Hydrol. Earth Syst. Sci. Discuss., 5, 903–933, 2008
www.hydrol-earth-syst-sci-discuss.net/5/903/2008/

Hydrology and Earth System Sciences Discussions

Papers published in *Hydrology and Earth System Sciences Discussions* are under open-access review for the journal *Hydrology and Earth System Sciences*

Effects of runoff thresholds on flood frequency distributions

A. Gioia[1], V. Iacobellis[1], S. Manfreda[2], and M. Fiorentino[2]

[1]Dipartimento di Ingegneria delle Acque e di Chimica, Politecnico di Bari, Bari, Italia
[2]Dipartimento di Ingegneria e Fisica dell'Ambiente, Università degli Studi della Basilicata, Potenza, Italia

Received: 17 January 2008 – Accepted: 17 January 2008 – Published: 1 April 2008

Correspondence to: A. Gioia (a.gioia@poliba.it)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Runoff generation during extreme floods usually occurs whenever rainfall forcing exceeds a given threshold. In many cases, different thresholds may be identified as responsible of the hydrological losses during ordinary events or extraordinary events at the basin scale. Such thresholds are shown to be related to the dynamics of soil saturation of the river basin and to account for the high skewness of their annual flood distributions. In basins where ordinary floods are mostly due to a small portion of the surface which is particularly prone to produce runoff, depending on permeability of a river basin and its antecedent soil moisture conditions, severe rainfall may exceed a basin-wide soil storage threshold and produce the so-called outlier events responsible of the high skewness of floods distributions. In this context, the derived theoretical model based on the concept of variable contributing area to peak flow proposed by Iacobellis and Fiorentino (2000) was generalized with the aim of incorporating such kind of dynamics in the description of the phenomena. The work produced a new formulation of the derived distribution where the two runoff components are explicitly considered. The present work was validated by using as test site a group of basins belonging to Southern Italy and characterized by flood distributions with high skewness. The application of the proposed model provided a good fitting to the observed distributions. Moreover, model parameters were found to be strongly related to physiographic basin characteristics giving consistency to the modelling assumptions.

1 Introduction

Operational methods for flood prediction try to maximize the exploitation of available information at regional scale (Chow et al., 1988; Cunnane, 1989; Moisello, 1989; Reed, 1999). Among these, regional analysis (e.g., index flood method, NERC, 1975) is based on transferability of hydrological information allowing prediction in ungauged basins. The transferability lies on the concept of hydrologic similarity whose analysis and detection is probably the most challenging task for the hydrologist. In fact, external knowledge about climate, geology, soil hydraulic properties and land cover is necessary in order to individuate homogeneous areas. Such kind of analysis is required in order to test the statistical homogeneity and\or the spatial variability of parameters (e.g., Lu and Stedinger, 1992; Hosking and Wallis, 1993, 1997; Viglione et al., 2007; Chebana and Ouarda, 2007).

Regionalization techniques allow the use of distributions with more than two parameters (e.g., GEV, TCEV), whose estimation procedures usually need extensive dataset, in particular for parameters dependent on the higher order moments.

However, the evolution of territory, anthropic impact and climate change require that methodologies for hydrological prediction become more reliable. Nowadays, it is common opinion that they should be based not only on the statistical analysis of the observed data, but also on the schematization of the physical processes acting at the basin scale. In this context, the analytical derivation of the probability distribution of floods with parameters related to climatic, geopedologic and morphologic basin features, plays a fundamental role that deserve to be explored deeper in details.

This can be done by deriving the flood distribution starting from the rainfall distribution and using a rainfall-runoff model which includes absorption and flow routing processes, as originally suggested by Eagleson (1972) who firstly tackled the problem analytically. After this first attempt, numerous scientists reported further analytical works (e.g. Haan and Edwards, 1988; Raines and Valdes, 1993; Kurothe et al., 1997; Goel et al., 2000; Franchini et al., 2005).

Among others, Sivapalan et al. (1990) implemented a model accounting for the effect of different mechanisms of runoff generation (namely hortonian and dunnian). Iacobellis and Fiorentino (2000) introduced the variability of partial contributing area in the basin schematization. De Michele and Salvadori (2002) evaluated the influence of the antecedent soil moisture condition on the flood frequency distribution by means of the SCS-CN method. Other authors performed the derivation of the flood frequency distribution through MonteCarlo numerical simulation (e.g., Consuegra et al., 1993; Muzik, 1993; Loukas, 2002) or by continuous simulation approach (e.g., Beven, 1987; Bras et al., 1985; Blazkova and Beven, 2002; Fiorentino et al., 2007).

The bibliography of the latest years shows large interest toward the improvement of these models, which hold a potential cognitive heritage not yet completely explored, and orientates the analysis toward the investigation of the relationships between hydrologic parameters and physiographic soil characteristics with the aim to improve the prediction in ungauged basins (e.g., Parajka et al., 2005; McIntyre et al., 2005; Lee et al., 2007; Bardossy, 2007).

These models are implemented combining a reliable representation of the physical dynamics and a simplified mathematical schematisation in order to provide an analytical solution of the problem. Thus, the development of the derived theoretical distribution of floods is strictly related to a deeper understanding of the processes due to climate, soil and vegetation interactions. For the specific case of runoff, its dynamic is controlled by the soil moisture state and characteristics of the surface. For instance, Fiorentino et al. (2007) observed that the mechanism of flood formation tend to be controlled by the basin pedology in arid basins, whereas morphology assumes a fundamental role in humid basins where the saturation excess process is more frequently dominant.

The effects of catchment thresholds have received attention in derived flood frequency analysis only in the last few years. Kusumastuti et al. (2007) investigated the effect of catchment storage thresholds and their impact on flood frequency. The authors derived the flood frequency distributions exploiting a Monte Carlo simulation approach combined with non-linear conceptual rainfall-runoff models. The study provides insights in the analysis of the interactions between climatic inputs, landscape properties and soil moisture antecedent condition, and runoff response analyzing the roles of non-linear thresholds on flood frequency curves.

The impact upon flood frequency due to spatial heterogeneity in non-linear thresholds, temporal variability in storm associated with seasonality, and space-time variability in the storage-discharge relationship associated with the rainfall-runoff process, was recently analysed by Struthers and Sivapalan (2007). They observed that temporal variability in storm properties with seasonality increases the frequency of threshold exceedence and the magnitude of the flood response associated with a given runoff process. Spatial variability in landscape and climatic properties provides a spatial variability in the local frequency of threshold exceedence, while the decreasing of soil depth towards the stream masks the impacts of threshold upon the resulting flood frequency.

In the present work, we focus on the role of soil losses in runoff generation emphasising the way they are affected by rainfall amount and intensity. The goal is to improve the descriptive properties of theoretically derived distributions with particular attention on their ability of coping with the Matalas condition of separation. With this aim we generalize the theoretical probability distribution proposed by Iacobellis and Fiorentino (2000) introducing a two component derived distribution where the role of rainfall thresholds is emphasized and reconnected to the analysis of soil-vegetation dynamics. In particular, two different mechanisms of runoff generation are considered, and their non-linear effects on the flood frequency distribution is explained and parametrized.

