**Hydrology and
Earth System
Sciences**

Forest cover influence on regional flood frequency assessment in Mediterranean catchments

F. Preti[1], G. Forzieri[2], and G. B. Chirico[3]

[1]Dipartimento di Economia, Ingegneria, Scienze e Tecnologie Agrarie e Forestali, Sezione di Ingegneria dei Biosistemi Agrari e Forestali, University of Florence, Florence, Italy
[2]Dipartimento di Scienze della Terra, formerly at Dipartimento di Ingegneria Civile e Ambientale, University of Florence, Florence, Italy
[3]Dipartimento di Ingegneria Agraria e Agronomia del Territorio, University of Napoli Federico II, Napoli, Italy

Abstract. The paper aims at evaluating to what extent the forest cover can explain the component of runoff coefficient as defined in a regional flood frequency analysis based on the application of the rational formula coupled with a regional model of the annual maximum rainfall depths. The analysis is addressed to evaluate the component of the runoff coefficient which cannot be captured by the catchment lithology alone. Data mining is performed on 75 catchments distributed from South to Central Italy. Cluster and correlation structure analyses are conducted for distinguishing forest cover effects within catchments characterized by hydromorphological similarities. We propose to improve the prediction of the runoff coefficient by a linear regression model, exploiting the ratio of the forest cover to the catchment critical rainfall depth as dependent variable. The proposed regression enables a significant bias correction of the runoff coefficient, particularly for those small mountainous catchments, characterised by larger forest cover fraction and lower critical rainfall depth.

1 Introduction

Flood peak assessment is fundamental for planning and design structural and non-structural flood risk mitigation actions. This is achieved by flood frequency analysis, which aims at estimating the probability distributions of flood peaks, so that the flood magnitude for any design return period can be easily determined. A direct assessment of flood frequency and magnitude is only possible for those limited number of catchments where stream flow gauges have been operating for a significant number of years. For ungauged catchments, the approach commonly employed in engineering hydrology is the regional analysis, which exploits the hydrological similarities among catchments and the scaling properties of flood statistics for exporting the information available in gauged catchments to ungauged catchments (e.g. Cunnane, 1988; Gupta and Waymire, 1990; Stedinger et al., 1993). The method of regional flood frequency analysis most widely applied by hydrologists and engineers is the index flood method, originally introduced by Dalrymple (1960). This method is based on the identification of homogeneous regions, where the probability distributions of the annual maximum floods are assumed invariant except for a site-specific scale parameter known as the index flood. For any site, the flood peak discharge with an assigned return period is defined as the product of two terms: a dimensionless probabilistic growth factor and a site-specific index flood. The statistics of the normalized annual maximum flood series within the homogeneous region are pooled to define the dimensionless probabilistic growth curve, which is assumed invariant within the homogenous region. The index flood is generally assumed coincident with the mean of annual maximum of flood peaks, while a few flood estimation procedures adopt the median (e.g. IH, 1999). The literature contains numerous studies on the identification of homogeneous regions and the estimation of dimensionless probabilistic growth curve, while relatively fewer studies can be found on estimating the index flood, particularly in ungauged catchments. In fact, while a simple arithmetic mean of the available observations can provide a direct estimate of the index flood in gauged catchments, indirect methods are required

---

Correspondence to: F. Preti
(federico.preti@unifi.it)

for estimating the index flood at ungauged sites. Many indirect methods are based on multiregression models linking the index flood to a selected set of catchment descriptors (e.g. Kjeldsen and Jones, 2010), representing morphological, climatic and land use catchment features. These regression methods take generally little consideration of the physical phenomena underlying the transformation of rainfall in runoff and of the dominant flood generating mechanisms, although more recent studies evidenced that flood generating mechanisms can be relevant in flood regional analysis (Iacobellis and Fiorentino, 2000; Mertz and Blöschl, 2003). Other indirect methods provide an estimate of the flood index based on a conceptual description of the hydrological response of the basin to intense rainfall events. Within the flood assessment procedures employed in Italy, an indirect estimation method largely applied is based on a conceptual model structured according to the well-known rational formula (Rossi and Villani, 1992). Other conceptual approaches have been also proposed, based on an analytical derivation of the probability distribution of floods (e.g. Becciu et al., 1993; Brath et al., 2001; Bocchiola et al., 2003). The index flood estimation method based on the rational formula implicitly assumes that the average value of the annual maximum peak discharges is related to the average value of the annual maximum rainfall depth within a critical time interval, which is assumed equal to a characteristic time scale of the catchment response. A key parameter is the runoff coefficient, which conceptually represents the fraction of the total rainfall contributing to the flood peak response. Runoff coefficient values are derived by regional models with respect to selected catchment descriptors that can be easily distinguished at regional scale. A general tendency is to employ the catchment lithology as principal catchment descriptor of the runoff coefficient, while assuming negligible the additional information attached to the land-cover patterns for the assessment of flood extreme values, at least in rural catchments (e.g. Iacobellis and Fiorentino, 2000; Brath et al., 2001).

As result of the index flood approach, the flood frequency curves within homogenous regions are shifted according to the value of the index flood, while keeping the same shape.

Another regionalisation procedure is based on the use of rainfall-runoff models (e.g. Maidment, 1993). In this case, annual maximum rainfall depths of various durations are treated through a regionalisation procedure, such as the index value approach, whereas the discharge with an assigned return period is estimated by applying a rainfall-runoff model. This second type of regionalization procedure based on rainfall-runoff models is more effective in representing the flood frequency distribution in regions that, although being homogeneous in terms of annual maximum rainfall depths, include catchments with significantly different flood frequency curves as result of different rainfall runoff processes controlling the flood response. In engineering applications, the most widely used rainfall-runoff model is still the rational formula, often applying the runoff

coefficient as function of the return period (Schaake et al., 1967; French et al., 1974; Pilgrim, 1989; Cannarozzo et al., 1995). In this case, the runoff coefficient does not simply represent a runoff rainfall ratio, rather it assumes the role of a probabilistic factor controlling not only the position but also the slope and the curvature of the flood frequency curve, by means of a (generally, non-linear) transformation of the catchment rainfall frequency curve for a catchment characteristic time scale. In some studies, the dependency of the runoff coefficient from the return period has been explored by a specific probabilistic model, considering the runoff coefficient independent from the rainfall depth (Gottschalck and Weingartner, 1998; Franchini et al., 2005).

The dependency of the runoff coefficient with the return period is consistent with the consideration that, particularly in rural catchments, there is not a perfect correspondence between annual maximum peak discharges and annual maximum rainfall depths (e.g. Hiemstra and Reich, 1967; Franchini et al., 2000; Haschemi et al., 2000). The maximum annual flood peak is in fact controlled by catchment antecedent conditions and thus can be also generated by rainfall events with catchment average depth even below the annual maximum. This aspect cannot be represented with a constant runoff coefficient, as it occurs in traditional applications of the rational formula.

Provided that vegetation patterns can have a significant influence on the catchment antecedent conditions as well as on other rainfall runoff processes in rural catchments, in this study we explore to what extent forest cover can be employed to predict the runoff coefficient, in the framework of a regional flood frequency analysis based on the rational formula coupled with a regional analysis of annual maximum rainfall depths.

The paper is structured as follows: the following section presents the regional flood frequency analysis; the third section illustrates the available data set; the fourth section explores the dependency of the runoff coefficient from the forest cover; in the fifth section we present a new regression model of the runoff coefficient accounting for the forest cover; the sixth section is devoted to the discussion and the last section to the conclusions.

2   Regional flood frequency analysis based on the rational formula

The flood peak for a given return period T, Q_T, can be defined as follows:

$$Q_T = K_{Q,T}\,\mu_Q \tag{1}$$

where $K_{Q,T}$ is the dimensionless probabilistic growth factor of the floods for a return period equal to T and μ_Q is the index flood.

A regional flood frequency analysis based on the index flood assumes that the dimensionless probabilistic growth

factor, controlling the slope and the curvature of the flood frequency distribution, is invariant within the homogenous region, while the index flood, controlling the position of the flood frequency distribution, is variable and can be predicted by catchment specific descriptors. A common approach for estimating the index flood μ_Q in Italian ungauged catchments is based on a conceptual model structured according to the well-known rational formula:

$$\mu_Q = \phi \, A \, \frac{\mu \, [h_A \, (t_c)]}{t_c} \qquad (2)$$

where A is the catchment area, while ϕ is the runoff coefficient for the index flood ($0 < \phi \leq 1$), i.e. the ratio of the mean flood runoff to the mean rainfall depth, assumed to be independent of rainfall intensity and duration; $\mu[h_A(t_c)]$ is the catchment areal rainfall depth within a critical duration t_c, obtained by multiplying the point depth-duration-frequency curve $\mu[h(t_c)]$ (referred to the centre of the storm) with the area reduction factor κ_A, which is expressed as function of the catchment area and the critical duration t_c (e.g. Brath et al., 2001).

