Exploring the climate-flood link by analyzing the underlying common climatic (spatio-temporal) structure in extreme floods in Germany

F. Cioffi[1], F. Conticello[1], B. Merz[2], U. Lall[3]

[1] University of Rome "La Sapienza" DICEA - Rome, Italy

[2] Helmholtz Centre Potsdam, GFZ German Research Centre for Geosciences, Section Hydrology, Potsdam, Germany

[3] Department of Earth and Environmental Engineering, Columbia University, New York, NY, USA

European Geosciences Union
General Assembly 2015
Vienna | Austria | 12 – 17 April 2015

Summary

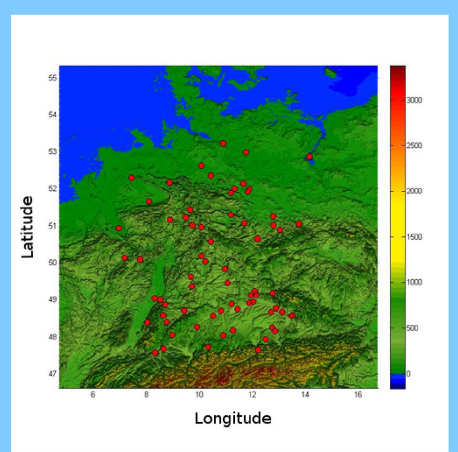
The existence of common spatial-temporal structures in extreme floods in Germany, as well as, how floods are modulated by specific meteorological teleconnections, was investigated. Linear and logistic principal component analyses (PCA), and coherence wavelet analysis were applied to investigate the spatio-temporal structure of extreme hydrological events and their link with atmospheric circulation patterns. Daily streamflow data from 68 stations (1950-2005) in Germany – made available by GFZ - are used. The influence of atmospheric circulation patterns, and related teleconnections, was examined through the structure of geopotential height at 1000 and 700 hPa for a region within 90°W-60°E, 20°-80° (reanalysis from NCEP-NCAR model). Three groups of streamflow stations with spatio-temporal structure are identified. Extreme floods, defined as those with flowrate over the 99th percentile threshold, are significantly modulated in frequency by geopotential patterns identified. These are shown to be related to the main climate indices generally used to characterize the Central Europe climate.

Geographic context and data

Daily Stramflow Data: 68 station in Germany from 1/1/1950 to 12/31/2005. Source: GFZ.

Monthly climate indexes: North Atlantic Oscillation (NAO); Arctic Oscillation (AO); East Atlantic (EA); East Atlantic/Western Russia (EA/WR); Scandinavia (SCAND); Polar / Eurasia (POL), El Nino Southern Oscillation (ENSO).

Daily Geopotential Height fields, at different hPa from 1000 to 500, bounded from 90W to 60E and from 20N to 80N (Source: NOAA: Climate Data Assimilation System I; NCEP-NCAR Reanalysis Project), have been deseasonalized by subtracting the interannual mean value of a date and by dividing by the interannual variance .

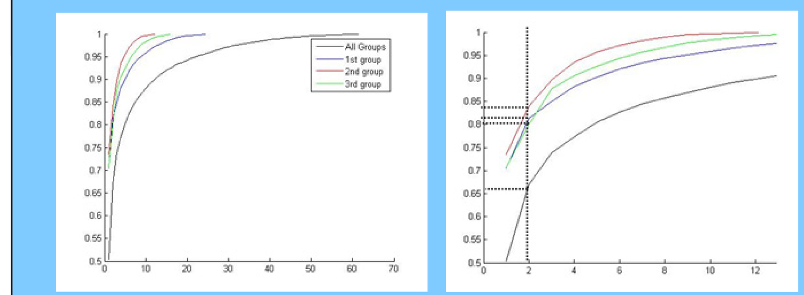


Physical Map of the Germany with red dots representing the position of each station analyzed in this study

Methods

Cluster analysis

The spatial coherence of daily streamflows of the different stations is explored following an approach similar to the one proposed by Cioffi et al. 2014.
The clustering procedure consists of the following steps:
1. PCA on daily streamflow dataset (including all the stations) is carried out ;
2. the leading n principal components (PCs) accounting the most percentage of explained variance are taken into consideration;
3. correlation matrices between each PC and daily streamflow time series are calculated; negative values are supposed to be time shifted so they are discarded from the cluster;
4. Stations with high correlation with a PC are then defined as a candidate cluster
5. For each such cluster compute a PC and keep 1st 2 PCs from this group
6. when the leading first two PCs of all the candidate groups, has an explained variance greater 80% and the 90% of stations are included in the ensemble of the groups, the clustering procedure stops; otherwise the number of PCs is increased until the two above said conditions are satisfied.