2 Theoretically derived flood frequency distribution (IF model)

Iacobellis and Fiorentino (2000) proposed a theoretical model for derived flood frequency distribution based on the concept of variable contributing (or source) area. This model is hereinafter referred to as the "IF model".

The IF model is based upon the following assumptions. The peak of direct streamflow Q is modelled as the product of two random variables strongly correlated, the (partial) source area contributing to runoff peak a and runoff peak per unit of a, u_a. Both random variables are controlled by: (*i*) rainfall intensity, duration and areal extension; (*ii*) runoff concentration; (*iii*) hydrological losses. The probability distribution of u_a, can be derived from the probability distribution of rainfall intensity conditional on a duration equal to τ_a which is a characteristic lag-time of a. The runoff peak per unit area,u_a, is considered linearly dependent from the areal net rainfall intensity in a time interval equal to τ_a. Fiorentino et al. (1987a) showed that, within a large range of observed intensity-duration-frequency (IDF) curves and of basin response functions, the basin lag-time is close to the critical rainfall duration. This last being defined as the rainfall duration, extracted from the same IDF, which maximize the flood peak. Moreover, the ratio between net rainfall intensity in τ_a and peak runoff can be assumed as constant routing factor ξ which can be confidently set to 0.7.

Thus, the runoff per unit area is

$$u_a = \xi(i_{a,\tau} - f_a), \tag{1}$$

where $i_{a,\tau}$ is the average areal rainfall intensity in τ_a covering the contributing area a, f_a is the corresponding space-time averaged hydrologic loss. The exceedance probability function of the peak of direct streamflow Q, $G'_Q(q)$, is found as the integral of the joint probability density function (pdf) of a and u_a

$$G'_Q(q) = \int_0^A \int_{\frac{q}{u_a}}^{\infty} g(u|a)\, g(a)\, du\, da, \tag{2}$$

where u_a is expressed as runoff peak u conditional on a.

The pdf of u_a is found from the pdf of area rainfall intensity $i_{a,t}$ which is assumed as a Weibull function with two parameters θ and k:

$$g\left(i_{a,\tau}\right) = \frac{k}{\theta_{a,\tau}} i_{a,\tau}^{k-1} \exp\left(-\frac{i_{a,\tau}^{k}}{\theta_{a,\tau}}\right), \tag{3}$$

with

$$\theta_{a,\tau} = E\left[i_{a,\tau}^{k}\right] = \left(E\left[i_{a,\tau}\right]/\Gamma\left(1+1/k\right)\right)^{k}. \tag{4}$$

The lag-time τ_a scales with a according to the power law

$$\tau_a = \tau_1\, a^{\nu}. \tag{5}$$

The IF model assumes that a power law relationship also exists between $E[i_{a,\tau}]$ and a

$$E\left[i_{a,\tau}\right] = i_1\ a^{-\varepsilon} = E\left[i_{A,\tau}\right]\ (a/A)^{-\varepsilon} \tag{6}$$

and between f_a and a

$$f_a = f_1\ a^{-\varepsilon'} = f_A(a/A)^{-\varepsilon'}, \tag{7}$$

where i_1 and f_1 are respectively the average rainfall intensity and the average hydrologic loss for contributing area equal to 1; $E[i_{A,\tau}]$ and f_A are referred to contributing area equal to the entire basin.

An estimate of f_A can be obtained by means of the relationship

$$\Lambda_q = \Lambda_p \exp\left(-\frac{f_a^k}{E[i_{a,\tau}^k]}\right) \tag{8}$$

where Λ_q is mean annual number of floods and Λ_p is the mean annual number of rainfall events and f_a is the corresponding runoff threshold.

With the aim of explaining the key role of f_A, Fiorentino and Iacobellis (2001) investigated its variability in real basins and found that Eq. (7) holds in arid and semi-arid basins of Puglia and Basilicata with parameter values f_1=37 [mm h^{-1}km$^{-2\varepsilon'}$] and ε'=0.5. In humid basins low values and low variability of f_A were found for basins with different area A, giving the following estimates: f_1=0.7 [mm h^{-1}km$^{-2\varepsilon'}$] and ε'=0. Such results are of particular interest because Eq. (7) represents a significant signature of basin hydrological response and its dependence on climate testifies the strong control of climate-soil-vegetation factors on flood frequency.

In particular, the different behaviour of arid basins compared to humid ones has been related to the characteristics of prevalent runoff generation mechanisms (Fiorentino and Iacobellis, 2001). In fact, as regards arid basins, it has been showed how the estimated value ε'=0.5 in Eq. (7) demonstrates that runoff occurs only when a soil storage capacity has been filled. Then, since in this climate intense rainfall are likely to occur after a long dry period, this capacity tends to be close to soil water content at saturation. An important related outcome is that peak-discharge contributing area is expected to be controlled by the basin pedology and its average to be mainly related to dominant storm extension and to heterogeneity of soil storage capacity with particular regard to the fraction of less permeable area. Conversely, in humid and semi-humid basins, ε'=0 indicate the existence of a constant threshold of rainfall intensity conventionally related to average infiltration rate of the soil-bedrock system in saturation conditions (Fiorentino and Iacobellis, 2001). Here in fact, due to persistent precipitations as well as to dense vegetation and forested hillslopes, a portion of the basin near the channel network is likely to be wet and close to saturation prior any intense storm event. Then, this portion can yield surface runoff even if rainfall is not particularly severe, because the surface runoff is there produced as soon as rainfall intensity is greater than the infiltration capacity of underlying soil-bedrock system. As a consequence, peak-discharge contributing area expands and contracts depending on the surface and subsurface conditions such as soil moisture and vegetation state at the time prior to the flood event, thus showing an expected value which is likely to be a small part only of the basin surface.

Such considerations also affect the distribution of variable contributing area a whose pdf $g(a)$ is found as sum of a continuous gamma function $\Gamma(\alpha, \beta)$ and the probability P_A=*prob[a=A]* times the Dirac function $\delta(.)$

$$g(a) = \frac{1}{\alpha\Gamma(\beta)}\left(\frac{a}{\alpha}\right)^{\beta-1}\exp\left(-\frac{a}{\alpha}\right) + \delta(a-A)P_A \tag{9}$$

Parameters α and β respectively represent position and scale of the Gamma distribution. Thus, the following relationship holds

$$\alpha = rA/\beta \tag{10}$$

where r is a dimensionless parameter

$$r = E[a]/A. \tag{11}$$

Thus, under the hypothesis that flood occurrence is of Poisson type, Iacobellis and Fiorentino (2000) derived the cumulative distribution function (cdf) of the flood annual maximum values of the flood peak $Q_p=Q+q_o$, where q_o is the base flow estimated as the average monthly flow observed in January and February (see Table 1).

Further analyses (Fiorentino and Iacobellis, 2001; Fiorentino et al., 2003; Fiorentino et al., 2007) highlighted the role of climate, geology, pedology and landuse factors on the frequency of extreme events.

Fiorentino and Iacobellis (2001) pointed out the main role of the hydrological losses on flood occurrences under different physical and climatic conditions. Such results were further confirmed by Fiorentino et al. (2007) that jointly analyzed the results of the IF model and the outputs of a continuous simulation scheme using a distributed hydrological model (DREAM–Manfreda et al., 2005) in cascade with a rainfall generator. Such methodology provided interesting insights on the role of physical factors, such as soil texture and basin morphology, on runoff production at basin scale.