The runoff coefficient ϕ is estimated by regression models against selected catchment descriptors, with invariant parameters within homogenous regions. These regional models are calibrated with data of gauged catchments, where runoff coefficients (ϕ_{obs}) are computed by inversion of Eq. (2) with index flood values ($\hat{\mu}_Q$) assessed by arithmetic average of the annual flood peak experimental values:

$$\phi_{\mathrm{obs}} = \frac{\hat{\mu}_Q \, t_c}{\mu \, [h_A \, (t_c)] \, A}. \qquad (3)$$

It is important to observe that ϕ conceptually describes both the fraction of rainfall volume retained by the soil and the vegetation (i.e. the transformation of the total rainfall in net rainfall as result of processes such as infiltration, canopy interception and surface detention), and the dampening effect of the catchment, which implies the reduction of the flood peak as compared to the net rainfall intensity (for example, if the basin is modeled as a linear reservoir with lag time equal to t_L, the flood peak reduction with respect to a constant rainfall intensity is equal to $1 - e^{-t_c/t_L}$).

As anticipated above, regional regression models of the runoff coefficient generally exploit only classes of catchment lithology as predictor variables. At regional scale, the lithological features of the catchments can be grouped according to different classes corresponding to different degrees of permeability, such as (e.g. Fiorentino and Iacobellis, 2001): (1) highly permeable lithoid complexes constituted by sediments and rocks with porosity based permeability, rocks with fissure-based permeability, and those having a mixed permeability; (2) lithoid complexes with medium permeability constituted by permeable lithologies which outcrop on a steep surface or lithologies more or less fractured and filled by clayey material; and (3) impermeable lithoid complexes represented by clayey formations.

In the case of a regional flood frequency analysis coupling a regional model of the annual maximum rainfall depths and a rainfall-runoff model structured according to the rational formula, the flood peak Q_T for a given return period is estimated in ungauged catchments as follows:

$$Q_T = C_T \, A \, \frac{h_{A,T} \, (t_c)}{t_c} = C_T \, K_{R,T} \, A \, \frac{\mu \, [h_A \, (t_c)]}{t_c} \qquad (4)$$

where $h_{A,T} \, (t_c)$ is the catchment areal rainfall with return period T, expressed as the product of the dimensionless probabilistic growth factor of the rainfall for a return period equal to T, $K_{R,T}$, and the index value $\mu \, [h_A \, (t_c)]$; C_T is the runoff coefficient for a return period equal to T. Sample values of the runoff coefficient ($C_{T,\mathrm{obs}}$) can be derived by inverting Eq. (4) applied to gauged catchments, where enough data are available for a direct assessment of the flood frequency distribution, defined by $K_{Q,T}$ and $\hat{\mu}_Q$. In particular, by combing Eqs. (1) and (4), $C_{T,\mathrm{obs}}$ can be expressed as follows:

$$C_{T,\mathrm{obs}} = \left(\frac{\hat{\mu}_Q \, t_c}{\mu \, [h_A \, (t_c)] \, A} \right) \left(\frac{K_{Q,T}}{K_{R,T}} \right) = \phi_{\mathrm{obs}} \left(\frac{K_{Q,T}}{K_{R,T}} \right). \qquad (5)$$

Equation (5) shows that C_T is the product of two terms: the first is the runoff coefficient for the index flood; the second term is the ratio of the rainfall and flow probabilistic growth factors, i.e. it describes the transformation of the slope and curvature of the flood frequency distribution with respect to the annual maximum rainfall depth frequency distribution.

It is important to note that C_T is not restricted to those values commonly attached to the traditional runoff coefficient as ϕ ($0 < \phi \leq 1$). According to Eq. (5), C_T is a probabilistic factor, which in principle could assume values even greater than 1, as for example when the flood frequency curves exhibit a significantly greater curvature than the corresponding rainfall frequency curves (Franchini et al., 2005).

We call C_L the value of the ϕ coefficient estimated by regression models accounting for the catchment lithology only, calibrated at regional scale against the observed values ϕ_{obs}, as defined by Eq. (3). Then we express the component of the runoff coefficient not explained by the cathment lithology, for a generic return period T, as follows:

$$\Delta C_T = \frac{C_{T,\mathrm{obs}} - C_L}{C_{T,\mathrm{obs}}}. \qquad (6)$$

In the hypothesis that ϕ is fully described by C_L, ΔC_T is only representative of the ratio of the rainfall and flow probabilistic growth factors.

3 Available data

We explore the possibility to explain ΔC_T as a function of the forest cover, in combination with other catchment descriptors that are already employed within the flood assessment procedure based on the rational formula, for a reference return period (T) of 20 years. We perform this analysis by examining the available data concerning 75 gauged

catchments in Central and Southern Italy (Fig. 1): 34 catchments in Toscana, 17 in Lazio, 12 in Campania and 12 in Sicilia. These catchments belong to different rainfall and flood homogeneous regions, as delineated by regional frequency analyses. The climate is quite variable among the study catchments, owing to significant variation in the geographic latitude. Sicilia is characterized by semiarid or dry subhumid climate, with mild, not very rainy winters, and warm and very dry summers. As one proceeds north (toward Campania, Lazio and Toscana), the climate turns to wet subhumid and humid, with a marked seasonality, characterized by very wet winters and dry summers.

The following hydro-morphological parameters have been collected from the available reports and papers (Birtone et al., 2008; Calenda et al., 2003; Di Stefano and Ferro, 2007; Rasulo et al., 2009; Regione Toscana, 2007):

- catchment area (A);

- catchment mean elevation (Z_m) above catchment outlet;

- catchment critical storm duration, assumed equal to the concentration time (t_c);

- catchment critical rainfall depth (h_c) for a reference return period of 20 years ($h_c = h_{A,T=20}(t_c)$);

- catchment fraction with highly permeable lithoid complexes (S_p);

- catchment fraction with forest cover (S_b);

- runoff coefficient of the index flood estimated by regional regression models accounting for the catchment lithology only (C_L);

- observed values of the runoff coefficient, defined according to Eq. (5), for a reference return period of 20 years ($C_{obs} = C_{T=20,obs}$).

We selected the catchment mean elevation (Z_m) above catchment outlet as this is one of the three catchment descriptors appearing in the *Giandotti* empirical formula employed for computing the catchment concentration time (e.g. Brath et al., 2001):

$$t_c = \frac{4\sqrt{A} + 1.5\,L}{0.8\sqrt{Z_m}} \qquad (7)$$

where L (km) is the main river length.

C_L is computed by multiregression equations against the catchment fractions with different degree of permeability, as derived from lithological maps at regional scale. Within the sample catchment set, S_p explains almost 80 % of the overall variability of C_L among the examined catchments. For this reason we selected only S_p among the parameters describing the lithological classes within the study catchments.

The forest cover fraction S_b is evaluated as the average value from land use maps of the same period analysed for

Fig. 1. Italian regions where the outlets of the 75 study catchments are located; from South (dark grey) to North: Sicilia (12 catchments), Campania (12 catchments), Lazio (17 catchments), Toscana (34 catchments).

assessing the observed runoff coefficient, mostly collected from 1960 and 1990 (e.g. Ferro and Porto, 2006). The spontaneous vegetation and land cover in Central and Southern Italy is quite consistent with climatic features and morphological characteristics of the territory. Arid and semiarid zones are characterized by scarce vegetation, which gradually turns into subhumid Mediterranean undergrowth and pasture land, to finally reach the mountain woods of humid and hyperhumid areas (Fiorentino and Iacobellis, 2001).

In the following section we examine the data set to disclose the dependency of $\Delta C = \Delta C_{T=20}$ from the forest cover fraction and the selected set of hydro-morphological parameters.