Explained Variance comparison. Black color represents the explained variance for all the groups; then 1st, 2nd and 3rd groups are represented in blue, red and green. Figure on the left shows an extended view for all the PC dimensions. Figure on the right focuses the first 12 PCs dimensions and return explain variance values for the first two PCs.

Definition of extreme floods

Extreme Flood = Daily Flow > Threshold

Threshold =Median over all years(99th percentile of daily flow for each year).

Then, binary time series f_{it} are obtained by the daily streamflow ones :

$$f_{it} = I(P_{it} > P_i^*)$$

where P_{it} is the value of daily flowrate at the station I at the time t, P_i^* is the threshold value of the i station, and I is an indicator function that takes the value 1 if the argument is true and 0 otherwise.

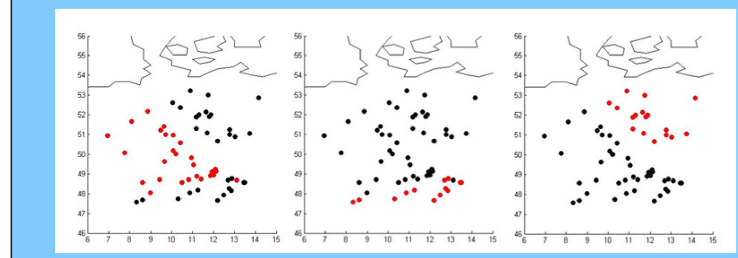
Frequency Analysis of extreme floods and linkage with geopotential fields

Dimensionality reduction (Geopotential Data and binary extreme flow time series) is carried out, then continuous and coherent wavelet transformations of the dimensionally reduced time series is performed.

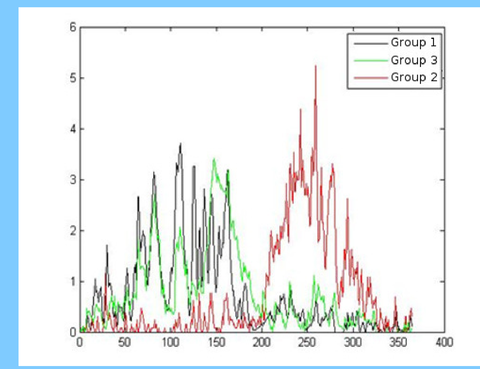
To reduce the binary extreme flood time series to few time series the approach suggested by Schein et al 2003 is used. Such approach, known as Logistic Principal Component Analysis, is a specific case of generalized criteria for dimensionality reduction by appealing to properties of distributions in the exponential family (Collins et al. 2001). In particular, the model exploits the log-odds as the natural parameter of the Bernoulli distribution and the logistic function as its canonical link.

Results

a) Cluster Analysis



The three clusters of figure shows a very clear spatial and temporal structure



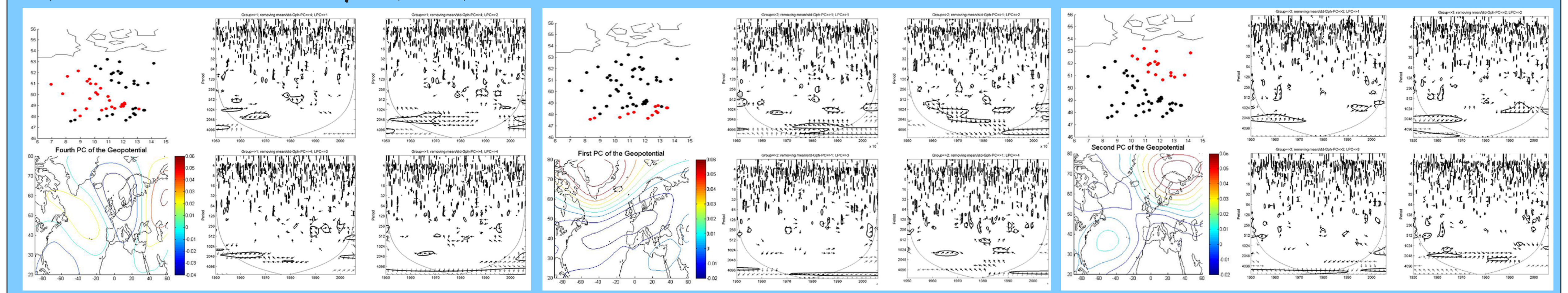
The clustering obtained by using daily streamflows holds also for extremes, because we are able to identify very distinct seasonal occurrence of extreme floods for each group.