In particular the "r" estimates, for arid river basins (Fiorentino and Iacobellis, 2001) showed a significant correlation with the permeability index defined as

$$\psi = \psi_h + 0.9\psi_m \tag{12}$$

where ψ_h and ψ_m are the fraction of the total area with outcrops belonging respectively to the highly permeable lithoid complexes and lithoid complexes with medium permeability. In the work by Fiorentino et al. (2003) the r estimates for arid basins in southern Italy were found to show a significant correlation also with the runoff coefficient "C", introduced by De Smedt et al. (2000) that depends on soil type, land-use and local slope. In humid basins characterized by more frequent rainfall, events which generate a high antecedent saturation condition, soil moisture pattern results organized reflecting the basin morphology. In fact, a dependence of the expected value of the contributing area (i.e. of parameter "r") on the variation coefficient of the topographic index proposed by Kirkby (1975) was found. Instead, in the case of arid basins, the basin morphology assumes a secondary role respect to the pedology, because floods are likely to occur when the soil moisture pattern is not organized, so that the surface runoff is more related to soil texture and land use (Fiorentino et al., 2007).

3 Two Component IF Model (TCIF)

The discussion above reported can be summarized in the following statements: (i) in humid basins ordinary floods are likely to be produced by a relatively small portion of the basin area and they occur when an infiltration rate threshold is exceeded; (ii) in arid basins, they likely occur when a soil storage capacity is filled and are mostly due to larger contributing areas, which are controlled by pedology rather than geomorphology. It is to point out that the saturation conductivity is generally much lower than the ratio of soil storage capacity to basin lag-time, then the first threshold is commonly lower.

Nevertheless, apart from what is prevailing, the two mechanisms may occur in all basins, with different weight and importance, depending on several factors including climate, geomorphology, soil hydraulic features, bedrock permeability. In all cases ordinary floods are provided by less severe storms insisting on small portion of the basin close to the channels while the remaining (greater) portion of the surface, including large hillslopes mostly contributes when large amount of precipitation occurs with intensity and duration sufficient to fill the basin-wide storage capacity during the event.

Thus, in humid basins, as well as in arid ones, floods are here assumed to be produced by two different mechanisms, the former responsible for ordinary flood peaks and the latter related to rarer and more intense events.

Such conjecture suggests an original phenomenological explanation leading to two-component probability distributions like the TCEV (Rossi et al., 1984), which is nowadays largely used for flood frequency analysis.

Based on the rationale above, a new two-component probability distribution is herein proposed based on a generalization of the IF model. It has been given the name of Two Component IF Model (TCIF), based on the hypothesis that runoff is produced according to two different threshold mechanisms both defined in accordance with Eq. (1):

an “L-type” mechanism, activated by the lower threshold;

$$u_{a,L} = \xi(i - f_{a,L}) \quad (13)$$

“H- type” mechanism, activated by the higher threshold:

$$u_{a,H} = \xi(i - f_{a,H}) \quad (14)$$

These equations describe the peak unit runoff provided by means of the same routing factor ξ and two different runoff thresholds $f_{a,L}$ and $f_{a,H}$, with $f_{a,H} > f_{a,L}$.

As regards the probability distribution of flood-peak contributing areas, they are obviously different for each runoff generation mechanism considered, and with particular regard to their expected value, it is assumed that L-type mechanism deals with smaller expected contributing areas. Thus, in analogy with IF model – see Eq. (11) – two variable a_H and a_L are here introduced, with different expected values, which allow to define two dimensionless parameters $r_L = E[a_L]/A$ and $r_H = E[a_H]/A$, with $r_H > r_L$.

Therefore, by means of the defined scheme, it is possible to derive two different peak flow distributions:

$$G'_{Q,L}(q) = \int_0^A \int_{\frac{q}{a}}^{\infty} g\left(u|a_L\right) g\left(a_L\right) du\, da_L \quad (15)$$

for L-type events and

$$G'_{Q,H}(q) = \int_0^A \int_{\frac{q}{a}}^{\infty} g\left(u|a_H\right) g\left(a_H\right) du\, da_H \quad (16)$$

for H-type events.

Then, considering a Poisson process of exceedances of the introduced thresholds, the *cdf* of the flood annual maximum is

$$F_{Q_p}(q_p) = \exp\left\{-\Lambda_L\left[G'_{Q,L}(q_p)\right] - \Lambda_H\left[G'_{Q,H}(q_p)\right]\right\}, \tag{17}$$

with Λ_L and Λ_H mean annual number of independent flood events arising from L-type and H-type runoff generation mechanisms.

As in the IF model – see Eq. (8) – also in this case the mean annual numbers of occurrences are related to the corresponding thresholds:

$$\Lambda_q = \Lambda_L + \Lambda_H = \Lambda_p \exp\left(-\frac{f^k_{A,L}}{E[i^k_{A,\tau}]}\right) \tag{18}$$

$$\Lambda_H = \Lambda_p \exp\left(-\frac{f^k_{A,H}}{E[i^k_{A,\tau}]}\right) \tag{19}$$

where $f_{a,L}$ and $f_{a,H}$ scale with contributing area according to the following power law relationships:

$$f_{a,L} = f_{A,L}(a/A)^{-\varepsilon'} \tag{20}$$

$$f_{a,H} = f_{A,H}(a/A)^{-\varepsilon''} \tag{21}$$

The proposed distribution includes, among others, 6 parameters, $f_{A,L}$, $f_{A,H}$, r_L, r_H, ε', ε", strictly related to the occurrence of the two main different mechanisms of runoff generation. Three parameters are added with respect to the IF model. Nevertheless, they have a strong physical meaning and much about their behaviour is known from previous applications of the IF model.

Thus, with the aim of investigating the effect of the two thresholds on flood frequency we analyzed the regional behavior of $f_{A,L}$, $f_{A,H}$, and r_L, r_H in order to characterize their variability and dependence on measurable factors related to climate-soil-vegetation dynamics.

4 Study cases

4.1 Climatic, Geologic and Morphologic characteristics of the investigated zones

The investigated area includes a number of basins with area ranging from 15 to 1140 km^2 in three regions of Southern Italy, namely Basilicata, Calabria and Puglia. These regions are quite heterogeneous for climate, geology and land use characteristics.

The climate is quite variable due to the morphological differences of land surfaces. Within the north-eastern sector (Puglia), characterized by low hills or flat lands, climate is of the hot-dry Mediterranean type (semiarid or dry sub-humid), with mild, not very rainy winters and warm-dry summers. As one proceeds to the West-Southern sector (Basilicata and Calabria), climate becomes more cold and humid (Southern Appennine).

The mean annual rainfall ranges from minimum values (about 600 mm) observed in Puglia and higher values (up to 1800 mm) in Basilicata and Calabria. Rainfall is distributed quite irregularly over the year with an average in the October–March semester that is more than twice the amount of the period April–September. July is the least rainy month, while the highest varies from October to January.