4 Data mining

The investigation of forest cover influence on ΔC is based on three sequential steps: (1) preliminary data analysis, to explore the hydro-morphological data distributions; (2) cluster analysis, to define catchment hydrological similarities and (3) correlation structure, to quantify possible interconnections between forest cover and ΔC. More details are described in the following sections by providing drivers to identify a new regression model of the runoff coefficient accounting for the forest cover.
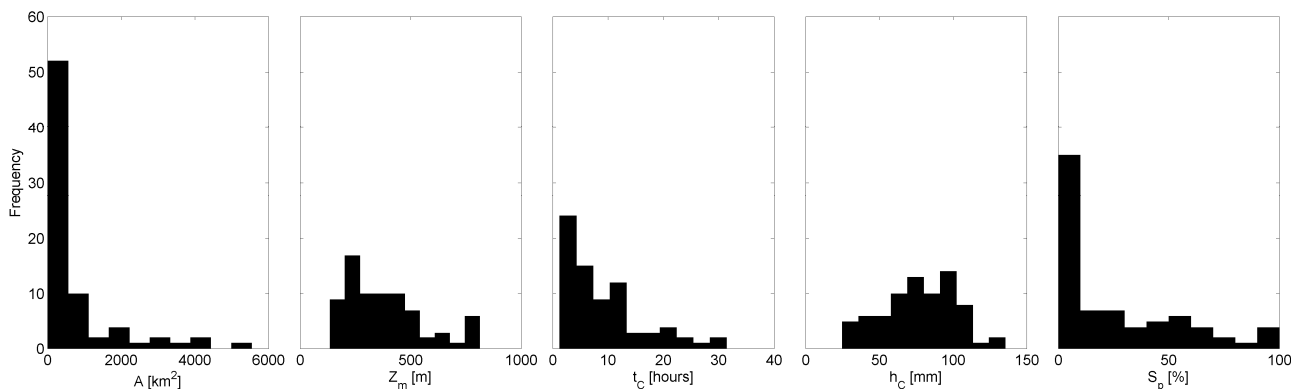
Fig. 2. Histograms of the hydro-morphological parameters: catchment extent (A); elevation above catchment outlet (Z_m); concentration time (t_c); maximum annual rainfall depth within a time interval equal to t_c and a return period of 20 years (hc); catchment fraction with highly permeable lithoid complexes (S_p).

Table 1. Mean (μ) and standard deviation (std) of the hydro-morphological variables for the examined catchments.

	A (km^2)	Z_m (m a.s.l.)	t_c (h)	h_c (mm)	S_p (%)	C_L	C_{obs}
μ	712	383	9.2	76.5	25.5	0.57	0.49
std	1129	177	7.2	25.4	29.6	0.29	0.27

Table 2. Correlation analysis among the hydro-morphological variables for the examined catchments: Spearman rank (r_k) correlation in the lower triangular matrix; corresponding p-values ($\underline{p}$) in upper triangular matrix.

r_k \ $\underline{p}$	A	Z_m	t_c	h_c	S_p	ΔC
A	1	0.003	<0.001	<0.001	0.048	0.216
Z_m	0.332	1	0.230	0.435	0.113	0.076
t_c	0.968	0.14	1	<0.001	0.068	0.115
h_c	0.731	0.091	0.744	1	0.410	0.430
S_p	0.228	0.184	0.211	0.096	1	0.394
ΔC	0.144	−0.205	0.183	0.092	0.099	1

4.1 Preliminary data analysis

The following preliminary analyses have been conducted:

- histogram analysis;

- non-parametric correlation analysis of the hydro-morphological variables;

- dependence analysis of ΔC from each of the hydro-morphological variables and the forest cover fraction.

Histograms show the large variability of the hydro-morphological features of the examined catchments (see Table 1 and Fig. 2). A large number of catchments have an extent smaller than 1000 km^2 and a surface fraction with high permeability (S_p) smaller than 20 %. The distributions of A and t_c are very similar and exhibit a large skweness, while h_c distribution is more uniform as direct consequence on the non-linearity of the rainfall depth-duration curve. Almost 50 % of the catchments are characterized by a concentration time smaller than 8 h: all 12 catchments in Sicily; 4 catchments in Campania; 2 in Lazio and 22 in Toscana. Only 10 catchments (13 %) have a concentration time larger than 18 h: 4 in Campania, 4 in Lazio and 2 in Toscana. The average C_L is 0.57, with a standard deviation of 0.29, while the average C_{obs} is 0.49 with a standard deviation of 0.27 (see Table 1).

The correlation analysis has been performed through the Spearman rank instead of the standard Pearson coefficient to better capture possible non-linear regressions (Wilks, 1995).

Many of the analyzed hydro-morphological variables exhibit significant cross-correlation (see p-values in Table 2), while none of these variables appears significantly correlated with ΔC. As expected, there is a significant positive correlation between A, t_c, and h_c. In fact, smaller catchments corresponds to lower order basins located in mountain areas, characterised by higher slope, smaller concentration time and smaller critical rainfall depth.

We also conducted an explanatory analysis among terns of variables $\Delta C - Y - S_b$, where Y represents one of the selected hydro-morphological variables A, Z_m, t_c, h_c and S_p, taken in turn. In Fig. 3, the contour maps represent the variability of each hydro-morphological variable with respect to the S_b values along the x-axis and the variable ΔC along the y-axis.

The contour maps show that a general dependence applicable to the entire data set cannot be found. These configurations indicate the need to explore the dependence between S_b and ΔC within group of basins exhibiting some hydrological similarities.
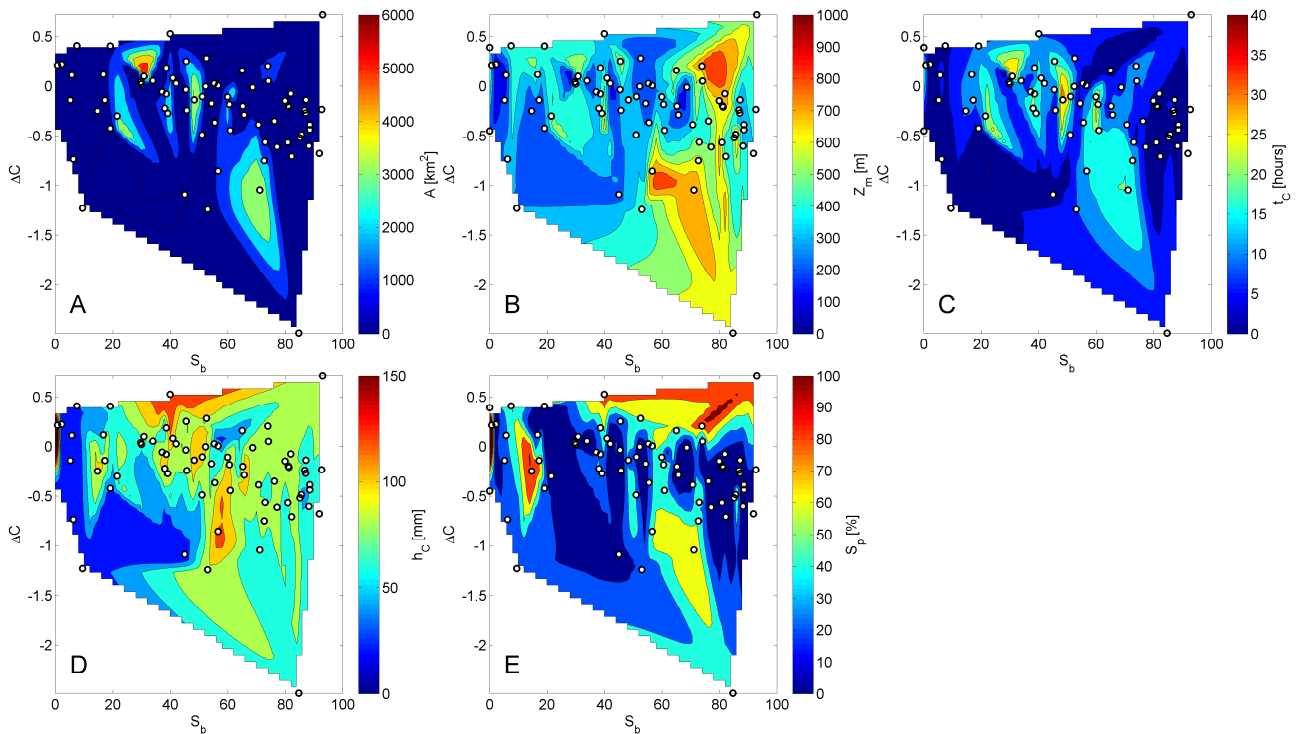
Fig. 3. Contour maps of S_b with respect to ΔC for different hydro-morphological variables: catchment extent (**A**), elevation (**B**), concentration time (**C**), reference rainfall intensity (**D**) and permeable lithoid fraction (**E**). White dots indicate the observed values.

4.2 Cluster analysis

The *k-means* method has been chosen in this study as a simple solution, characterized by short computation times, for the characterization of possible hydrological similarities (MacQueen, 1967). The method performs an unsupervised classification based on the frequency distribution of the hydro-morphological variables. Catchments are grouped according to the *k-means* cluster analysis following two different procedures: (i) clustering based on individual hydro-morphological variables, in order to assess the role of each parameter in the $S_b - \Delta C$ relation; (ii) clustering including all parameters (hereafter indicated as HP case), to explore the effect of the reciprocal interaction among different parameters in the $S_b - \Delta C$ relation. Catchment clusters are identified by maximizing the mean of the silhouette plot (Sh), which is a distance metric based on the squared of the Euclidean distance. Sh indicates the distance of each catchment value within a given cluster from the catchment values belonging to other clusters and ranges from +1 to −1. Sh provides an indirect measure of the inter-cluster separability (Rousseeuw, 1987): +1 suggests a correct catchment clustering, whereas −1 indicates a possible misclassification. Examples of silhouette plots are depicted in Fig. 4.

