On the role of experimental evidences in the description of the distribution of annual rainfall maxima

A. Parodi & G.Boni

Centro di Ricerca in Monitoraggio Ambientale, University of Genova, Cadorna, 7 - 17100 Savona, Italy. E-mail: boni@cima.unige.it, antonio@cima.unige.it

ABSTRACT

Any regional rainfall frequency analysis needs for the definition of a theoretical frequency distribution that can be assumed as parent distribution for the analyzed variable. In Italy many investigations have been done in order to identify the best form of the parent distribution in terms of capabilities of reproducing the statistical properties of the observed time series. The TCEV distribution applied to Italian hydrological data has been demonstrated to be able to take into account both for the observed presence of outliers and high variability of the skewness. Furthermore its parameters have a clear physical meaning. Despite of this good properties of TCEV no satisfactory explanation of the physical reasons that produces the presence of two components in observed rainfall distribution has been given so far. The aim of this work is to provide a phenomenological validation of the experimental evidences on the observed data. The experimental evidence of the Mediterranean storm characteristics provided by remote sensing can give the elements for this validation. It is shown that using a simple model for storm simulation to generate synthetic time series, the different characteristics of the two kind of storms, especially in terms of advection speed, can produce the two components in time series frequency distribution and reproduce some other global characteristics of the time series.

1 INTRODUCTION

Regional frequency analysis of the hydrological extremes requires the identification of the probabilistic model able to describe the statistical behavior of the variable under analysis and the regional model for the description of the spatial variability of the distribution parameters. There are few theoretical studies that provide some indication for the choice of a particular frequency distribution or regional model rather than one another (WMO, 1981).

For the case of the frequency distribution, it must have an adequate theoretical basis and it must be able to explain the principal statistic properties of the real processes (Versace et al., 1989).

Many probabilistic models based on the theory of the extreme value, such as GEV (Jenkinson, 1955) and TCEV (Rossi et al, 1984) distributions, have been proposed.

The TCEV distribution has demonstrated to have good performance both in reproducing the presence of outliers and the variability of the higher order moments (coefficient of variation and skewness) observed in the italian annual flood series (AFS) and rainfall maxima (ARM) (Rossi et al, 1984; Arnell & Beran, 1984). The distribution showed the capability to provide unbiased fit to other parents such as Wakeby parents and performs better than other distributions (GEV,Wakeby) in the sense that its worst efforts are less bad than those of the other distributions, when the sample size is sufficiently extended (Arnell & Gabriele, 1988).

Furthermore the clear physical meaning of the TCEV parameters provides a greater confidence on the statistically obtained results. For this reason the TCEV distribution has been applied successfully in the VAPI (flood evaluation in italy) Project for the regional flood frequency analysis within different italian regions (Versace et al., 1989; Copertino & Fiorentino, 1992; Cannarozzo et al., 1993; Rossi & Villani, 1994).

Despite of many proofs of the goodness of the TCEV in fitting the observed statistical behavior of hydrological data, no analysis about physical reasons that produces the double component has been carried out until now. In the present work the characteristics of the high impact weather conditions of the north western mediterranean are analysed in order to provide some physical explanation of the presence of the two components.

A simple model for rainfall storm generation is used here to provide synthetic point rainfall records for a sequence of storms. It is shown that the use of two different advection velocities for rainstorms, derived from the analysis of the high impact weather characteristics of the western Mediterranean, produces the two components in the rainfall depth maxima distributions.

2 HIGH IMPACT WEATHER IN THE WESTERN MEDITERRANEAN

The European atmospheric blocking has long been recognized as a physical process of profound dynamical and meteorological interest.

This phenomenon influences in crucial way the meteorology of the european region and a lot of observational, theoretical and numerical works as been devoted to study the blocking condition (Tibaldi and Buzzi, 1983; Tibaldi and Molteni, 1990).

In particular observational studies have demonstrated that advection velocity of convective systems, generated by cyclones and interacting with orographical obstacles, is influenced by blocking condition. The experimental evidence provided by geostationary satellites shows, from a synoptic point of view, that the spatial evolution of convective systems causing heavy rainfall in northern Tyrrhenian region is affected by the blocking condition. The presence or absence of the blocking condition produces a bimodality on the displacement velocity distribution with peaks around 1 m/s (“slow” storms) and 10 m/s (“fast” storms) (Boni, 1996; Conti et al., 1994).