The climatic pattern is fairly reflected by the climatic index (Thornthwaite, 1948):

$$I = \frac{h - E_p}{E_p} \tag{22}$$

with h mean annual rainfall depth and E_p mean annual potential evapotranspiration calculated according to Turc's formula (Turc, 1961) dependent only on the mean annual temperature.

Vegetation and land cover in the region investigated is quite consistent with climatic features and morphological conditions. Arid and semi-arid zones are characterized by scarce vegetation, which gradually turns into sub-humid Mediterranean undergrowth

(Macchia Mediterranea), wheat crops and pasture land, to finally reach the mountain woods of humid and hyper-humid areas.

The different lithological units of the investigated regions and the nature of the outcrops show characters of permeability of different type and degree. In fact, within the area examined, sediments and rocks permeable because of porosity and fissuring can be distinguished. In some areas the system also shows communicating cracks like bedding joints, faults, and intense circulation of underground water. SubAppennine clays and flysch formations are also present with interbedding of marls and sandstones.

In order to investigate the thresholds effect on the flood probability distributions we focused on series of annual maximum flood characterized by high skewness coefficient (Ca>1.7). Thus records of ten gauged sites (Ca) in the area were selected including 8 humid basins in Basilicata and Calabria and two arid basins in Puglia (Fig. 1). In Table 1 climatic and morphometric characteristics of the investigated basins, as well as some significant statistics of their recorded annual maximum flood series, are reported.

Figures 2 and 3 provide a description of the spatial variability of the mean runoff coefficient "C" and the permeability index "ψ" computed for all the investigated basins. The estimates of "C" and "ψ" (reported in Table 1) were computed within a GIS environment adopting lythological, pedological, land cover and local slope maps. In particular for Puglia and Basilicata regions "C" and "ψ" values are those reported respectively in Fiorentino at al. (2003) and in Fiorentino and Iacobellis (2001). For Calabria Basin they are estimated using Corinne Land-Cover map, geological map (scale 1:50 000) and DTM of Italy (250 m).

5 Model application to gauged basins

The structure of the proposed TCIF model of Eq. (17) and its probabilistic construction suggest interesting insights into the analysis of the right hand tail of the annual flood distribution. In particular, Eq. (17) arises from the application of the compound Poisson processes as well as the TCEV distribution (Rossi et al., 1984). In Sect. 3, in analogy with the theory of the TCEV distribution, we identified two main event mechanisms as responsible of ordinary and extraordinary events respectively. The former characterized by frequent occurrences (higher mean annual number Λ) and lower average of exceedances. The latter provided by rare events (lower Λ) and higher average of exceedances. In this section we exploit the TCEV model in order to obtain reliable estimates of some important distribution parameters and for comparison purposes.

The TCEV distribution has four parameters, namely Λ_1, θ_1, Λ_2, θ_2, and its cumulative distribution function is

$$F_x(x) = P[X \leq x] = \exp\left[-\Lambda_1 \exp\left(-{}^{x}/_{\theta_1}\right) - \Lambda_2 \exp\left(-{}^{x}/_{\theta_2}\right)\right], \quad (23)$$

with $\theta_2 > \theta_1 > 0$ and $\Lambda_1 > \Lambda_2 > 0$ the model has been mostly used in statistical regional analysis. In facts, at-site estimation techniques (Fiorentino et al., 1987b) in principle are not recommended for short length data series because of the very high estimator variability, with particular regard to parameters dependent on the second and third order moments.

Nevertheless, based on the analogy between Eqs. (17) and (23), for estimating Λ_L and Λ_H we assumed

$$\Lambda_L = \Lambda_1 \text{ and } \Lambda_H = \Lambda_2, \quad (24)$$

and exploited the available robust techniques for at-site estimation of TCEV parameters.

In other words, in order to obtain reliable estimates for parameters to be used within the derived model applied to observed annual flood series, we used at-site estimation of TCEV parameters performed by Maximum Likelyhood Estimator. Such a choice was made also to assess that differences between data series as well as their high skewness were mostly due to a physical control rather then to sample variability. Thus, an attempt is made to substitute the basic approach of statistical regionalization models aimed to look for basin's similarities with a more physically based approach thought to identify their differences. Furthermore, in this frame one should note that an effort is made to reduce uncertainty due to short data records, at the price of introducing uncertainty related to soil information which, in a way, is more prone to be knocked down by the advent of new technologies for earth observation.

For estimating other TCIF parameters we used the following information on rainfall and geomorphological features of selected basins. In particular, exploiting the relationship between the averages of annual maxima and the base process, the expected value of the space-time averaged rainfall intensity $E[i_{A,\tau}]$, with regard to the total area of the basin, occurring in the duration τ_A may be evaluated by means of:

$$E\left[i_{A,\tau}\right] = \frac{p_1 \tau_A^{n-1} \left[1 - \exp\left(-1.1 {\tau_A}^{0.25}\right) + \exp\left(-1.1 \tau_A^{0.25} - 0.004\,A\right)\right]}{\Lambda_p S_{\Lambda_p}} \tag{25}$$

in which the U.S. Weather Bureau areal reduction factor is used (see Eagleson, 1972), $\Lambda_q S_{\Lambda q}$ is a factor that allows to elapse from the mean of the base process to the mean of maxima (see Iacobellis and Fiorentino, 2000), p_1 and n (Table 1) are the parameters of the at-site IDF curves referred to the expected value of the annual maximum rainfall intensity in the duration τ_A (Table 1).

Assuming the hypothesis of Weibull distribution of rainfall intensity and poissonian occurrence of events, the distribution of annual maxima turns out to be a Power Extreme Value (PEV) type.

A regional estimation based on a PEV-ML procedure was applied to the study area with the aim to evaluate the exponent (k) of the Weibull distribution of rainfall intensity; regional values of k (e.g. Fiorentino and Iacobellis, 2001) are equal to 0.8 in Puglia and Basilicata (resulting by regional analysis performed on the annual maxima rainfall series recorded at 178 gauging stations) and 0.53 (e.g. Claps et al., 2000) in Calabria (resulting by regional analysis performed on the annual maxima in 225 raingauge stations with record length not less than 20 years).

The values of Λ_p used here, are regional estimates reported by Fiorentino and Iacobellis (2001) and Claps et al. (2000) and are displayed in Table 1.

HESSD
5, 903–933, 2008

Once parameters dependent on rainfall were assigned the threshold values $f_{A,L}$ and $f_{A,H}$ were estimated by exploiting Eqs. (18), (19) and (4). Results are shown in table 2 and their analysis allows the individuation of significant regional patterns and the assignment of physically consistent values for their scaling exponents ε' and ε''.

In fact, Fig. 4 shows that, within the investigated region, the first mechanism threshold corresponds to a low and constant infiltration rate practically independent from basin area, while the second runoff threshold follows, with a good approximation, a scaling law relationship with exponent equal to 0.5 representative of a soil water storage control.

Thus, following Fiorentino and Iacobellis (2001), we assumed ε'=0 in Eq. (20) for the L -type scaling relationship and ε'=0.5 in Eq. (21) for the H-type mechanism.

Finally, parameters r_L and r_H were found as heuristic estimates providing the best fitting of the theoretical Annual maximum flood distribution to the observed time series.