Table 3 shows the Sh for a number of clusters between 2 and 5, for each of the parameter examined and for the HP case. The values in bold indicate the Sh threshold values

Table 3. Sh values for different clustering levels. In bold Sh threshold value above which further clustering is not acceptable.

	Sh (CL2)	Sh (CL3)	Sh (CL4)	Sh (CL5)
A	**0.906**	0.874	0.79	0.785
Z_m	0.754	**0.756**	0.747	0.729
t_c	0.754	**0.812**	0.71	0.748
h_c	0.747	**0.77**	0.719	0.706
S_p	0.835	**0.838**	0.81	0.795
HP	**0.9**	0.858	0.708	0.569

above which further clustering does not add much to the catchment classification. We identified 2 clusters for the catchment extent A and three clusters for each of the remaining 5 parameters. When we analyzed all parameters (HP), we selected 2 clusters.

Tables 4 and 5 show mean, standard deviation, sample size and range of the parameters within each cluster. The differences between clusters have been also verified by a statistical test on the mean value of the hydro-morphological parameters belonging to each cluster, with 0.01 significance level. All clusters, except in HP case, are significantly different in mean, i.e. the null hypothesis that they belong to the same population can be rejected with a significance level equal to 0.01. The value range of each parameter represented
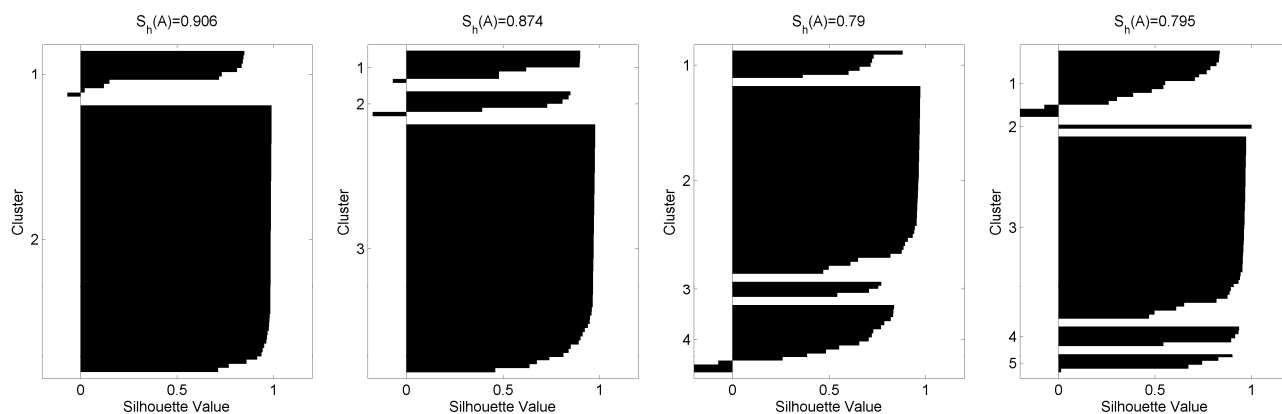
Fig. 4. Example of silhouette plots for cluster analysis.

Table 4. Catchment clustering based on the analysis of single parameters. Statistics of the parameter values in each cluster: mean (μ), standard deviation (std), number of samples (n), and value ranges corresponding to the quantiles 0.05 and 0.95 (Min and Max, respectively).

	CL1					CL2					CL3				
	μ	std	n	Min	Max	μ	std	n	Min	Max	μ	std	n	Min	Max
A	299.98	322.58	64	15.584	980	3114.4	1161.2	11	1921.1	5469.9	–	–	–	–	–
Z_m	735.56	67.082	10	637	813	256	61.475	42	154.77	346.32	463.11	57.115	23	375.47	561.72
t_c	3.858	1.571	40	1.515	6.47	11.965	2.611	25	8.597	16.776	23.616	4.173	10	18.65	31.416
h_c	33.576	5.295	11	24.848	39.617	66.993	8.784	31	52.582	80.232	99.767	11.254	33	87.06	123.09
S_p	87.461	11.176	8	69.801	99.989	46.441	11.127	21	27.731	60.806	5.221	7.634	46	0.01	22.491

Table 5. Catchment clustering based on the analysis of the entire set of parameters (HP). Statistics of the parameter values in each cluster: mean (μ), standard deviation (std), number of samples (n), and value ranges corresponding to the quantiles 0.05 and 0.95 (Min and Max, respectively).

	CL1					CL2				
	μ (CL1)	std (CL1)	n (CL1)	Min	Max	μ (CL2)	std (CL2)	n (CL2)	Min	Max
HP (A)	854.57	1341.2	32	12.1	4014.2	607.25	944.94	43	31.28	3078.7
HP (Z_m)	322.55	129.15	32	154.42	548.35	428.79	195.85	43	196.92	789.45
HP (t_c)	10.103	7.967	32	1.409	27.699	8.519	6.57	43	1.872	24.472
HP (h_c)	69.775	30.291	32	26.607	111.71	81.526	19.882	43	52.981	112.55
HP (S_p)	28.984	27.187	32	0.01	93.485	22.968	31.333	43	0.01	85.436

within each cluster is identified by the quantiles 0.05 and 0.95 of the corresponding sample distributions in each cluster. These ranges do not overlap when the clusters are identified by analyzing one parameter at a time, except for the clusters identified with the parameter A only. When all parameter are considered in the clustering process (HP), the distinction of each cluster is more difficult, since value ranges overlap, as we might expect by examining the corresponding mean and std values (see Table 5).

4.3 Correlation structure

We calculated the Spearman's rank correlation and the corresponding significance level (p-values) between $S_b - \Delta C$ for each combination of parameters and for each cluster identified (Table 6). Significant correlation occurs for the cluster with the largest number of samples for each fixed parameter. This suggests that for those clusters with a limited number of samples, the correlation might be underestimated. For the HP case no significant correlation has been identified.

We also assessed the dependence between ΔC and S_b by linear regression, in order to evaluate the component of the
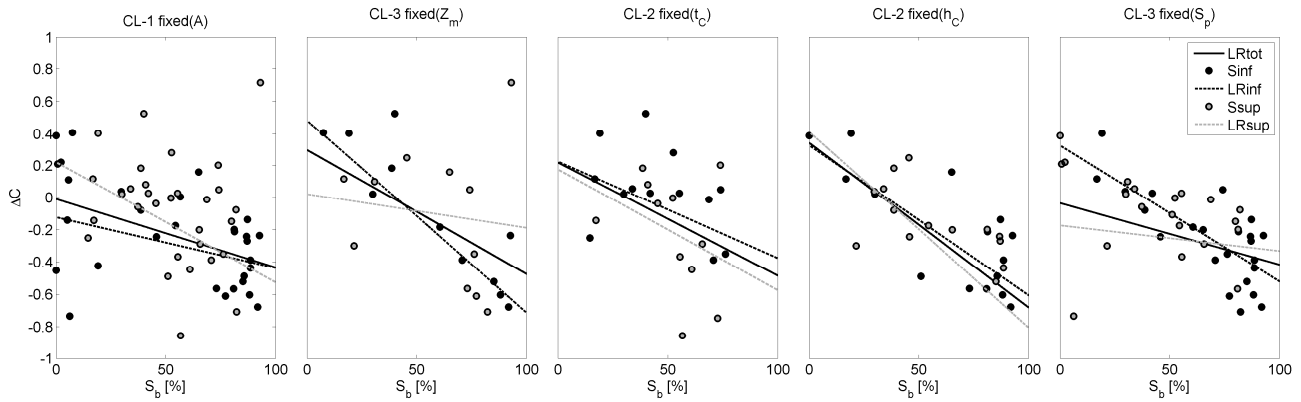
Fig. 5. S_{inf} represents those (S_b, ΔC) couples which values of the corresponding hydro-morphological parameters are below its intra-cluster average. S_{sup} represents those (S_b, ΔC) couples which values of the corresponding hydro-morphological parameters are above its intra-cluster average. Regression lines between ΔC and forest cover fraction for selected clusters of the study catchments: (i) including all (S_b, ΔC) couples belonging to the examined cluster (LRtot); (ii) including only S_{inf} couples (LRinf); (iii) including only S_{sup} couples (LRsup).

Table 6. Spearman's rank correlation $S_b - \Delta C$ and corresponding p-values, within each cluster for a given parameter set. In bold those with p-values below 0.05.

