

# Flood quantiles estimation based on theoretically derived distributions: regional analysis in Southern Italy

V. Iacobellis<sup>1</sup>, A. Gioia<sup>1</sup>, S. Manfreda<sup>2</sup>, and M. Fiorentino<sup>2</sup>

<sup>1</sup>Dipartimento di Ingegneria delle Acque e di Chimica, Politecnico di Bari, Italy

<sup>2</sup>Dipartimento di Ingegneria e Fisica dell'Ambiente, Università degli Studi della Basilicata, Italy

Received: 30 July 2010 – Revised: 4 January 2011 – Accepted: 10 January 2011 – Published: 7 March 2011

**Abstract.** A regional probabilistic model for the estimation of medium-high return period flood quantiles is presented. The model is based on the use of theoretically derived probability distributions of annual maximum flood peaks (DDF). The general model is called TCIF (Two-Component IF model) and encompasses two different threshold mechanisms associated with ordinary and extraordinary events, respectively. Based on at-site calibration of this model for 33 gauged sites in Southern Italy, a regional analysis is performed obtaining satisfactory results for the estimation of flood quantiles for return periods of technical interest, thus suggesting the use of the proposed methodology for the application to ungauged basins. The model is validated by using a jack-knife cross-validation technique taking all river basins into consideration.

## 1 Introduction

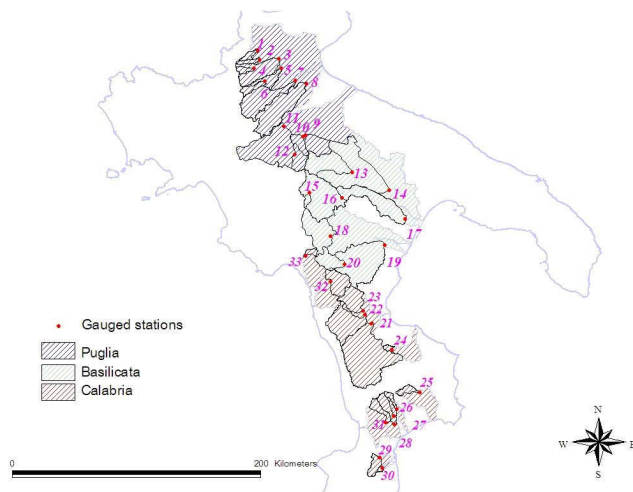
Estimating the flood peaks of an assigned return period is required in several environmental and engineering applications. The methods for predicting flood peaks can be broadly divided into statistically and physically based methods and, obviously, there is not a sharp division between the two groups. Among the first are those that perform the statistical regional analysis of annual maximum flood series (AMFS). These often suffer from the poor availability of long and reliable time series and try, not without difficulty, to exploit physical information within this context. The second category includes distributed or semi-distributed models able to account for an accurate representation of the flood processes. Also these models often suffer from a lack of information necessary for their implementation, calibration and valida-

tion. Nevertheless, investigation efforts in both categories of models, has lead to important advances of knowledge that help to indicate factors depending on climate, soil and vegetation as signatures for the identification of hydrological heterogeneity and similarity. In this context, regional analysis is useful for analyzing, on a regional scale, the behaviour of some key parameters of flood frequency distribution, with the aim of reducing the uncertainty in flood prediction in ungauged basins (Sivapalan et al., 2003).

The limitations of probabilistic and other operational models, such as a regression relation, a polynomial fit, a transfer function, etc. have been already discussed, for example, by Klemes (1982). Eagleson (1972) was the first to propose an alternative approach based on the use of theoretically derived distribution of flood frequency (DDF). DDFs allow the evaluation of the  $T$ -year flood, where  $T$  is the return time, accounting for the observed rainfall probability distribution and exploiting a rainfall-runoff model parameterized by means of geomorphological information. This approach has been exploited by several authors (e.g. Haan and Edwards, 1988; Raines and Valdes, 1993; Kurothe et al., 1997; Gottschalk and Weingartner, 1998; Goel et al., 2000; Iacobellis and Fiorentino, 2000; De Michele and Salvadori, 2002; Franchini et al., 2005). The original approach was widely developed, nevertheless its application has always been devoted to recognizing the main hydrological characteristics controlling flood frequency and to analyzing the distribution of floods at-site. In this paper, a regional analysis is performed on the parameter values of a theoretical probability distribution (recently applied to southern Italian river basins) which is based on the concept that the partial (source) area a contributing to surface runoff is a main random variable of the model. The aforesaid model was proposed by Iacobellis and Fiorentino (2000), and is based on the IF distribution which was generalized by the TCIF distribution (Gioia et al., 2008), assuming that the infiltration process may be produced through two different mechanisms.