This bimodal distribution of velocity, by the point of view of the ground effects, can generate different quantities of precipitation and so produce a different distribution of rainfall depths observed on assigned time interval: the blocking action of the high pressure cell causes long permanence of the phenomena over the same area, allowing the release of large amounts of rainfall. The physical meaning of the TCEV parameters allows to provide a phenomenological validation of the experimental evidence on the observed ARM, through the experimental evidence on the bimodality of the advection velocity provided by satellite imagery: the presence of two populations in the rainfall process, that can be produced by that bimodality, is the basic hypothesis of the TCEV (Rossi et al, 1984). Then a clear relation can be found between the frequency distribution parameters and the two rainfall population characteristics through a simple model for rainfall generation based on velocity bimodality. In the following the structure of this model is described and the results of the simulations are reported, with special emphasis on the capabilities of reproducing the statistics of the ground observed data.

3 A SIMPLE MODEL FOR RAINFALL EXTREMES

For each of M (M=10000) simulation years n rainfall events are generated using N(E[N], σ_N). Each event *i* belongs to the "slow" population with probability p such that:

P[i $\in$ "slow population"]=p
P[i $\notin$ "slow population"]=1-p

Each event has an average spatial extension D, whose value is assumed deterministic.

The inner structure of the event is simulated with a simple model using the Bartlett-Lewis random pulses scheme in the space domain with duration D_S, interarrival Int and intensity I. These three quantities are random variables having an exponential distribution with expected value μ_{DS}, μ_{Int} e μ_I.

The switch from the spatial domain to the temporal domain is done using the advection velocity of the event that it is assumed bimodal,, as a linear combination of two normal distributions with parameters μ_{V1}, σ_{V1} and μ_{V2}, σ_{V2}, to take into account the presence/absence of the blocking condition. The definition of the advection velocity is just one of the most important aspects of this model and it is closely related to the distinction between "slow" and "fast" storms, supposed to be able to produce annual rainfall maxima respectively belonging to extreme component and medium component.

When the rescaling operation has been done, for each simulation year the model provides n rainfall time series: the annual maxima for different durations are obtained from these synthetic time series.

In order to calibrate the model the parameters have been classified into three different categories:

- Parameters calibrated from high resolution raingauge data: E[N], σ_N, μ_{Int},, μ_{V1}, σ_{V1} and μ_{V2}, σ_{V2};
- Parameters calibrated from experimental evidence or inferred from literature P, D and μ_{DS};
- Parameters calibrated from annual rainfall maxima expected values μ_I.

The values of the parameters are reported in table 1.

E[N]	16
σ_N	5
μ_{Int}	18 Km
μ_{V1}	6 Km/h
σ_{V1}	1 Km/h
μ_{V2}	25 Km/h
σ_{V2}	3 Km/h
P	5%
D	500 Km
μ_{DS}	5 Km
μ_I	10 mm/h

Table 1. Values of parameters used in order to calibrate the model (best set).

4 CASE STUDY ANALYSIS AND RESULTS

From the synthetically generated time series the ARM for 1, 3, 6, 12 and 24 hours are derived in order to compare them with the recorded ones. Historical series of short duration ARM for the Liguria region are available for analyzed durations.

The data are assumed TCEV parent. The expression (1) used for the TCEV is:

$$F_X(x) = \exp\left(-\Lambda_1 * \exp\left(-\frac{x}{\theta_1}\right) - \Lambda_2 * \exp\left(-\frac{x}{\theta_2}\right)\right) \qquad (1)$$

where

Λ_1=expected number of independent medium component events
Λ_2=expected number of independent extreme component events
θ_1=expected value of medium component
θ_2=expected value of extreme component

First the model capability of reproducing the behavior of real world data, as regards the parameters Λ_1, Λ_2, θ_1 and θ_2, has been tested. In particular the sample and model ratios Λ_2/Λ_1 and θ_2/θ_1, that affect the shape of the rainfall frequency curves (growth curves), have been compared for different durations.

The results are shown (Fig.1 and Fig.2): the comparison between the real world data and the synthetic ones is quite successful. In fact the ratio between expected number of independent medium component events and extreme component events is conserved and its variability with observation time interval is correctly modeled (Fig.1). As regards the ratio between expected value of medium component and extreme component, the sample and model values are quite different, but their variability with observation time interval is correctly modeled (Fig.2). This means that relying only on a simple model based on the observed differences in time evolution of the heavy rainfall in the Mediterranean, due to the European blocking, it is possible to reproduce correctly the observed variation with the time interval of TCEV parameters and the observed rainfall frequency curves (growth curves) (Fig.3).