Parameter	r_k			p		
Set	CL1	CL2	CL3	CL1	CL2	CL3
A	**−0.36**	−0.29	NaN	**0.003**	0.386	–
Z_m	0.078	−0.233	**−0.466**	0.838	0.136	**0.026**
t_c	−0.285	**−0.403**	−0.066	0.074	**0.046**	0.864
h_c	−0.318	**−0.636**	−0.203	0.341	**0.000**	0.255
S_p	0.19	−0.271	**−0.474**	0.664	0.233	**0.000**
HP	−0.052	−0.201	–	0.774	0.196	–

Table 7. Results of linear regression analysis: slope (m) and intercept (q) of the regression lines; coefficient of determination (R^2).

		m	q	R^2
Lrtot	CL-1 fixed (A)	−0.004	−0.01	0.065
	CL-3 fixed (Z_m)	−0.007	0.297	0.198
	CL-2 fixed (t_c)	−0.007	0.218	0.116
	CL-2 fixed (h_c)	−0.01	0.342	**0.291**
	CL-3 fixed (r)	−0.01	0.312	**0.223**
	CL-3 fixed (S_p)	−0.003	−0.035	0.087
Lrinf	CL-1 fixed (A)	−0.003	−0.123	0.086
	CL-3 fixed (Z_m)	−0.011	0.482	**0.452**
	CL-2 fixed (t_c)	−0.006	0.224	0.091
	CL-2 fixed (h_c)	−0.009	0.326	**0.496**
	CL-3 fixed (r)	−0.009	0.254	**0.227**
	CL-3 fixed (S_p)	−0.008	0.325	**0.557**
Lrsup	CL-1 fixed (A)	−0.007	0.22	0.079
	CL-3 fixed (Z_m)	−0.002	0.02	0.015
	CL-2 fixed (t_c)	−0.007	0.175	0.121
	CL-2 fixed (h_c)	−0.012	0.411	**0.218**
	CL-3 fixed (r)	−0.01	0.383	**0.224**
	CL-3 fixed (S_p)	−0.001	−0.173	0.009

runoff coefficient which could be explained by S_b. The goodness of fit of the linear models are estimated by the coefficient of determination (R^2). The regression analysis is applied to three different sets: (i) including all (S_b, ΔC) couples belonging to the examined cluster (LRtot); (ii) including (S_b, ΔC) couples which values of the corresponding hydro-morphological parameters are below its intra-cluster average (LRinf); (iii) including (S_b, ΔC) couples which values of the corresponding hydro-morphological parameters are above its intra-cluster average (LRsup). Results are listed in Table 7. Figure 5 shows the computed regression lines for the selected clusters.

The possibility to explain ΔC with S_b is highly variable, particularly in the case all catchments belonging to a cluster are included in the regression analysis (LRtot). Higher R^2 can be gathered if only those catchments in the lower range of the corresponding parameter values are included in the regression analysis (LRinf). The best fitting is obtained for LRinf within CL-3 fixed (Z_m), CL-2 fixed (h_c) and CL-3 fixed (S_p).

The regression lines show that ΔC is always decreasing for increasing forest cover fraction, consistently with the expected result that forest cover contributes to the decrease of flood peaks. The highest linear correlation is observed within CL-2 fixed (h_c). It is also interesting to observe that there is a high linear correlation for all three sub-sets of CL-2 fixed (h_c), with consistent regression coefficients m and q. Moreover, within CL-2 fixed (h_c) it is also possible to observe the highest non-parametric correlation with ΔC (see Table 6).

Table 8. Results of the regression analysis with four different catchment data sets: all catchments; catchments included in the cluster CL-1 fixed (t_c); catchments with an area $A < 500\,\mathrm{km}^2$; catchments with an area $A < 100\,\mathrm{km}^2$.

Data set	$V_{b,T=20}$	95 % Confidence limits		R^2	t-statistics	p-value
	(mm)	Lower value	Upper value			
All	13.9	10.9	16.9	0.357	9.266	<0.0001
CL-1 fixed (t_c)	15.3	12.0	18.6	0.447	9.422	<0.0001
$A < 500\,\mathrm{km}^2$	15.7	11.7	19.8	0.392	7.902	<0.0001
$A < 100\,\mathrm{km}^2$	15.2	11.2	19.1	0.565	7.972	<0.0001

These results suggest the possibility to explore new regression models for predicting the runoff coefficient $C_{T=20}$, accounting for the forest cover S_b, at least within catchments classes which can be considered homogeneous with respect to $h_{A,T=20}$ (t_c).

5 A new simple conceptual model for runoff coefficient assessment

Predicted values of the runoff coefficients from the catchment lithology (C_L) are generally negatively biased with respect to the observed runoff coefficient C_{obs}. Moreover, the absolute difference $C_{obs} - C_L$ tends to be higher for catchments with smaller concentration time, as illustrated in Fig. 6. This suggests that there are larger margins for correcting the prediction of the runoff coefficients in catchments with smaller concentration time, which are generally those basins characterized by larger forest cover fractions, as these catchments are of lower order and higher slope, mostly located in mountainous areas.

Following the results of the previous cluster analysis, we explore the possibility to correct the bias $C_{obs} - C_L$ with a runoff coefficient $C_{T=20}$ expressed by a simple regression model of the forest cover S_b scaled by the critical rainfall depth $h_{A,T=20}$ (t_c):

$$C_{T=20} = C_L - V_{b,T=20}\,\frac{S_b}{h_{A,T=20}\,(t_c)}. \tag{8}$$

According to Eq. (8), the runoff coefficient estimated from the catchment lithology only (C_L) is reduced by a factor $V_{b,T=20}\,S_b/h_c$, which conceptually could be interpreted as an additional abstraction loss of the total rainfall due to the storage capacity of the forested fraction of the catchment, with a specific volume equal to $V_{b,T=20}$. Moreover, recalling Eq. (5), the term $[1 - (V_{b,T=20}\,S_b)/(h_{A,T=20}\,(t_c)\,C_L)]$ describes the ratio $K_{Q,T=20}/K_{R,T=20}$, if we assume that C_L represents the ratio between the corresponding index values.

The optimal value $V_{b,T=20}$ can be assessed by least squares method applied to a linear regression model with the dimensionless ratio $S_b/h_{A,T=20}$ (t_c) as independent variable
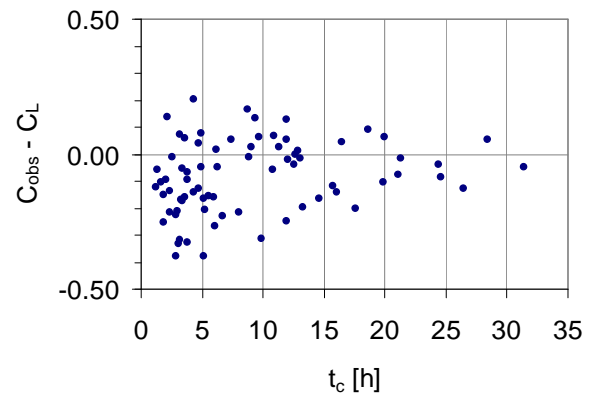

Fig. 6. Differences between C_{obs} and C_L versus the catchment concentration time.

and ($C_{T=20} - C_L$) as dependent variable, exploiting ($C_{obs} - C_L$) as sample data.

Figure 7a shows the fitted linear regression model with the corresponding scatter plot. The standardized residuals appear normally distributed, as depicted in Fig. 7b and c. The optimal $V_{b,T=20}$ results equal to 13.9 mm, with a determination coefficient (R^2) equal to 0.357 (see Table 8). The null hypothesis of no linear correlation is rejected with a p-value below 0.0001. The 95 % confidence interval spans 6 mm around the expected value.

If we restrict the analysis to catchment clusters selected with respect to t_c as illustrated in the previous section, the prediction performance improves with small changes in the optimal $V_{b,T=20}$. For example, as illustrated in Table 8, limiting the analysis to CL-1 fixed (t_c), the optimal $V_{b,T=20}$ is equal to 15.3 mm with a R^2 equal to 0.447.

The prediction performance also improves if we restrict the analysis to smaller catchments, still with slight changes in the predicted $V_{b,T=20}$ (see Table 8): for catchment areas smaller than $500\,\mathrm{km}^2$, we get $V_{b,T=20} = 15.7$ mm and $R^2 = 0.392$; for catchment areas smaller than $100\,\mathrm{km}^2$, we get $V_{b,T=20} = 15.2$ mm and $R^2 = 0.565$.

