysis was then performed for this period for both, predictors and predictand.

2. *CCA analysis*- has been performed using input data the EOFs of 692 index and SLP. Similar analysis is made using as predictors specific humidity, averaged over all levels, together with SLP.

3. Finally, *a* subset of CCA pairs is used in a regression model in order to estimate the local extreme events of precipitation, represented by 692 index (*predictand*) from the large scale SLP or SLP combined with specific humitity(*predictor*). The skill of the model is dependent on the number of EOFs retained for the CCA and the number of CCA components used in the regression model. Various alternatives are tested to select the number of EOF and CCA time components. The model was constructed for the period 1958-1980 and validation for the period 1981-2000.The linear trend in the original data was removed prior to the statistical downscaling

3. Results

The EOF analysis performed for the 692 winter index over the period 1958-2000 (44 stations) reveals the following characteristics:
- EOF1 (Figure2a), which explains 39% of the total variance, presents a monopole structure with higher variability in the mountain area. The time series associated to the EOF1(Figure2b) reveals the presence of a negative trend.



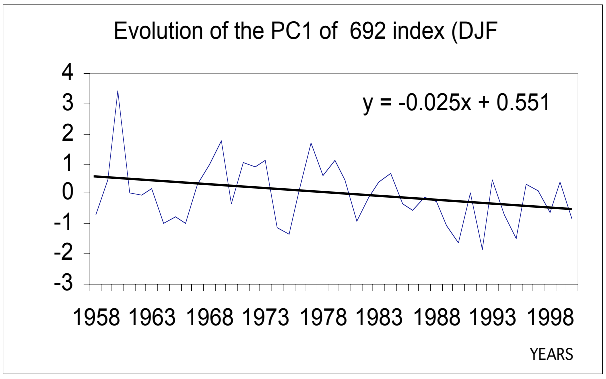


Figure 2a Configuration of the EOF1-692 winter index

Figure 2b

- EOF2 (not shown), which explains 18% of the total variance, presents a dipole structure, revealing the influence of the mountains. No significant trend is present in the PC2 time series. The influence of hill and sea appears in the configuration of EOF3 and EOF4(not shown).
Taking into account the results presented above for these stations a cluster analysis has been performed a using complete linkage method. As result of this analysis, Emilia-Romagna was divided into 5 separate areas (figure 1).For each area we compute the mean index, these time series being our predictands.

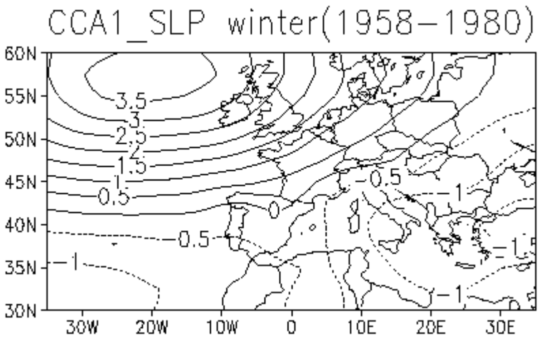
A similar EOF analysis was performed for 692 index and winter SLP, but for the period 1958-1980, which represent the period for which it has been applied the CCA and then constructed the regression model. The number of events greater than 90th percentile of raindays reveals patterns similar with those presented above. The variance explained for each parameter is presented in table 1.

Table 1 Variance(%) explaind by the first EOFs of 692 and SLP (DJF-1958-1980)

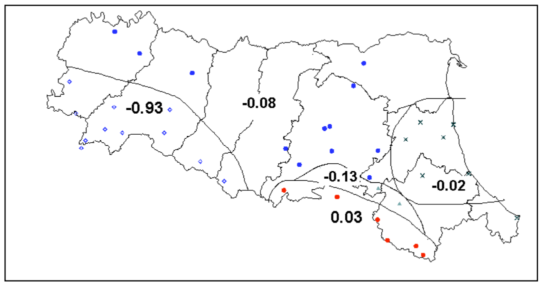
Parameter	EOF1(%)	EOF2(%)	EOF3(%)	EOF4(%)	EOF5(%)	EOF6(%)	Total(%
692 index	62	*28*	*6*	*3*	*1*	-	*100*
SLP	*43*	*28*	*15*	*7*	*2*	*1*	*96*

The EOF patterns of the 692 index and SLP were then used in the CCA analysis using different numbers of EOFs. The number of EOFs retained for the CCA has been chosen such that the addition of more EOFs did not change the CCA patterns and their correlations substantially. The best combination is formed by the first 5EOFs of 692 index and first 6 EOFs of SLP.
The first CCA pattern, which have a correlation between the 692 index and SLP time series equal to 0.85, is associated with a north-easterly circulation, advecting cold and dry continental air over Emilia-Romagna which implies a decrease in the 692 index (figure 3a and 3b), affecting less the sea area.