This is an important phenomenological validation of this experimental evidence, renforced by the analysis of the capabilities of the model to reproduce other observed statistics such as scale invariance.

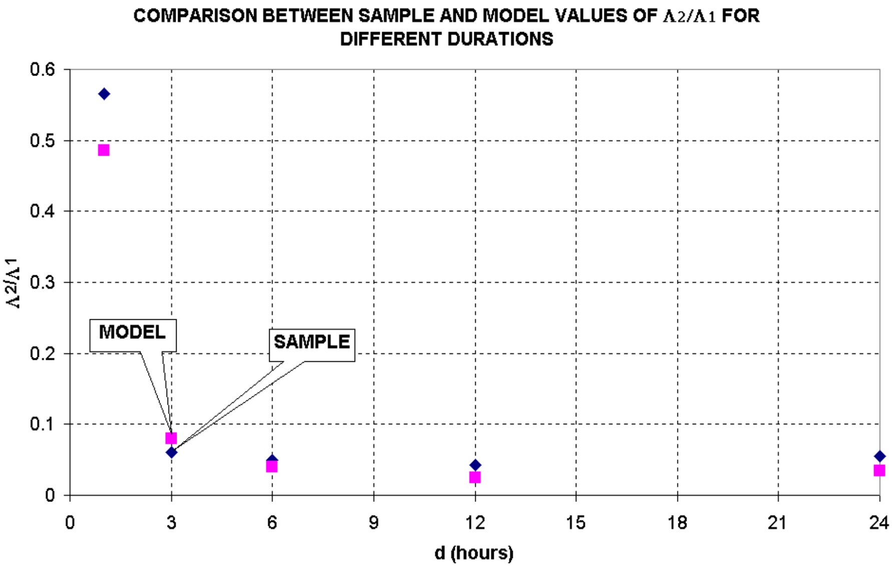

Figura 1: The comparison between sample and model values of $\Lambda 2/\Lambda 1$ for different durations is successful. The value of this ratio and its variability with time interval observation are correctly modeled.

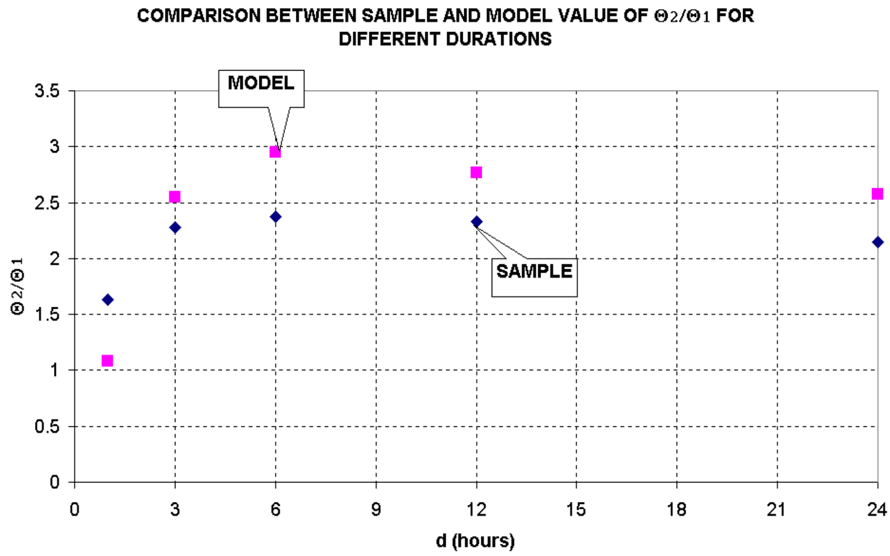

Figura 2: the comparison between sample and model values of θ2/θ1 for different durations is quite successful. The sample and model values are quite different, but their variability with observation time interval is correctly modeled.