Correlation between Geopotential PCs and Teleconnection Indexes

	AO	EA	EA/WR	NAO	PNA	POL	SCA	WP	ENSO
PC1	-0.759	-0.090	-0.003	-0.807	0.0535	-0.194	0.0683	-0.036	-0.091
PC2	-0.281	-0.191	-0.101	0.05	-0.05	-0.209	0.673	-0.011	-0.029
PC3	-0.066	0.376	-0.248	-0.041	0.088	-0.055	5e-06	-0.047	0.058
PC4	-0.185	-0.187	-0.671	-0.071	0.022	-0.120	0.003	0.019	0.010

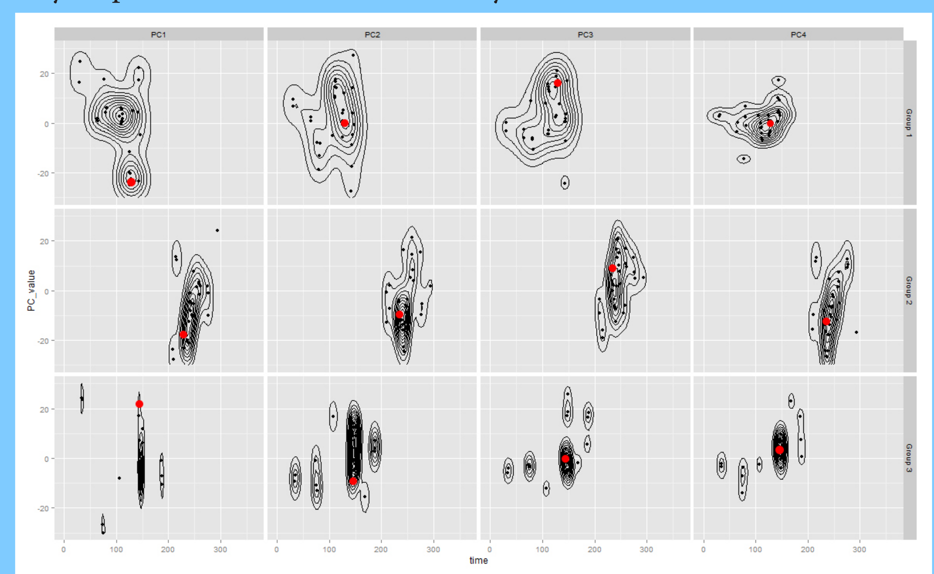
Correlation has been performed between first 4 PCs and the principal Northern Hemisphere Teleconnection Indexes: Arctic Oscillation (AO), East Atlantic Pattern (EA), East Atlantic/West Russia Pattern (EA/WR), North Atlantic Oscillation (NAO), Pacific/ North American Pattern (PNA), Polar/ Eurasia Pattern (POL), Scandinavia Pattern (SCA), West Pacific Pattern (WP), El Nino South Oscillation (ENSO). Highest correlation are shown in red.

b) Coherence Wavelet Analysis (WTC)