Figure 3 The spatial configuration of the first CCA pairs.



CCA1 692

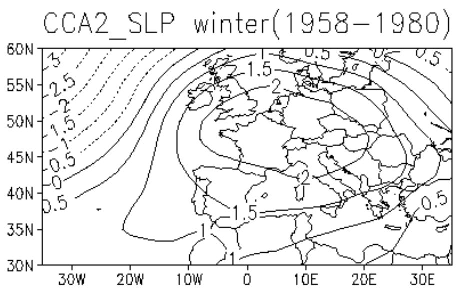


a) var.=20%

b) var.=8%

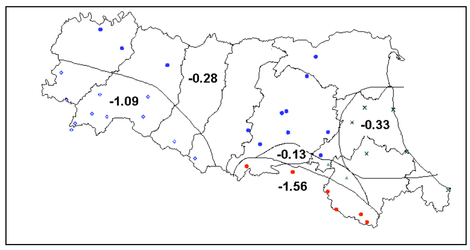
The second CCA pair, with correlation equal to 0.77, is associated with the occurrence of a blocking centred over continental Europe (figure 4a and 4b) connected with a decrease in the frequency of extremes over Emilia-Romagna(figure 4b).
The above patterns seems to be the mainly mechanism that control the distribution of the frequency of the extremes over the region.

a)



b)

692 index



Var.=15%

Var.=33%

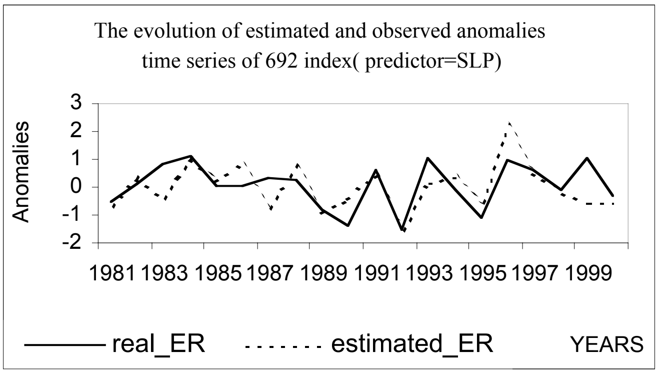
Figure 4 The spatial configuration of the second CCA pairs.

The time series associated to the first 5 CCA pairs of SLP and 692 index are then used in a multivariate regressive model in order to estimate the local climatic anomalies for the period 1981-2000. In order to identify the model with the best performance different combinations of CCA were tested. The model has been constructed and validated for each zone (figure 1) and for the whole region. Table 2 (second and third column) describes the skill of the model with different combinations.

Table 2 *Skill of the statistical downscaling model expressed as correlation coefficient between the observed and estimated anomalies (bold-5% significance)*

Zone	SLP (5CCA)	SLP(4CCA)
1.	**0.45**	0.22
2.	**0.45**	0.36
3.	0.13	0.15
4.	0.34	**0.45**
5.	**0.37**	**0.41**
Emila –Romagna	**0.59**	**0.58**

The results presented in table 2 reveals a good performance of the model for the combination of 5 CCAs of SLP. The skill is significant over the mountain, plain and hill area (zone 1,2,5. Good results are also obtained for the whole region (r= 0.59).Figure 5 presents the evolution of the frequency of extreme events over Emilia-Romagna, estimated and observed data, in the period 1981-2000.



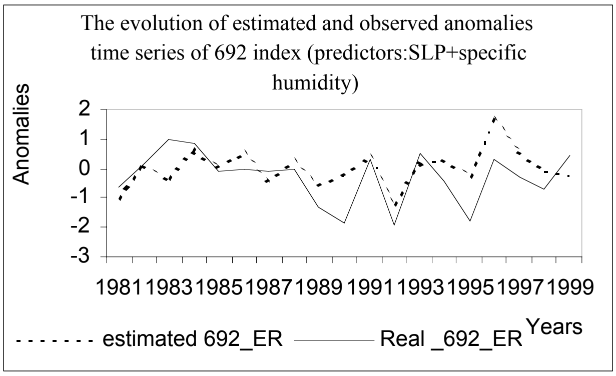


Figure 5 r=0.59 r=0.52

A similar analysis has been started using winter specific humidity combined with SLP, good performance being obtained for all Emilia-Romagna(figure 5b) but poorer performance for the single zones.

4. Conclusion

The analysis developed for winter season reveals that sea level pressure and specific humidity could be potential predictors of the frequency of events greater than 90^{th} percentile. An extension of this study to the other seasons will be done in the near future.

5. References

Barnett T. P., Preisendorfer R., 1987, Origins and Levels of Monthly and Seasonal Forecast Skill for United States Surface Air Temperatures Determined by Canonical Correlation Analysis, American Meteorological Society, **115,** 1825-1850.