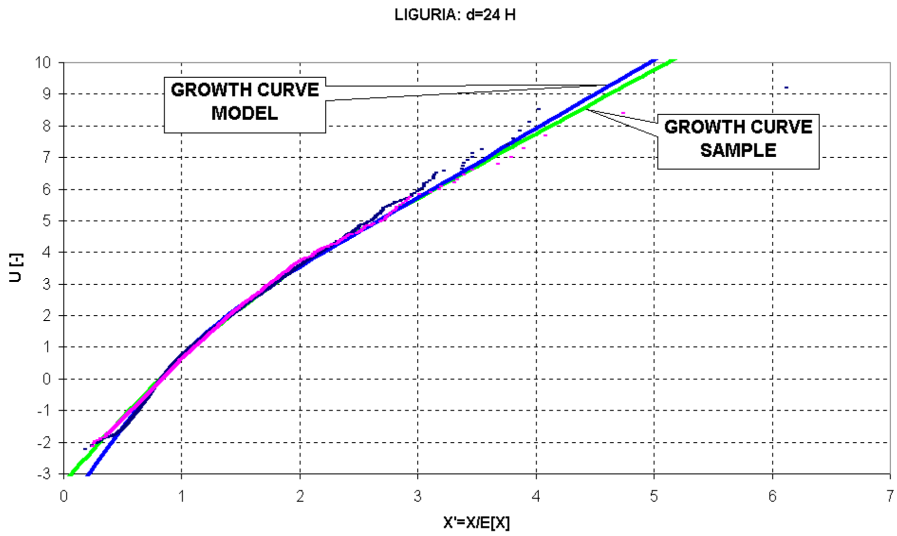

Figure 3: comparison between sample and model growth curve on Gumbel chart (d=24 hours). The synthetic and sample data are plotted in order to highlight the good agreement with the respective curves.

The observed ARM show self similarity for at least a limited interval of durations d. That means that the expected value varies with duration following:

$$\frac{E[H_d]}{E[H_1]} = d^n \qquad (2)$$

and that the coefficient of variation CV is constant with the duration.

The synthetic data have been analyzed and compared with the observed ones in order to test if the model respects this property. It is shown that there is a good agreement between synthetic and observed ARM in term of scale invariance (Fig.4 and Fig.5).

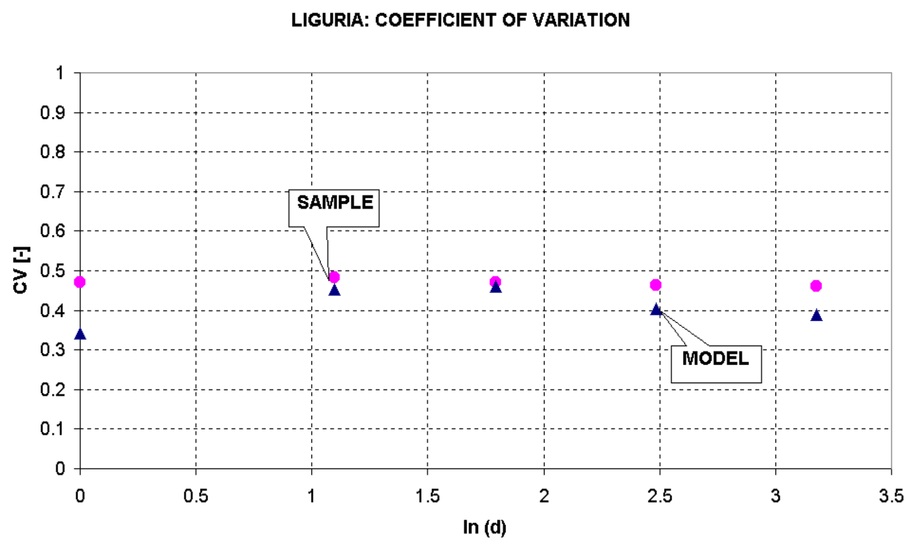

Figure 4: comparison between sample and model generated CVs: the scale invariance is quite well reproduced.

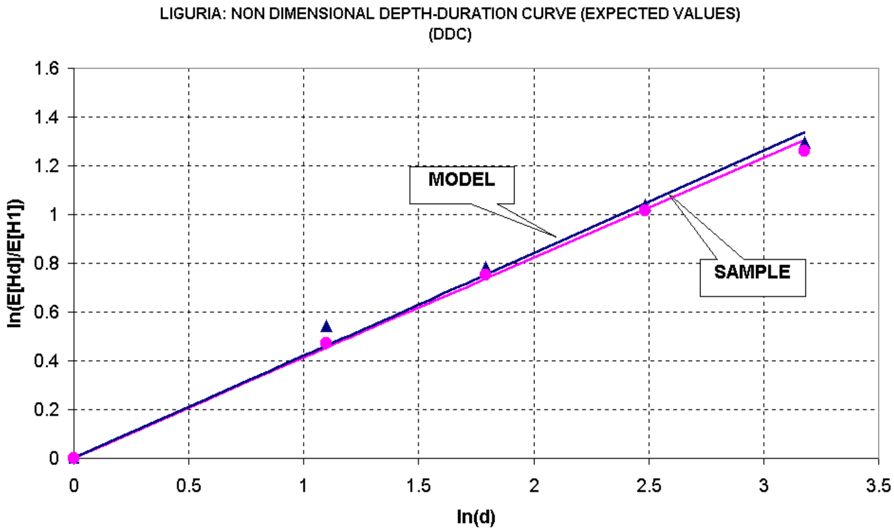

Figure 5: comparison between scale curves related to synthetic and sample data. The scale curves are very close: the process, by means of which the rainfall depths are rescaled depending on rainfall duration, is reproduced correctly.