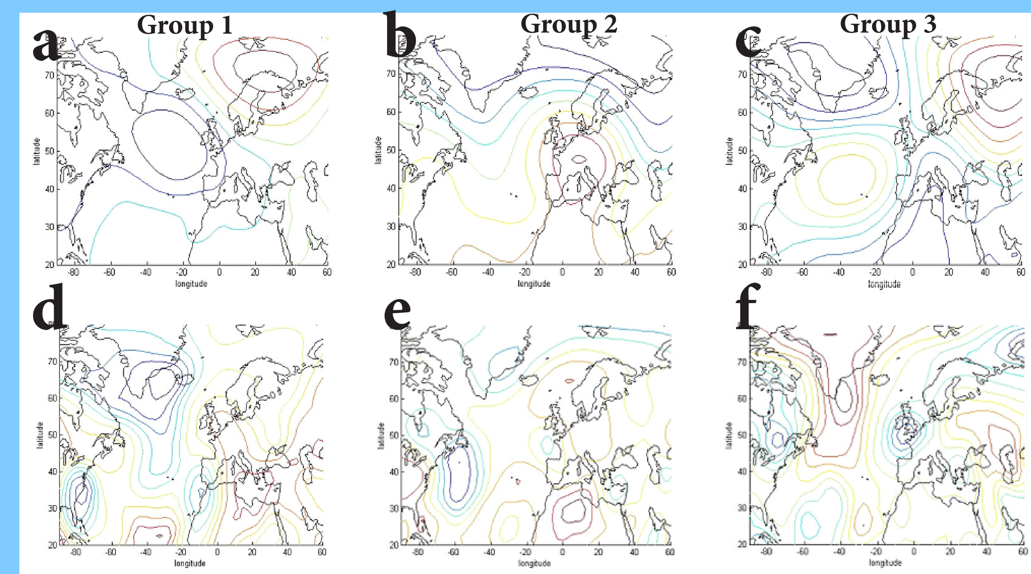


In the left side of the figure are resumed stations of the group (red) and the Geopotential PC taken into account. In the right side of the figure Wavelet Coherence Analysis Results are shown (a – LPC1; b– LPC2; c – LPC3; d – LPC4), high (low) frequency domain is in the upper (lower) part, black contour lines enclose peaks greater than 95% of confidence, whitened areas exclude a domain by the cone of influence of the wavelet analysis (Torrence and Compo 1998), arrows represent the relative phasing of two time series (right: in phase; left: anti-phase; down).

WTC scalograms show that significant locally phase-locked correlations between the PCs and LPCs time series can be recognized. Low frequencies in the range (10 to 15 year period) modulate extreme floods on the entire time interval, while intermittent correlations occur for minor frequencies. Looking at the different frequency ranges we can infer that a) for the group 1, extreme floods in the frequency range of 3-6 and 7-10 year period are associated with PC4 , in the frequency range of 11-15 year period they are mainly associated with PC1; b) for the group 2, in the frequency range of 3-6 year period extreme floods are mainly associated with PC1; in the frequency range of 7-10 they are modulated by PC2, in the frequency range of 11-15 year period they are modulated by PC3; c)for the group 3, in the frequency range 3-6 year period floods are modulated by PC3, in the frequency range 7-10 year period they are modulated by PC1, and finally in the frequency range 11-15 year period floods are modulated by PC1 and PC2



Joint probability density between PC value and day of the year for extreme floods events (> 99 th percentile) occurring simultaneously for more the 70% of the stations of each group. Red dots refer to the extreme event that involves the higher number of stations.



a), b), c) geopotential fields reconstructed by the values of the four PCs corresponding to the maximum joint probability (see figure on the left). d), e), f) geopontial fields obtained by the PCs corresponding to the red dots of figure on the left.

Acknowledgments

This work was supported by a grant from the DAAD - Deutscher Akademischer Austauschdienst at German Research Center for Geosciences, Section Hydrology in Potsdam.

References

Cioffi et al., 2014, Space-time structure of extreme precipitation in Europe over the last century, International Journal of Climatology
Torrence, C., and G. P. Compo. 1998. A practical guide to wavelet analysis. Bulletin of the American Meteorological Society 79:61–78.
Collins, M., Dasgupta, S., & Schapire, R. (2001). A generalization of principal component analysis to the exponential family. NIPS2001
A. Schein, L. Saul, and L. Ungar, "A generalized linear model for principal component analysis of binary data," in Proceedings of the 9th International workshop on artificial Intelligence and Statistics, Key West, FL, January 2003

Conclusion

Analysis showed the existence of three different groups of stations that share similar space-temporal structure both in daily streamflows and in extreme floods (over the 99th percentile threshold). Extreme floods of each group of streamflow stations are significantly modulated in frequency by different 700 hPa geopotential patterns associated to few PCs and related EOFs.