The proposed model shows good performances in terms of descriptive ability. It was compared with the at-site TCEV probability distribution (see Fig. 5 and Table 2). The comparison between the CDFs obtained by the TCEV and TCIF models (whose parameters are reported in Table 2) for all 10 basins with corresponding plotting positions is reported in Fig. 5. Skewness of the observed distributions is always captured by the TCIF model.

Further analyses were carried out with the aim of investigating main physical controls on model parameters r_L and r_H. In particular we searched for significant dependence between the estimated parameter values and physical features of the investigated basins. Although these analyses are affected by the approximation used in the estimation procedure, we would like to remark that parameter r_L shows a linear dependence on the permeability index, for area ranging between 2% and 22% of the entire basin area and obviously decreases for higher values of permeable bedrocks. For the H-type mechanism, instead, the expected value of the contributing area ratio (r_H) ranges between 4% and 70% and is linearly related to the runoff coefficient "C" confirming that hydrological losses mainly depend on soil type and land cover (Fig. 6).

These results are in strong agreement with patterns observed in previous studies for basins belonging to Puglia, Basilicata and Calabria stating that in all the investigated basins, with more or less pronounced effect, depending on prevailing controls, we observe:

- a first (L-type) component characterized on average by small contributing area and constant rainfall threshold depending on infiltration rate in saturation condition;
- a second (H-type) component typically showing larger mean contributing area and rainfall threshold scaling with area mainly accounting for soil storage capacity.

6 Conclusions

This work focuses on the dynamic of flood generation processes, with the principal aim of individuating main controls on flood frequency distribution and the second purpose of investigating regional patterns and spatial variability of distribution model parameters.

A simple analytical schematization of the different mechanisms of runoff production is proposed in order to explain the high skewness in flood distributions. The strong non-linearity in flood formation is mainly ascribed to the presence of different runoff thresholds linked to two different mechanisms of runoff generation and their impact on flood frequency distribution is investigated exploiting the derived distribution scheme proposed by Iacobellis and Fiorentino (2000).

We started from the basic assumption that in humid basins, as well as in arid ones, the soil storage capacity can be filled for storms that are rather rare. In such basins then, floods can be produced by two different mechanisms, the former responsible for more ordinary flood peaks and the latter related to rarer and more intense events. This conjecture deserves deep attention and provides a phenomenological explanation for two-component probability distributions as for instance the TCEV model (Rossi et al., 1984), which is nowadays largely used for flood frequency analysis.

Hence, a two component derived distribution model (TCIF) is proposed in order to account for highly skewed flood distributions under the hypothesis that, as it more frequently happens in humid basins, ordinary floods may arise from partial areas which have been mostly saturated by antecedent precipitations, while more rare and intense rainfall may generate the activation of a soil storage capacity threshold which is responsible for the contribution of a much wider part of the drainage area to overland flow.

We should remark that this interpretation highlights the role of infiltration control in different kind of basins and in accordance with mechanisms of runoff generation. In particular, in humid basins, where the saturation excess runoff generation mechanism is commonly thought to prevail, the infiltration control is operated on a narrow contributing area that is already close to saturation at the event time, and furthermore, it is organized around the river network thanks to antecedent rainfall and significant subsurface flow prior the event. On the other hand, in arid basins, where the infiltration excess mechanism is generally supposed to be mainly responsible for runoff generation, average infiltration is mainly affected by initial adsorption and soil storage capacity. Nevertheless, a mixed mechanism may occur in arid basins also, due to particular soil-bedrock conditions in small but not negligible portion of the basin. Two particular cases are here shown. In basin #1 (I=−0.24) a deep and very fine textured clay soil surrounds most of the channel network. In basin #2 (I=−0.17) a bedrock of low permeability due to lake deposits of piroclastic sediments and alluvial deposits of fine texture outcropping nearby the entire channel network. In both cases, in fact, ordinary floods could be provided by not severe storms insisting on that portion of the basin while the remaining (greater) portion of the surface would contribute only if precipitations would be intense enough to fill the soil storage capacity of the remaining part of the basin during the event.

The relationships observed in Figs. 4 and 6 support the role played by basin physiographic characteristics on model parameters and allow to recognize and classify the behaviour of basins within such hydrological schemes. Such relationships between model parameters and physical characteristics of a river basin may provide also a useful tool to reduce the uncertainty in flood prediction especially at the higher return periods.

The role of runoff thresholds reserves strong attention and an alternative way for deeper investigation may be provided by the comparison with virtual laboratories and continuous simulation distributed modelling.

Acknowledgements. Authors gratefully acknowledge the support of MIUR (Italian Ministry of Instruction, University and Research) under the grant PRIN CoFin2005 entitled "Climate-soil-vegetation interaction in hydrological extremes".

References

Bárdossy, A.: Calibration of hydrological model parameters for ungauged catchments, Hydrol. Earth Syst. Sci., 11, 703–710, 2007, http://www.hydrol-earth-syst-sci.net/11/703/2007/.

Beven, K.: Towards the use of catchment geomorphology in flood frequency predictions, Earth Surf. Proc. Land., 12, 6–82, 1987.

Bras, R. L., Gaboury, D. R., Grossman, D. S., and Vicen, G. J.: Spatially varying rainfall and flood risk analysis, J. Hydraul. Eng., 111, 754–773, 1985.

Blazkova, S. and Beven, K. J.: Flood frequency estimation by continuous simulation for a catchment treated as ungauged (with uncertainty), Water Resour. Res., 38, 8, 1–14, 2002.

Chebana, F. and Ouarda, T. B. M. J.: Multivariate L-moment homogeneity test, Water Resour. Res., 43, W08406, doi:10.1029/2006WR005639, 2007.

Chow, V. T., Maidment, D. R., and Mays, L. W.: Statistical and Probability Analysis of Hydrologic Data, McGraw Hill (Ed.), Appl. Hydrol., 1988.

Claps, P., Fiorentino, M., and Iacobellis, V.: Regional flood frequency analysis with a theoretically derived distribution function, Proc. of the EGS Plinius Conference on Mediterranean Storms, 2000.

Consuegra, D., Meylan, P., Musy, A.: A probabilistic approach to estimate design parameters for flood control projects, Kuo, C. Y. (Ed.), Engineering Hydrology. ASCE, San Francisco, 455–460, 1993.