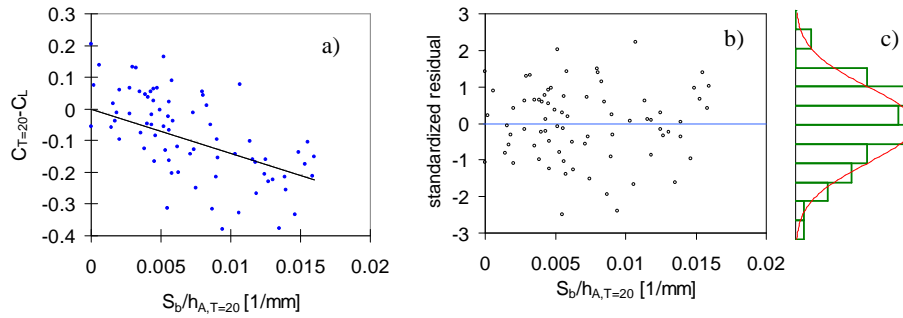Figure 8 compares the observed ($C_{T=20,obs}$) to the predicted runoff coefficients, C_L and $C_{T=20}$. Prediction

Fig. 7. **(a)** Scatter plot and linear regression between $(C_{T=20} - C_L)$ and $S_b/h_{A,T=20}$ (t_c); **(b)** scatter plot of standardized residuals; **(c)** frequency distribution of the standardized residuals and corresponding normal fit.
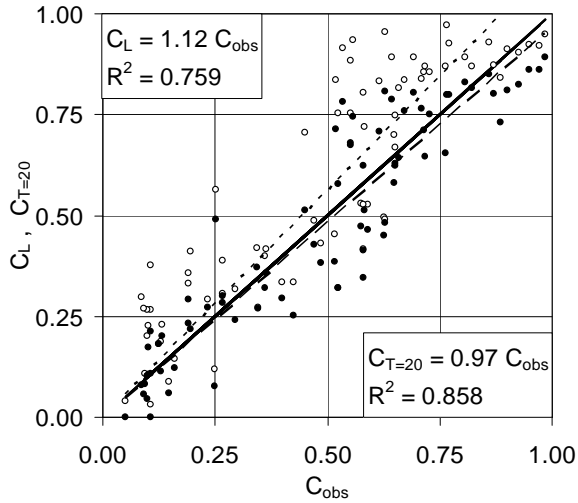

Fig. 8. Comparison between observed and predicted runoff coefficients: open circles, (C_{obs}, C_L); filled circles, $(C_{obs}, C_{T=20})$. The dotted line is the linear regression between C_{obs} and C_L; the dashed line is the linear regression between C_{obs} and C_T. The linear regressions are compared to the continuous line, representing the perfect agreement.

performances are evaluated by computing the bias, the absolute bias (Abias) and the root mean square error (RMSE), as follows:

$$\text{Bias} = \frac{\sum\limits_{i=1}^{N} \text{err}_i}{N}$$

$$\text{Abias} = \frac{\sum\limits_{i=1}^{N} |\text{err}_i|}{N} \qquad (9)$$

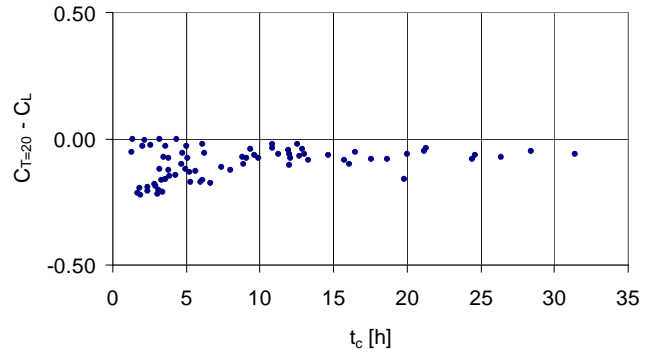$$\text{RMSE} = \sqrt{\frac{\sum\limits_{i=1}^{N} (\text{err}_i)^2}{N}}$$

Fig. 9. Differences between $C_{T=20}$ and C_L versus the catchment concentration time.

where N is the total number of catchment data and err_i represents the deviation between the observed and the predicted runoff coefficients, $\text{err}_i = (C_{obs} - C_{T=20})$ or $\text{err}_i = (C_{obs} - C_L)$. As illustrated in Table 8, all three performance statistics indicate that there is an improvement of the predicted runoff coefficient with $C_{T=20}$. The bias correction is particularly high in catchments with higher $S_b/h_{A,T}$ (t_c) ratios. Figure 9 shows how the (negative) bias correction tends to zero as the catchment concentration time increases. Higher catchment concentration time values correspond to higher $S_b/h_{A,T}$ (t_c) ratios, which belong to small mountainous catchments, with higher slope and higher forest cover, which are characterized by small concentration time and small critical rainfall depth.

6 Discussion

The effect of forest cover on flood regime has been largely studied, but it is still a controversial argument (e.g. Sorriso-Valvo et al., 1995; Robinson, 1989; Robinson et al., 2003; Cognard-Plancq et al., 2001; Cosandey et al., 2005; Bathurst et al., 2011). Despite the public perception that forests reduce flood hazard, there is a large sector of the scientific community asserting that forest cover, although being relevant

within the hydrological cycle and in catchment response to small storms, does not mitigate significantly floods during extreme rainfall events (e.g. Calder, 2007; van Diijk and Keenan, 2007). This opinion is also prompted by influential United Nation Policy documents (e.g. FAO and CIFOR, 2005; Hamilton, 2008), which confine the public perception to a misconception conceived by those who are not directly involved in studying hydrological extreme events, including environmentalists and conservation agencies. Experimental studies show that forest cover reduces the annual catchment discharge as result of increased rainfall interception, increased transpiration and lower soil moisture regime during interstorm periods and higher permeability of forest soil (e.g. Bosch and Hewlett, 1982; Cornish, 1993; Rowe and Pearce, 1994; Stednick, 1996; Fahey and Jackson, 1997; Bruijnzeel, 2004). On the other hand, it is more difficult to assess the impact of forests on floods catchment response response to rainfalls with low frequency, due to lack of experimental data for quantifying the effect of forests on the catchment response to rainfalls with low frequency (Nelson and Chomitz, 2007). Moreover, the effect of forest cover on flood peaks is difficult to be isolated, being the flood discharge influenced by other factors, such as initial catchment conditions, forestry and agricultural activities, road constructions, etc. (Moore and Wondzell, 2005).

Alila et al. (2009) pointed out that some contrasting conclusions about the relation between forests and floods is the result of catchment paired studies, which do not account for the effect of forest cover on the non-linear dependency between magnitude and frequency of floods. Alila et al. (2009) challenged decades of peer reviewed paired watershed study literature, arguing that the experimental design and statistical methods used in paired watershed studies have overlooked a fundamental part of the physics of the relation between forests and floods, namely that forest affects not only the magnitude but most importantly the frequency of floods. Moreover, because of the strong non-linearity of the flood frequency distribution, small changes in the magnitude of extreme floods can translate into large changes in the corresponding return periods (Alila et al., 2009, 2010).

Previous studies paid much attention to the effect of forest cover changes. In fact, forest patterns have been subjected to significant changes worldwide, with different trends, depending specifically on local socio-economic and environmental factors. There are some areas of the world where forest cover has been reducing as results of logging and land claim for agriculture or urban infrastructures. Other areas, such as Mediterranean landscapes, forest patterns are experiencing a significant expansion in the last decades, as consequence of the abandonment of the agricultural lands in marginal areas, mostly located in hilly and mountainous areas, providing space for the natural development of forest (Mazzoleni et al., 2004).

It is important to point out that forest cover change is a source of non-stationarity and its effect on catchment response cannot be analyzed in studies such as the present one. In fact, herein we explore the role of the forest cover on the runoff coefficient as defined in a regional flood frequency analysis, based on the application of the rational formula coupled with a regional model of the annual maximum rainfall depths, which inherently assumes that catchments are in stationary conditions.

Within this framework, the runoff coefficient looses its original meaning of output-input ratio of a rainfall-runoff model, as it occurs in the original interpretation of the rational formula, while it assumes the role of a probabilistic factor, which describes the ratio of the flood peak to the maximum annual rainfall depth for a given return period, i.e. the ratio between two random variables paired by the same frequency as dictated by the corresponding cumulative probability distributions. Thus, the runoff coefficient affects not only the first moment of the flood frequency distribution, as it occurs within the regional flood frequency analysis according to the index flood method, but also the higher moments with respect to those of the maximum annual rainfall depth.

We explored the possibility to exploit the forest cover fraction (S_b) to explain the part of the runoff coefficient which is not already described by C_L, which is the runoff coefficient defined on the basis of the catchment lithology only and employed for identifying the index flood. The *k-means* cluster analysis evidenced the possibility to explain an additional component of the runoff coefficient ($C_{obs} - C_L$) by the forest cover fraction S_b, scaled with the corresponding critical areal rainfall depth, $h_{A,T}$ (t_c). In this analysis, we avoid considering other potential catchment descriptors, beside those already employed for computing the index flood, in order to keep the overall flood assessment procedure as simple as possible. We proposed a linear function of the $S_b/h_{A,T}$ (t_c) ratio for correcting C_L, with just one additional empirical parameter, $V_{b,T}$. Although the interpretation of $V_{b,T}$ as an additional abstraction loss attached to the forest fraction for different return periods might be suggestive from a conceptual point of view, it is important to keep in mind that $V_{b,T}$ is an empirical parameter which values are strictly valid for the region where have been estimated, based on the local rainfall and discharge frequency distributions. Moreover, beside the ratio $S_b/h_{A,T}$ (t_c), other factors might contribute to the observed difference ($C_{obs} - C_L$), as this is subjected to several error sources, such as: the model structure, resembling the rational formula, adopted for describing the correspondence between the annual maximum rainfall depth and the peak flow for a given return period, as depicted by Eq. (5); the assumption that a catchment critical storm duration can be identified by a catchment morphological parameters as in Eq. (7), assumed representative of the catchment concentration time; the regional model for representing the maximum annual rainfall; the approach employed for defining the areal rainfall reduction factor κ_A.