Finally, the model performance has been validated analyzing that synthetic data generated using a unique set of advection velocity, high or low, the same being the values of the remaining parameters should not show the double-component behavior, whatever duration is considered. Figures from 6 to 7 clearly demonstrate that in such a case the parent distribution is not TCEV. The result supports the hypothesis of a link between presence of double component behavior of annual rainfall maxima and bimodal storm advection velocity.

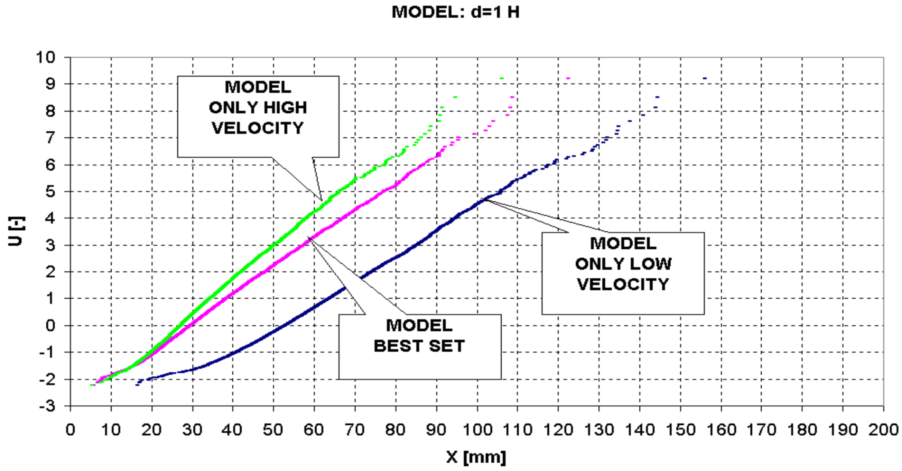

Figure 6: the hourly synthetic annual rainfall maxima generated using a unique set of advection velocity (high or low), the same being the values of the remaining parameters, or the best set parameters do not show the double-component behavior: generation process of hourly annual rainfall maxima is weakly influenced by the storm advection velocity.

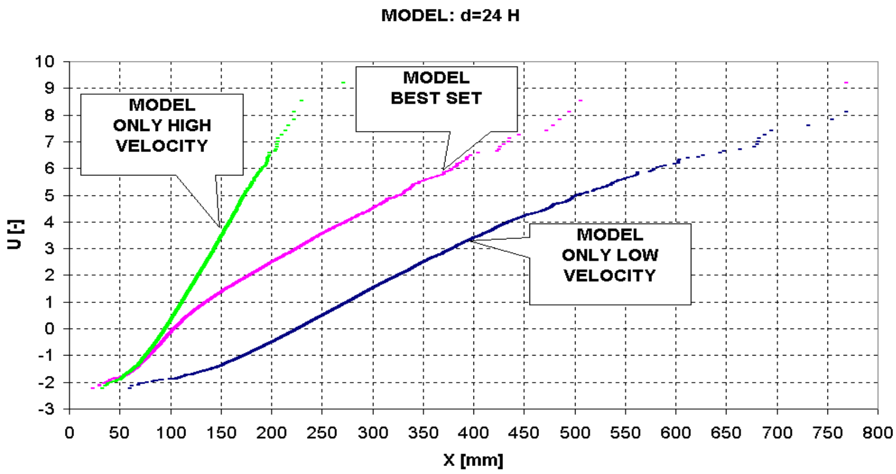

Figure 7: the synthetic annual rainfall maxima (d=24 hours) generated using a unique set of advection velocity (high or low), the same being the values of the remaining parameters, do not show the double-component behavior. This data are compared with synthetic maxima generated using the best set parameters and showing the double component behavior.

The mean and standard deviation of synthetic and sample data, for the different durations, have been compared. The agreement between sample and synthetic values is very good and the mean square error related to mean and standard deviation (MSE) confirm this fact (Fig.8).