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



**Fig. 1.** Investigated basins in Southern Italy.

Thus, we propose a complete regional application of DDFs which provides an operational procedure for prediction in ungauged basins in a large region of Southern Italy. Results are relevant because this approach allows exploiting the robustness of regional analysis and the potential of DDFs, whose parameters may be controlled by several basin features such as geology, geomorphology, soil-use, etc.

An application of the IF model to basins of Puglia and Basilicata, respectively the northern and central regions of Southern Italy (see Fig. 1), was proposed by Fiorentino and Iacobellis (2001), which also provided a preliminary investigation of the regional variability of some key parameters. The TCIF model has been tested by Gioia et al. (2008) only in highly-skewed basins of Puglia, Basilicata and Calabria (southern region). The application was extended to 33 river basins in Southern Italy by Fiorentino et al. (2011), who used the numerical procedure suggested by Iacobellis et al. (2010) for the estimation of parameters and best-fit selection between the IF and TCIF nested distribution, and focused on the regional behaviour of the runoff thresholds. We show results of the regional analysis of parameter estimates and provide relationships for the evaluation of flood quantiles in ungauged basins.

Iacobellis et al. (2010) divided the model parameters into two categories: those dependent on information other than AMFS (among these are all parameters dependent on precipitation), and parameters calibrated by means of AMFS. Estimates of parameters of the first group are based on a priori information, thus, our attention here is given to the investigation of regional relationships between the calibrated parameters and physical (geomorphoclimatical) descriptors.

First insights into the role played by geomorphoclimatical descriptors were gained, in this context, by Fiorentino and Iacobellis (2001); they highlighted the role of climatic and geological factors. More details were provided by Fiorentino

et al. (2002) who used, among others, the runoff coefficient ( $C_{\text{runoff}}$ ) proposed by De Smedt et al. (2000), the coefficient of variation of the wetness index proposed by Beven and Kirkby (1979) and the basin Normalized Difference Vegetation Index (NDVI) averaged on an annual scale. For the purpose of this paper we have also considered other descriptors that are introduced in Sect. 3.

Another important issue which we investigated in order to find reliable regional relationships between model parameters and geomorphoclimatical descriptors regards the partitioning of river basins in homogeneous groups or sub-regions, within which the best regression equations are provided. Moreover, we tested the model in ungauged basins by means of a jack-knife cross-validation technique. In order to focus attention on the regional analysis, we report the main model features in Appendix A and the procedure for the estimation of the first group of parameters in Appendix B.

## 2 Case studies and probabilistic model

The regional analysis includes 33 gauged basins in Southern Italy (see Fig. 1) differing in climate, geomorphology, vegetation coverage, soil type and permeability. Some important features of these basins are reported in Table 1, where  $A$  is the basin area,  $\mu$ ,  $C_v$ ,  $C_s$  and  $N$  are, respectively the mean, the coefficient of variation, the coefficient of skewness and the sample size of the observed AMFS,  $I = (P - E_p)/E_p$  is the Thornthwaite climatic index (Thornthwaite, 1948), which compares annual precipitation  $P$  and annual potential evapotranspiration  $E_p$ . The climatic index distinguishes, in general, between dry ( $I < 0$ ) and humid ( $I > 0$ ) basins, the minimum value observed in the study area is  $-0.28$  (*semi-arid*) in Puglia while the maximum is  $1.66$  (*hyper-humid*) in Calabria. The values of basin area range from  $15$  to  $1657 \text{ km}^2$  and the skewness coefficient  $C_s$  ranges from  $0.08$  to  $3.18$ .