The regression model suggests that the effect of forest cover decreases with catchment extent. In fact, for a given

Table 9. Performance statistics of the predicted runoff coefficients by exploiting: the catchment lithology only (C_L); both the catchment lithology and the forest cover ($C_{T=20}$).

Predicted values	Bias	RMSE	Abias	$(\text{err}_i)_{\min}$	$(\text{err}_i)_{\max}$
C_L	-0.08	0.16	0.13	-0.38	0.20
$C_{T=20}$	0.01	0.10	0.08	-0.25	0.24

$$S_b \frac{C_L h_{A,T}(t_c)}{V_{b,T}}$$

Fig. 10. Changes in return period of a 20 year return period flood as function of the forest cover fraction S_b scaled by the dimensionless number $C_L\, h_{A,T}(t_c)/V_{b,T}$ for a generic catchment with a flood probabilistic growth factor as defined for the Campania region.

fraction of forest cover (S_b), as the catchment size increases, the critical rainfall depth $h_{A,T}(t_c)$ increases and therefore the relative contribute of the forest cover on the runoff coefficient reduces. This result is consistent with the observation that the effect of forest cover decreases in larger catchments, i.e. in catchments characterized by larger concentration time, where forest cover is more fragmented and the flood response is dominated by other hydrological and hydraulic processes (Blöschl et al., 2007).

It is interesting to observe that the prediction performance of the runoff coefficient also improves without a significant change in the estimated $V_{b,T=20}$, if we exclude from the analysis those catchments with an extent larger than those limit values (e.g. 500 km^2), above which a model structure as the rational formula is not considered appropriate.

From a flood risk perspective, according to the proposed model for predicting the runoff coefficient, neglecting the effect of forest cover can correspond to a significant change in the return period attached to a predicted flood peak value. Figure 10 shows the relative changes in the estimated return period of a 20 year return period flood peak when neglecting the contribute of the forest cover, as function of the forest cover fraction S_b scaled by the dimensionless number $C_L\, h_{A,T}(t_c)/V_{b,T}$, for a generic catchment with a dimensionless flood growth factor as defined for the Campania region (Rossi and Villani, 1992). For example, neglecting the effect of a 30 % (scaled) forest cover fraction can lead to underestimating the corresponding return period by almost 3 times, i.e. attributing a return period of 7 years to a flood peak of 20 years. This figure shows that relatively small changes in the estimated flood peak magnitude can translate into surprisingly large changes in their return periods, as a consequence of the strong non-linearity of the flood frequency curve (Alila et al., 2010).

The range of the uncertainty bounds attached to the estimated $V_{b,T=20}$ value is quite small (6 mm) if compared to the uncertainty attached to the critical areal rainfall depth $h_{A,T}(t_c)$. The prediction performance of the observed runoff coefficient improves if $V_{b,T=20}$ is calibrated for catchment clusters, as those indentified by the k-means procedure with respect to the catchment concentration time t_c. This suggests the possibility of regionalizing the parameter $V_{b,T}$ with respect to selected catchment descriptors, by also examining the spatial distribution of the selected catchment cluster.
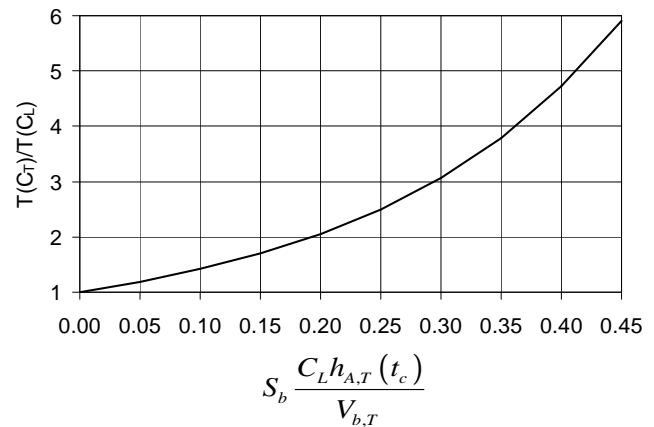
7 Conclusions

In a regional flood frequency analysis based on the application of the rational formula coupled with a regional model of the annual maximum rainfall depths, the runoff coefficient assumes the role of a probabilistic factor, being defined by the product of two components: the first is the runoff coefficient of the index flood; the second is the ratio of the rainfall and flow probabilistic growth factors and is dependent from the return period.

In this paper we evaluated the effect of forest cover on the second component, provided that the first component is assessed by the catchment lithology only.

The results of a k-means cluster analysis applied to a data set of 75 catchments distributed from South to Central Italy, evidenced that the second component of the runoff coefficient can be partly explained by the forest cover fraction, scaled with the corresponding critical areal rainfall depth. Thus, we proposed a linear regression model to improve the prediction of the runoff coefficient, exploiting the ratio of the forest cover to the catchment critical rainfall depth as dependent variable, with just one additional empirical parameter. The proposed regression enables a significant bias correction of the runoff coefficient, particularly for those small mountainous catchments, characterised by larger forest cover fraction and lower critical rainfall depth.

In this study we restricted our analysis to a reference return period of 20 years. Preliminary investigations limited to the Toscana region, show that the parameter regression model is not sensitive to the return period as compared with the uncertainty attached to the estimated parameter. In this case, for higher return periods, the forest cover fraction, as introduced in the proposed regression model, does not provide any additional information on the ratio between the probabilistic growth factors of the flood peak and the maximum

annual rainfall depth, respectively. This does not necessarily mean that the effect of forest cover on flood peaks is not relevant for higher return periods. In fact, owing to the correlation between the forest cover fraction and the other hydro-morphological parameters employed in the examined flood regionalization procedure, the forest cover can also indirectly influence the ratio of the corresponding index values and thus the position of the entire flood frequency distribution (Alila et al., 2010).

Further studies will be addressed to verify, over a larger data-set, both the efficiency of the proposed regression model for different return periods and the sensitivity of the parameter $V_{b,T}$, in order to develop a regional model of its spatial variability to be integrated into the regional flood assessment procedure.

References

Alila, Y., Kuraś, P. K., Schnorbus, M., and Hudson, R.: Forests and floods: a new paradigm sheds light on age-old controversies, Water Resour. Res., 45, W08416, doi:10.1029/2008WR007207, 2009.

Alila, Y., Hudson, R., Kuraś, P. K., Schnorbus, M., and Rasouli, K.: Reply to comment by Jack Lewis et al. on "Forests and floods: A new paradigm sheds light on age-old controversies", Water Resour. Res., 46, W05802, doi:10.1029/2009WR009028, 2010.

Bathurst, J. C., Birkinshaw, S. J., Cisneros, F., Fallas, J., Iroumé, A., Iturraspe, R., Gaviño Novillo, M., Urciuolo, A., Alvarado, A., Coello, C., Huber, A., Miranda, M., Ramirez, M., and Sarandón, R.: Forest impact on floods due to extreme rainfall and snowmelt in four Latin American environments 1: Field data analysis, J Hydrol., 400, 281–291, 2011.

Becciu, G., Brath, A., and Rosso, R.: A physically based methdology for regional flood frequency analysis, in: Engineering Hydrology, edited by: Kuo, C. Y., Am. Soc. Civil Engng, New York, USA, 461–466, 1993.

Birtone, M., Ferro, V., and Pomilla, S.: A contribute to the application of the kinematic model in small catchments of Sicily, Italia Forestale Montana, 6, 505–518, 2008.

Blöschl, G., Ardoin-Bardin, S., Bonell, M., Dorninger, M., Goodrich, D., Gutknecht, D., Matamoros, D., Merz, B., Shand, P., and Szolgay, J.: At what scales do climate variability and land cover change impact on flooding and low flows?, Hydrol. Process., 21, 1241–1247, 2007.

Bocchiola, D., De Michele, C., and Rosso, R.: Review of recent advances in index flood estimation, Hydrol. Earth Syst. Sci., 7, 283–296, doi:10.5194/hess-7-283-2003, 2003.

Bosch, J. M. and Hewlett, J. D.: A review of catchment experiments to determine the effect of vegetation changes on water yield and evapotranspiration, J. Hydrol., 55, 3–23, 1982.

Brath, A., Castellarin, A., Franchini, M., and Galeati, G.: Estimating the index flood using indirect methods, Hydrolog. Sci. J., 46, 399–418, 2001.