HESSD

5, 903–933, 2008

Effects of runoff thresholds on flood frequency distributions

A. Gioia et al.

Cunanne, C.: Statistical distributions for flood frequency analysis, Wmo Operational Hydrology report, 33, Geneva, 1989.
De Michele, C. and Salvadori, G.: On the derived flood frequency distribution: analytical formulation and the influence of antecedent soil moisture condition, J. Hydrol., 262, 245–258, 2002.
De Smedt, F., Yongbo, L., and Gebremeskel, S.: Hydrologic modelling on a catchment scale using GIS and remote sensed land use information, Brebbia, C. A. (Ed.), Risk Analysis II, WIT press, Southampton, Boston, 295–304, 2000.
Eagleson, P. S.: Dynamics of flood frequency, Water Resour. Res., 8, 878–898, 1972.
Fiorentino, M., Rossi, F., and Villani, P.: Effect of the basin geomorphoclimatic characteristics on the mean annual flood reduction curve, Proc. of the 18th Annual Conference on Modeling and Simulation, Pittsburgh, part5, 1777–1784, 1987a.
Fiorentino, M., Gabriele, S., Rossi, F., and Versace, P.: Hierarchical approach for regional flood frequency analysis, Singh, V. P. (Ed.), Regional flood frequency analysis, 35–49, D. Reidel, Norwell, Mass, 1987b.
Fiorentino, M. and Iacobellis, V.: Non-linearity in the process of floods generation, Proc. of the EGS Plinius Conference, Maratea, Italy, 2000.
Fiorentino, M. and Iacobellis, V.: New insights about the climatic and geologic control on the probability distribution of floods, Water Resour. Res., 37, 721–730, 2001.
Fiorentino, M., Carriero, D., Laguardia, G., Manfreda, S., Margiotta, M., Rosano, R., Sole, A., and Iacobellis, V.: Una proposta metodologica per la mappatura della variabilità spaziale delle perdite idrologiche durante i fenomeni di piena, Proc. of "Convegno Nazionale sulla Conservazione dell'ambiente e rischio idrogeologico", CNR-GNDCI, 2830, 147–154, 2003.
Fiorentino, M., Manfreda, S., and Iacobellis, V.: Peak runoff contributing area as hydrological signature of the probability distribution of floods, Adv. Water Resour., 30, 2123–2134, 2007.
Franchini, M., Galeati, G., and Lolli, M.: Analytical derivation of the flood frequency curve through partial duration series analysis and a probabilistic representation of the runoff coefficient, J. Hydrol., 303, 1–15, 2005.
Gabriele, S. and Iiritano, G.: Alcuni aspetti teorici ed applicativi nella regionalizzazione delle piogge con il modello TCEV, GNDCI-Linea 1 U.O. 1.4, Pubblicazione N., 1089, Rende (Cs), 1994 (in Italian).
Gioia, A., Iacobellis, V., Manfreda, S., and Fiorentino, M.: Climate and soil controls on flood frequency, Proc. of WSEAS International Conferences, 223–230, 2007.

Goel, N. K., Kurothe, R. S., Mathur, B. S., and Vogel, R. M.: A derived flood frequency distribution for correlated rainfall intensity and duration, J. Hydrol., 228, 56–67, 2000.

Haan, C. T. and Edwards, D. R: Joint probability estimates of return period flows, Trans. American Society of Agricultural Engineers, 31, 4, 1115–1119, 1988.

Hosking, J. R. M. and Wallis, J. R.: Regional Frequency Analysis:An Approach Based on L-Moments, Cambridge Univ. Press, New York, 224 pp., 1997.

Hosking, J. R. M. and Wallis, J. R.: Some statistics useful in regional frequency analysis, Water Resour. Res., 29, 271–281, 1993.

Iacobellis, V. and Fiorentino, M.: Derived distribution of floods based on the concept of partial area coverage with a climatic appeal, Water Resour. Res., 36, 469–482, 2000.

Kirkby, M. J.: Hydrograph modelling strategies, in: Progress in Physical and Human Geography, edited by: Peel, R. F., Chisholm, M. D., and Haggett, P., Heinemann, London, 69–90, 1975.

Kurothe, R. S., Goel, N. K., and Mathur, B. S.: Derived flood frequency distribution for negativaly correlated rainfall intensity and duration, Water Resour. Res., 33, 2103–2107, 1997.

Kusumastuti, D. I., Struthers, I., Sivapalan, M., and Reynolds, D. A.: Threshold effects in catchment storm response and the occurrence and magnitude of flood events: implications for flood frequency, Hydrol. Earth Syst. Sci., 11, 1515–1528, 2007, http://www.hydrol-earth-syst-sci.net/11/1515/2007/.

Lee, H., Zehe, E., and Sivapalan, M.: Predictions of rainfall-runoff response and soil moisture dynamics in a microscale catchment using the CREW model, Hydrol. Earth Syst. Sci., 11, 819–849, 2007, http://www.hydrol-earth-syst-sci.net/11/819/2007/.

Loukas, A.: Flood frequency estimation by a derived distribution procedure, J. Hydrol., 255, 69–89, 2002.

Lu, L. and Stedinger, J. R.: Sampling variance of normalized GEV/PWM quantile estimators and a regional homogeneity test, J. Hydrol., 138, 223–245, 1992.

Manfreda, S., Fiorentino, M., and De Smedt, F.: An investigation on flood generation mechanisms by coupling distributed models at different time scales, Proc. of 3rd EGS Plinius Conference on Mediterranean Storms, Baja Sardinia, 309–314, 2001.

Manfreda, S., Fiorentino, M., and Iacobellis, V.: DREAM: a distributed model for runoff, evapotranspiration, and antecedent soil moisture simulation, Adv. Geosci., 2, 31–39, 2005, http://www.adv-geosci.net/2/31/2005/.

Matalas, N. C., Slack, J. R., and Wallis, J. R.: Regional skew in search of a parent, Water Resour. Res., 11, 815–826, 1975.

McIntyre, N., Lee, H., Wheater, H., Young, A., and Wagener, T.: Ensemble predictions of runoff in ungauged catchments, Water Resour. Res., 41, W12434, doi:10.1029/2005WR004289, 2005.

Moisello, U.: Regional statistical analysis of floods in Italy, Excerpta, 4, 39–65, 1989.

Muzik, I.: The method of derived distributions applied to peak flows, Kuo, C. Y., (Ed.), Engineering Hydrology, American Society of Civil Engineers, San Francisco, 437–442, 1993.

NERC (Natural Environment Research Council): Flood Studies Report, I, Hydrologic studies, NERC Natural Environment Research Council, London, 1975.

Parajka, J., Merz, R., and Blöschl, G.: A comparison of regionalisation methods for catchment model parameters, Hydrol. Earth Syst. Sci., 9, 157–171, 2005, http://www.hydrol-earth-syst-sci.net/9/157/2005/.

Raines, T. H. and Valdes, J. B.: Estimation of flood frequencies for ungauged catchments, J. Hydraul. Eng, 119, 1138–1154, 1993.

Reed, D. W.: Flood Estimation Handbook-Overview, 1., Institute of Hydrology, Wallingford, UK, 1999.

Rossi, F., Fiorentino, M., and Versace, P.: Two component extreme value distribution for flood frequency analysis, Water Resour. Res., 20, 847–856, 1984.

Sivapalan, M., Wood, E. F., and Beven, K. J.: On hydrologic similarity, 3, A dimensionless flood frequency model using a generalized geomorphologic unit hydrograph and partial area runoff generation, Water Resour. Res., 26, 43–58, 1990.

Struthers, I. and Sivapalan, M.: A conceptual investigation of process controls upon flood frequency: role of thresholds, Hydrol. Earth Syst. Sci., 11, 1405–1416, 2007, http://www.hydrol-earth-syst-sci.net/11/1405/2007/.

Thornthwaite, C. W.: An approach toward a rational classification of climate, Am. Geograph. Rev., 38, 1, 55–94, 1948.

Turc, L.: Estimation of irrigation water requirements, potential evapotranspiration: a simple climatic formula evolved up to date, Annals of Agronomy, 12, 13–14, 1961.