The physical characteristics of the basins investigated are described in Table 2 using a list of descriptors that will be used in the regional analysis. They can be grouped in factors depending on:

- basin geomorphology: basin area  $A$ , total length of the stream network  $L$ , fraction of floodplane areas  $Flp$ , average slope of the basin  $Sl$ , wetness index distribution  $P(WI)$ , mean and standard deviation of  $WI$ ;
- basin geology: permeability index  $\psi$ ;
- basin landcover: % of Forest in basin area ( $Fst$ ), % of Grass ( $Gr$ ) in basin area, annual Normalized difference vegetation index (NDVI);
- basin hydrology:  $C_{\text{runoff}}$  depending on soil type, landcover and local slope,  $CN$  depending on soil type and landcover.

**Table 1.** Main features of the investigated basins.

No.	River basin	Sub-region	$A$ (km <sup>2</sup> )	$\mu$	$C_v$	$C_s$	$N$	$I$
1	Santa Maria at Ponte Lucera Torremaggiore	Puglia	58	18.44	0.92	0.99	15	-0.28
2	Triolo at Ponte Lucera Torremaggiore	Puglia	56	35.44	0.70	0.45	16	-0.25
3	Salsola at Ponte Foggia San Severo	Puglia	455	76.14	0.54	0.28	40	-0.27
4	Casanova at Ponte Lucera Motta	Puglia	57	26.53	0.82	1.32	16	-0.14
5	Celone at Ponte Foggia San Severo	Puglia	233	45.70	0.72	2.43	39	-0.24
6	Celone at San Vincenzo	Puglia	92	31.82	0.61	1.27	15	-0.06
7	Cervaro at Incoronata	Puglia	539	215.80	0.58	0.64	53	-0.19
8	Carapelle at Carapelle	Puglia	715	283.74	0.57	1.34	36	-0.23
9	Venosa at Ponte Sant' Angelo	Puglia	263	55.84	1.18	2.26	34	-0.17
10	Arcidiaconata at Ponte Rapolla Lavello	Puglia	124	45.06	0.64	0.85	32	-0.04
11	Ofanto at Rocchetta Sant' Antonio	Puglia	1111	456.62	0.57	0.46	52	0.16
12	Atella at Ponte sotto Atella	Puglia	176	61.27	0.57	0.96	45	0.17
13	Bradano at Ponte Colonna	Basilicata	462	201.56	0.76	1.21	32	-0.08
14	Bradano at San Giuliano	Basilicata	1657	507.06	0.79	1.03	17	-0.17
15	Basento at Pignola	Basilicata	42	34.86	0.43	1.13	28	0.7
16	Basento at Gallipoli	Basilicata	853	352.61	0.63	2.25	38	0.28
17	Basento at Menzena	Basilicata	1382	400.63	0.63	1.57	24	0.08
18	Agri at Tarangelo	Basilicata	511	189.24	0.49	1.19	32	0.47
19	Sinni at Valsinni	Basilicata	1140	554.91	0.56	2.42	22	0.57
20	Sinni at Pizzutello	Basilicata	232	232.63	0.35	0.82	17	1.26
21	Crati at conca	Calabria	1339	441.42	0.52	1.44	31	0.61
22	Esaro at La musica	Calabria	520	328.84	0.82	1.64	19	0.77
23	Coscile at Camerata	Calabria	285	80.34	0.74	1.65	29	0.65
24	Trionto at Difesa	Calabria	32	8.73	1.09	3.18	16	0.90
25	Tacina at Riviotto	Calabria	79	81.16	1.27	2.79	25	1.43
26	Alli at Orso	Calabria	46	16.66	0.72	2.74	47	1.26
27	Melito at Olivella	Calabria	41	17.18	0.62	1.42	16	0.72
28	Corace at Grascio	Calabria	182	151.65	0.70	1.83	38	0.90
29	Ancinale at Razzona	Calabria	116	82.35	0.73	1.34	50	1.34
30	Alaco at Mammone	Calabria	15	13.61	0.75	1.76	19	1.66
31	Amato at Marino	Calabria	113	79.19	1.18	2.43	26	0.86
32	Lao at piè di Borgo	Calabria	280	214.31	0.59	0.98	24	1.16
33	Noce at la Calda	Calabria	43	30.68	0.41	0.08	30	1.58