Bruijnzeel, L. A.: Hydrological functions of tropical trees: not seeing the soil for the trees?, Agr. Ecosyst. Environ., 104, 185–228, 2004.

Calder, I. R.: Forests and water – ensuring forest benefits outweigh water costs, Forest Ecol. Manage., 251, 110–120, doi:10.1016/j.foreco.2007.06.015, 2007.

Calenda, G., Mancini, C. P., Cappelli, A., Gaudenti, R., and Lastoria, B.: Studi per l'aggiornamento del piano stralcio per l'assetto idrogeologico: relazione Tecnica – rapporto finale, Regione Lazio – A. B. R., available at: http://www.abr.lazio.it/PAI.pdf/Microsoft%20Word%20-%20ABR%20PAI%20rel%20finale%203.pdf, (last access: 5 October 2011), 2003.

Cannarozzo, M., D'Asaro, F., and Ferro, V.: Regional rainfall and flood frequency analysis for Sicily using the two component extreme value distribution, Hydrolog. Sci. J., 40, 19–42, 1995.

Cognard-Plancq, A., Marc, V., Didon-Lescot, J., and Normand, M.: The role of forest cover on stream?ow down sub-Mediterranean mountain watersheds: a modelling approach, J. Hydrol., 254, 229–243, 2001.

Cornish, P. M.: The effects of logging and forest regeneration on water yields in a moist eucalypt forest in New South Wales, J. Hydrol., 242, 43–63, 1993.

Cosandey, C., Vazken, A., Martin, C., Didon-Lescot, J. F., Lavabre, J., Folton, N., Mathys, N., and Didier, R.: The hydrological impact of the mediterranean forest: a review of French research, J. Hydrol., 301, 235–249, doi:10.1016/j.jhydrol.2004.06.040, 2005.

Cunanne, C.: Methods and merits of regional flood frequency analysis, J. Hydrol., 100, 269–290, 1988.

Dalrymple, T.: Flood frequency analyses, US Geol. Surv. Wat. Supply Pap. 1543-A, 1960.

Di Stefano, C. and Ferro, V.: A contribute to the application of the rational formula with a probabilistic approach in small catchments of Sicily, Quaderni di Idronomia Montana, 27, 389–399, 2007.

Fahey, B. D. and Jackson, R. J.: Environmental effects of forestry at Big Bush forest, South Island, New Zealand, I: changes in water chemistry, J. Hydrol. New Zealand, 36, 43–71, 1997.

FAO and CIFOR: Forests and floods: drowning in fiction or thriving on facts?, RAP Publication 2005/03, Bangkok, Thailand, FAO Regional Office for Asia and the Pacific, 2005.

Ferro, V. and Porto, P.: Flood frequency analysis for Sicily (South Italy), J. Hydrol. Eng.-ASCE, 11, 110–122, 2006.

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.

Franchini, M., Hashemi, A. M., and O'Connell, P. E.: Climatic and basin factors affecting the flood frequency curve: PART II - A full sensitivity analysis based on the continuous simulation

approach combined with a factorial experimental design, Hydrol. Earth Syst. Sci., 4, 483–498, doi:10.5194/hess-4-483-2000, 2000.

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.

French, R., Pilgrim, D. H., and Laurenson, E. M.: Experimental examination of the rational method for small rural catchments, Civ. Eng. Trans. Inst. Engrs. Aust. CE, 16, 95–102, 1974.

Gottschalck, L. and Weingartner, R.: Distribution of peak flow derived from a distribution of rainfall volume and runoff coefficient, and a unit hydrograph, J. Hydrol., 208, 146–162, 1998.

Gupta, V. K. and Waymire, E.: Multiscaling properties of spatial rainfall and river flow distributions, J. Geophys. Res., 95, 1999–2009, 1990.

Hamilton, L.: Forests and water. A Thematic study prepared in the framework of the Global Forest Resources Assessment 2005, FAO Forestry Paper 155, Rome, 2008.

Hashemi, A. M., Franchini, M., and O'Connell, P. E.: Climatic and basin factors affecting the flood frequency curve: PART I - A simple sensitivity analysis based on the continuous simulation approach, Hydrol. Earth Syst. Sci., 4, 463–482, doi:10.5194/hess-4-463-2000, 2000.

Hiemstra, L. A. V. and Reich, R. M.: Engineering judgment and small area flood peaks, Hydrology Paper 191967, Colorado State University, Fort Collins, 1967.

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.

IH – Institute of Hydrology: Flood Estimation Handbook, IH, Wallingford, UK, 1999.

Kjeldsen, T. R. and Jones, D. A.: Predicting the index flood in ungauged UK catchments: On the link between data-transfer and spatial model error structure, J. Hydrol., 387, 1–9, 2010.

MacQueen, J. B.: Some methods for classification and analysis of multivariate observations, Proceedings of Fifth Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, 281–297, 1967.

Maidment, D. R.: Handbook of Hydrology, McGraw-Hill, New York, 1993.

Mazzoleni, S., Di Pasquale, G., and Mulligan, M.: Conclusion: Reversing the Consensus on Mediterranean Desertification, in: Recent Dynamics of Mediterranean Vegetation Landscape, edited by: Mazzoleni, S., Di Pasquale, G., Mulligan, M., Di Martino, P., and Rego, F., Wiley, London, 281–286, 2004.

Mertz, R. and Blöschl, G.: A process typology of regional floods, Water Resour. Res., 36, 469–482, doi:10.1029/2002WR001952, 2003.

Moore, R. D. and Wondzell, S. M.: Physical hydrology and the effects of forest harvesting in the Pacific Northwest: a review, J. Am. Water Res. Assoc., 41, 763–784, 2005.

Nelson, A. and Chomitz, K. M.: The forest-hydrology-poverty nexus in Central America: an heuristic analysis, Environ. Develop. Sustainabil., 9, 369–385, 2007.

Pilgrim, D. H.: Development of the rational method for flood design for small rural basins in Australia, in: Channel Flow and Catchment Runoff, edited by: Yen, B. C., Department of Civil Engineering, University of Virginia, 1989.

Rasulo, G., Del Giudice, G., and Calcaterra, D.: Flood index assessment. Link between the runoff coefficient and the soil characterisation according to SCS, L'Acqua 2/2009, 37–44, 2009.

Regione Toscana: Aggiornamento e sviluppo del sistema di regionalizzazione delle portate di piena in Toscana – AlTo, available at: http://www.rete.toscana.it/sett/pta/suolo/difesa_suolo/ alto/index.htm, (last access: 5 October 2011), 2007.

Robinson, M.: Small catchment studies of man's impact on flood flows: agricultural drainage and plantation forestry, FRIENDS in Hydrology, IAHS Publ., 187, 299–308, 1989.

Robinson, M., Cognard-Plancq, A.-L., Cosandey, C., David, J., Durand, P., Fürer, H.-W., Hall, R., Hendriques, M. O., Marc, V., McCarthy, R., McDonnell, M., Martin, C., Nisbet, T., O'Dea, P., Rodgers, M., and Zollner, A.: Studies of the impact of forests on peak flows and baseflows: a European prospective, Forest Ecol. Manage., 186, 85–97, 2003.

Rossi, F. and Villani, P.: Regional methods for flood estimation, in: Coping with floods, edited by: Rossi G., Harmancioglu, N., and Yevjevich, V., NATO-ASI Series, Kluwer Academic Publ., 135–169, 1992.

Rousseeuw, P. J.: Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, J. Comput. Appl. Math., 20, 53–65, 1987.

Rowe, L. K. M. and Pearce, A. J., Hydrology and related changes after harvesting native forest catchments and establishing Pinus radiata plantations, Part 2 the native forest water balance and changes in streamflow after harvesting, Hydrol. Process., 8, 281–297, 1994.

Schaake, J. C., Geyer, J. C., and Knapp, J. P.: Experimental examination of the rational method, J. Hydraul. Div., 121, 353–370, 1967.

Sorriso-Valvo, M., Bryan, R. B., Yair, A., Iovino, F., and Antronico, L.: Impact of afforestation on hydrological response and sediment production in a small Calabrian catchment, Catena, 25, 89–104, 1995.

Stedinger, J. R., Vogel, R. M., and Foufoula-Georgiou, E.: Frequency analysis of extreme events, in: Handbook of Hydrology, edited by: Maidment, D., McGraw-Hill, New York, 1993.

Stednick, J. D.: Monitoring the effects of timber harvest on annual water yield, J. Hydrol., 176, 79–95, 1996.

Van Dijk, A. I. J. M. and Keenan, R.: Planted forests and water in perspective, Forest Ecol. Manage., 251, 1–9, doi:10.1016/j.foreco.2007.06.010, 2007.

Wilks, D. S.: Statistical methods I the atmospheric sciences – an introduction, International Geophysics Series, Vol. 59, Academic Press, 467 pp., 1995.