Viglione, A., Laio, F., and Claps, P.: A comparison of homogeneity tests for regional frequency analysis, Water Resour. Res., 43, W03428, doi:10.1029/2006WR005095, 2007.

HESSD

5, 903–933, 2008

Effects of runoff thresholds on flood frequency distributions

A. Gioia et al.

HESSD

5, 903–933, 2008

Effects of runoff thresholds on flood frequency distributions

A. Gioia et al.

Table 1. Climatic and morphological features of the investigated basins.

n.	Basins	A (km^2)	I	Runoff coefficient "C"	ψ	Λ_p	Ca	$\tau_A(h)$	p_l (mm/h)	n	q_o (m^3/s)
1	Celone at Ponte Foggia San Severo	233	−0.24	0.48	0.98	44.60	2.43	5.20	23.33	0.27	2.20
2	Venosa at Ponte Sant' Angelo	263	−0.17	0.52	0.85	44.60	2.26	5.60	24.13	0.26	1.4
3	Sinni at Valsinni	1140	0.57	0.40	0.41	21.00	2.42	5.60	23.13	0.40	45.00
4	Basento at Gallipoli	853	0.28	0.52	0.40	21.00	2.25	4.80	21.00	0.31	25.00
5	Alli at Orso	46	1.26	0.18	0.98	20.00	2.74	3.00	33.20	0.52	2.34
6	Corace at Grascio	182	0.90	0.24	0.94	20.00	1.83	3.80	29.80	0.45	8.84
7	Alaco at Mammone	15	1.66	0.13	1.00	10.00	1.76	1.30	39.60	0.63	0.96
8	Tacina at Rivioto	79	1.43	0.12	0.97	10.00	2.79	3.00	32.70	0.59	3.40
9	Trionto at Difesa	32	0.90	0.22	0.99	20.00	3.18	2.80	31.00	0.50	1.17
10	Amato at Marino	113	0.86	0.29	0.95	20.00	2.43	4.60	28.80	0.43	5.32

Table 2. Parameters of the investigated models.

		TCIF Parameters					TCEV Parameters			
n.	Basins	$f_{A,L}$ (mm/h)	$f_{A,H}$ (mm/h)	r_l	r_H	k	Λ_1	θ_1	Λ_2	θ_2
1	Celone at Ponte Foggia San Severo	2.07	7.83	0.07	0.60	0.80	6.61	15.64	0.19	67.93
2	Venosa at Ponte Sant' Angelo	1.93	6.35	0.05	0.70	0.80	6.74	11.10	0.38	103.41
3	Sinni at Valsinni	0.00	6.90	0.15	0.50	0.80	20.51	127.52	0.49	481.35
4	Basento at Gallipoli	0.95	6.34	0.22	0.67	0.80	8.33	110.11	0.38	337.22
5	Alli at Orso	1.13	20.67	0.02	0.05	0.53	8.18	4.94	0.38	18.43
6	Corace at Grascio	0.95	7.37	0.10	0.28	0.53	6.52	39.63	1.19	114.14
7	Alaco at Mammone	0.31	18.33	0.03	0.04	0.53	6.65	3.23	1.11	11.29
8	Tacina at Rivioto	0.00	10.24	0.02	0.25	0.53	8.90	6.52	1.10	85.15
9	Trionto at Difesa	1.26	25.75	0.02	0.10	0.53	7.75	2.26	0.21	19.05
10	Amato at Marino	0.60	8.92	0.06	0.52	0.53	8.30	11.71	0.70	102.90

HESSD

5, 903–933, 2008

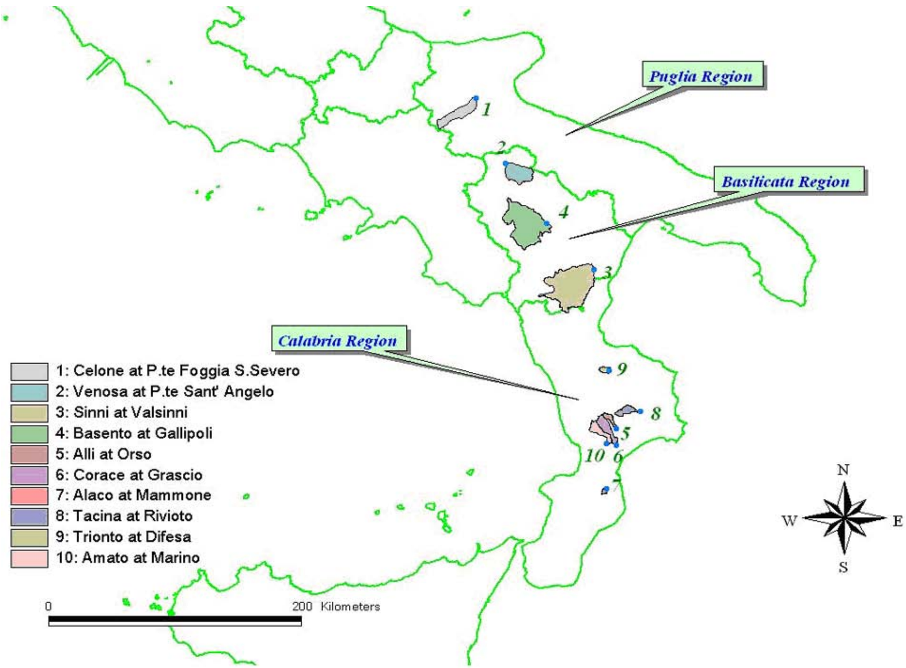

Fig. 1. The investigated basins.

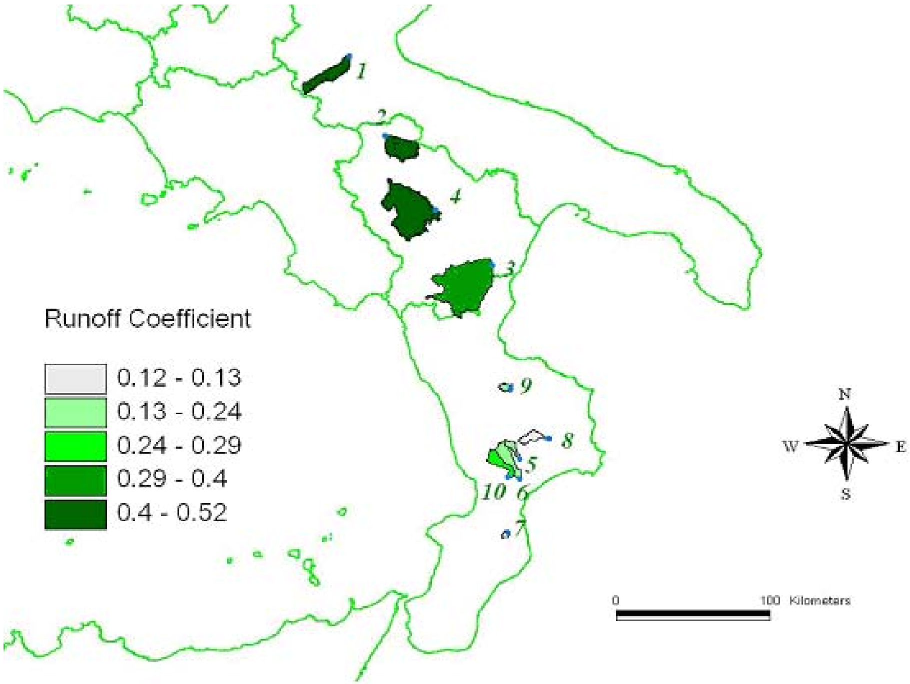

Fig. 2. Map of mean runoff coefficient "*C*" for the investigated basins.