Descriptors were obtained from the following sources: the Corine Land-Cover 2000 map, geological maps (scale 1:50 000), DEM (cellsize 90 m), pedological maps produced at a scale of 1:100 000, NDVI images from AVHRR sensor on NOAA mission (with a resolution of 1.1 Km and acquired in 1998).

All geomorphological descriptors have been derived in a GIS environment. In particular, we exploited the distribution of the topographic index which is a measure of basin storage capacity (Manfreda and Fiorentino, 2008; Manfreda, 2008). Two indices are certainly new in this context: one is the probability of finding a topographic index  $WI > \mu_{WI} + \sigma_{WI}$  in a basin and the second is the portion of a basin assumed as flood prone defined with the procedure proposed by Manfreda et al. (2011). In order to give a description of these two adopted indices, we reported in Fig. 2 the portion of the basin with  $WI > \mu_{WI} + \sigma_{WI}$  and the flood prone areas for the Crati basin at Conca (Calabria).

## 2.1 Model description and parameter estimation

The TCIF distribution was developed in analogy with the well-known TCEV distribution (Rossi et al., 1984), with the aim of exploiting all available information regarding physical processes, such as rainfall distribution and river basins features, that significantly affect the Annual Maximum Flood Series (AMFS). The TCIF model stems from the consideration that runoff generation occurs when rainfall intensity and/or rainfall depth exceed given thresholds. Different runoff thresholds may be due to different values of permeability and soil moisture dynamics, thus providing a new phenomenological interpretation for ordinary and extraordinary components of the AMFS. The TCIF distribution is a generalization of the IF distribution (Iacobellis and Fiorentino, 2000) previously developed on the basis of a single runoff threshold leading to a single (ordinary) component of the flood distribution. This new model is based on the