This result confirms the fact that the model is able to generate, in correct way, synthetic rainfall series and respective statistic moments.

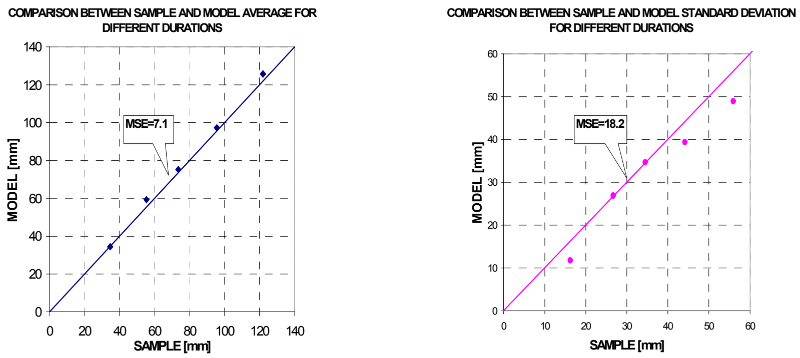

Figure 8: comparison between sample and model mean and standard deviation for different durations: the model is able to generate in correct way synthetic rainfall series and respective statistic moments.

5 COMMENTS AND CONCLUSIONS

The experimental evidence on the ARM of the Tyrrhenian coast of Italy shows the presence of a double component on the frequency distribution with the expected number of events belongings to the two component variable with the duration. Recent observation of extreme events in the Mediterranean from a synoptic point of view showed a possible interaction between permanence time of the phenomena producing extreme rainfall and the presence of European blocking through the blocking effects on frontal systems approaching the region This effect has been considered as a responsible for the production of the two components in rainfall distribution and a simple model for rainfall generation made to take into account for two different advection speed of the frontal systems has been used. The model demonstrated to be able to reproduce all the significant statistics of the observed ARM series. The most important result is that the model is able to correctly reproduce, for all durations, the observed ratios between medium and extreme component parameters and, most important, its variation with the duration, being the model parameters inferred from indipendent sources of information. This represents an important validation for the phenomenological explanation of both the presence of the second component and variation with duration of ratios between first and second component based on the European blocking effects.

REFERENCES

Arnell, N.W. & Gabriele, S. The performance of the two component extreme value distribution in regional flood frequency analysis, *Water Resources Research*, 1984, 20, pp.879-887.

Boni, G., Bolla, R., LaBarbera, P., Lanza, L., Marchese, M. & Zappatore S. The tracking and prediction of high intensity rainstorms, *Remote sensing Reviews*, 1996, 14, pp. 151-183.

Cannarozzo, M., D`Asaro, F. & Ferro, V. Valutazione delle onde di piena in Sicilia, CNR-GNDCI, Palermo, 1993.

Conti, M., Lanza, L. & Siccardi, F. Predictability of heavy rainfall patterns over the southern European region. Storm 93 and GNDCI recent research experiences, *Flood inundations Related to Large Earth Movements,* Trento, October 4-7, 1994.

Copertino, V. & Fiorentino, M. Valutazione delle onde di piena in Puglia, CNR-GNDCI, Potenza, 1992.

Jenkinson, A.F. The frequency distribution of the annual maximum (or minimum) values of meteorological events, *Q. J. R. Meteorol. Soc.,* 1955, 81, pp. 158-171.

Rossi, F., Fiorentino, M. & Versace, P. Two-component extreme value distribution for flood frequency analysis, *Water Resources Research*, 1984, 20, pp. 847-856.

Rossi, F. & Villani, P. Valutazione delle onde di piena in Campania, Rapporto Regionale Campania, CNR-GNDCI, 1994.

Tibaldi, S. & Buzzi, A. Effects of orography on Mediterranean lee cyclogenesis and its relationship to European blocking, *Tellus*, 1983, 35 A, pp.269-286.

Tibaldi, S. & Molteni, F. On the operational predictability of blocking, *Tellus*, 1990, 42 A, pp. 343-365.

Versace, P., Ferrari, E., Fiorentino, M., Gabriele, S. & Rossi, F. La valutazione delle piene in Calabria. CNR-GNDCI, LINEA 1, CNR-IRPI, Geodata, Cosenza, 1989, Cartografia.

WMO Selection of distribution types for extremes precipitation, Operational Hydrology Report, 15, Word Meteorological Organisation, 560, Geneva, Switzerland, 1981.