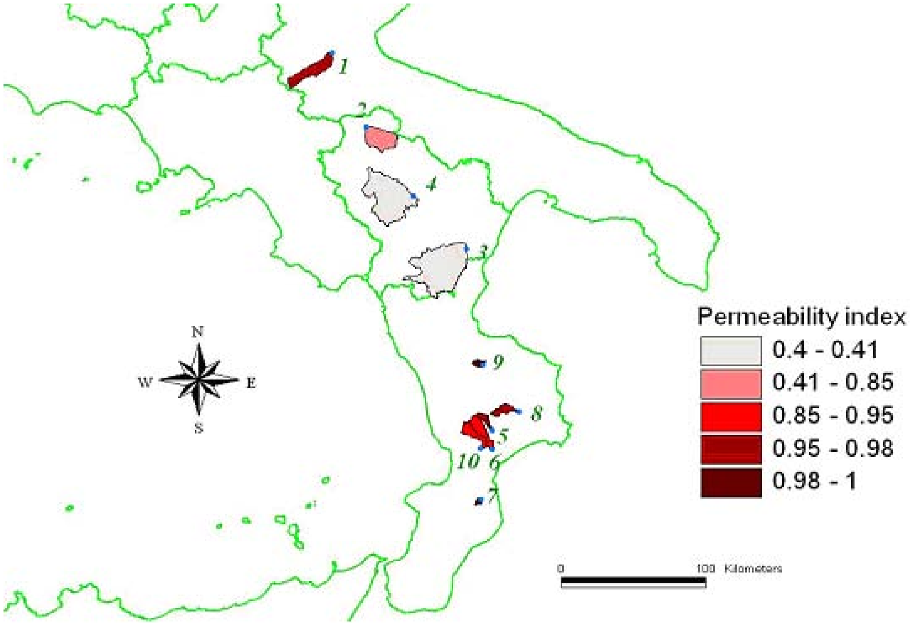

Fig. 3. Map of permeability index (ψ) for the investigated basins.

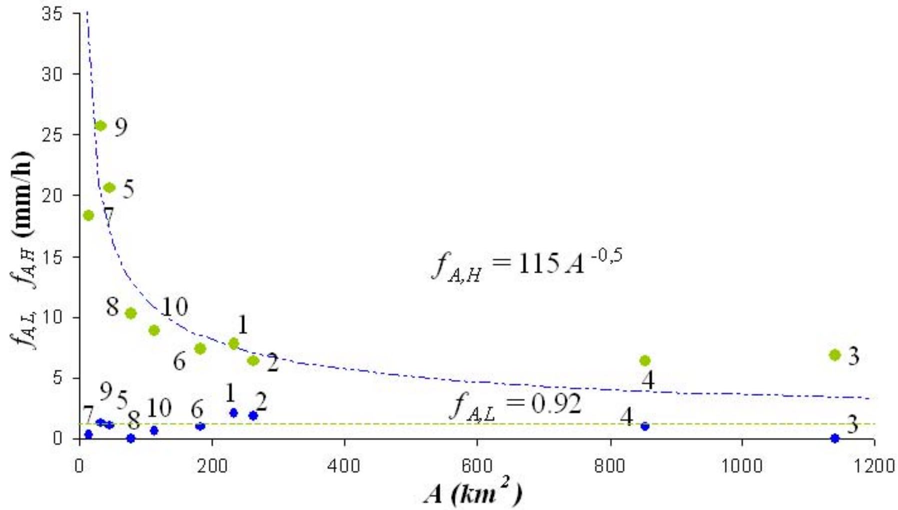

Fig. 4. Relationships between the two different runoff thresholds $f_{A,H}$ and $f_{A,L}$ and basin area.

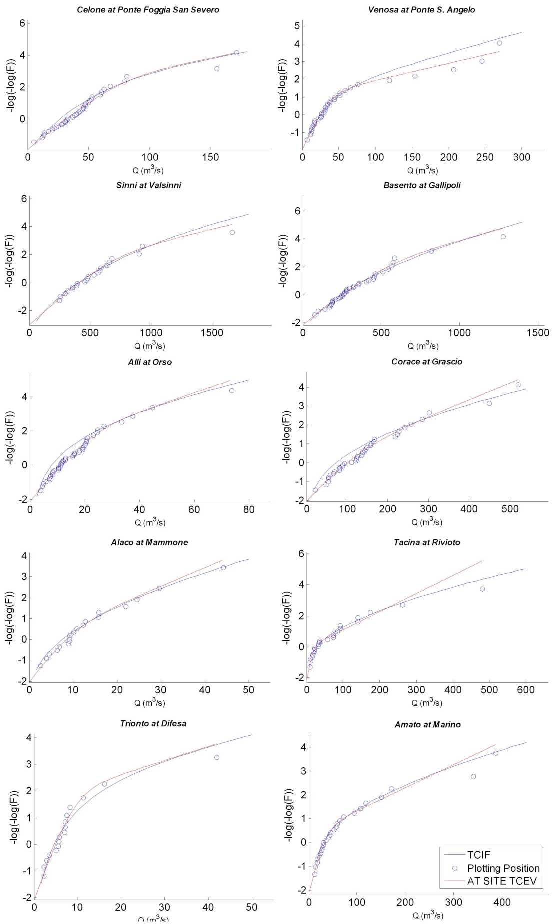

Fig. 5. A comparison between the TCIF model and at site TCEV model for the investigated basins.

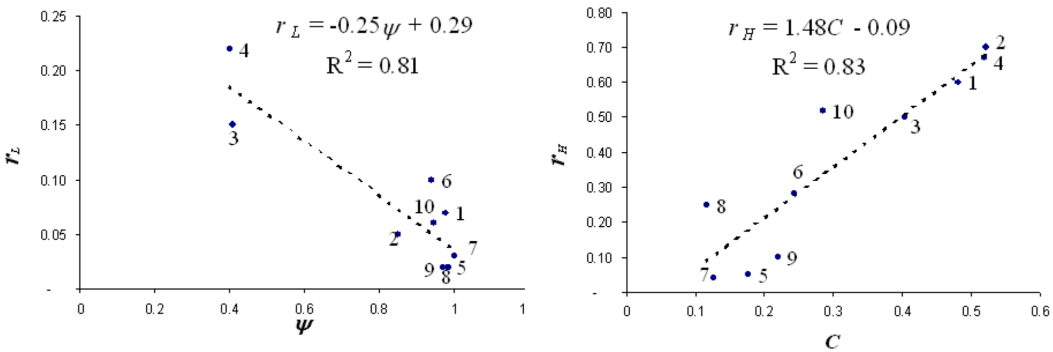

Fig. 6. Relationship between r_L and the permeability index "ψ" and between r_H and runoff coefficient "C".