**Table 2.** Descriptors adopted for the regional parameter estimation.

No.	A (km <sup>2</sup> )	$\psi$	Flp	Sl	C <sub>runoff</sub>	$P(WI > \mu_{WI} + \sigma_{WI})$	$P(WI > \mu_{WI})$	$P(WI > \sigma_{WI})$	L (km)	Fst	Gr	CN	NDVI
1	58	0.99	0.4	0.05	0.46	0.08	0.63	0.55	27.87	0	0	86.32	0.29
2	56	0.98	0.34	0.06	0.49	0.09	0.64	0.54	29.84	4.24	0	85.36	0.3
3	455	0.96	0.51	0.05	0.46	0.11	0.62	0.51	198.62	5.8	0.37	82.98	0.31
4	57	0.92	0.21	0.09	0.51	0.08	0.63	0.55	32.39	10.69	0	81.53	0.3
5	233	0.98	0.46	0.03	0.48	0.1	0.62	0.52	106.56	6.6	4.05	83.86	0.3
6	92	0.96	0.21	0.1	0.5	0.07	0.61	0.55	59.94	14.88	10.1	83.35	0.33
7	539	0.55	0.25	0.09	0.63	0.07	0.62	0.55	270.33	12.06	7.81	75.08	0.34
8	715	0.59	0.33	0.07	0.58	0.08	0.61	0.53	389.74	3.79	1.94	80.63	0.33
9	263	0.85	0.18	0.06	0.52	0.05	0.65	0.6	188.88	8.33	0	78.37	0.34
10	124	0.83	0.03	0.12	0.49	0.03	0.67	0.64	72.41	20.98	0	52.51	0.39
11	1111	0.25	0.05	0.12	0.52	0.04	0.65	0.61	651.13	23.75	7.34	71.89	0.4
12	176	0.45	0.06	0.11	0.52	0.05	0.65	0.59	128.57	23.37	9.95	71.92	0.42
13	462	0.46	0.11	0.1	0.55	0.04	0.64	0.6	255.68	20.93	1.88	75.74	0.36
14	1657	0.48	0.14	0.09	0.56	0.04	0.65	0.61	785.75	15.21	2.3	79.03	0.32
15	42	0.63	0.09	0.15	0.4	0.05	0.64	0.59	32.62	72.09	0.44	74.95	0.44
16	853	0.4	0.03	0.16	0.52	0.04	0.64	0.6	439.82	41.71	13.93	70.52	0.4
17	1382	0.41	0.07	0.12	0.51	0.03	0.65	0.62	771.28	40.02	8.74	72.91	0.35
18	511	0.63	0.13	0.15	0.46	0.05	0.63	0.59	303.61	48.29	15.99	74.66	0.43
19	1140	0.41	0.05	0.16	0.4	0.03	0.65	0.62	664.67	57.79	10.21	73.01	0.39
20	232	0.3	0.02	0.2	0.44	0.04	0.65	0.61	112.93	64.52	13.9	76.96	0.43
21	1339	0.89	0.1	0.2	0.44	0.03	0.66	0.63	956.38	37.74	1.6	76.96	0.43
22	520	0.88	0.11	0.16	0.44	0.05	0.64	0.59	382.13	41.53	4.24	64.3	0.42
23	285	0.97	0.07	0.2	0.4	0.09	0.61	0.53	160.52	46.93	16.6	68.14	0.39
24	32	0.9	0.07	0.06	0.34	0.07	0.64	0.57	19.44	40.04	1.99	54.83	0.43
25	79	0.91	0.02	0.17	0.31	0.03	0.67	0.64	70.04	86.16	10.13	36.93	0.47
26	46	0.9	0.04	0.15	0.32	0.03	0.69	0.66	37.47	65.42	16.74	35.78	0.47
27	41	0.91	0.02	0.16	0.38	0.02	0.67	0.64	32.46	67.5	5.75	51.22	0.47
28	182	0.9	0.03	0.16	0.4	0.01	0.68	0.67	119.08	64.89	6.58	51.22	0.46
29	116	0.9	0.1	0.16	0.3	0.04	0.66	0.61	63.62	63.44	11.22	51.22	0.48
30	15	0.9	0.17	0.09	0.26	0.08	0.63	0.55	4.69	73.03	13.48	36.39	0.5
31	113	0.9	0.07	0.15	0.41	0.02	0.67	0.65	62.17	59.78	1.72	55.74	0.46
32	280	0.76	0.04	0.19	0.52	0.04	0.66	0.62	170.32	53.83	11.85	71	0.41
33	42.53	0.79	0.01	0.2	0.47	0.05	0.65	0.6	20.82	84.63	3.95	63.59	0.42

assumption that runoff generation may be due to two different generation mechanisms: one L-type (frequent) response occurring when a lower threshold is exceeded, and the second named H-type (rare) response, occurring when a higher threshold is exceeded providing extraordinary floods. More details on the theoretical background of the two distributions are reported in Appendix A.

Fiorentino et al. (2011) performed the application of the two distributions to the same set of river basins in Southern Italy and highlighted the influence of the soil and climate physical characteristics on the probability distribution of floods, the skewness coefficients and the mean annual number of ordinary and extraordinary independent events. The two probability distributions IF and TCIF are characterized respectively by twelve and fifteen parameters (see Gioia et al., 2008); most of these are evaluated by a priori knowledge based on data different from the AMFS, mainly rainfall records and other geomorphoclimatic information (see Appendix B and Table 3). Nine of these parameters are common to both distributions while the values of the loss threshold scaling factors  $\varepsilon'$ ,  $\varepsilon_L$ ,  $\varepsilon_H$ , were assigned according to basin climate.

The parameters that need calibration by means of AMFS are two ( $\Lambda_q, r$ ) for the IF model ( $\Lambda_q$  is the mean annual number of independent flood events and  $r$  is the ratio between the average contributing area  $E[a]$  and the total basin area  $A$ ) and four ( $\Lambda_L, \Lambda_H, r_L, r_H$ ) for the TCIF model ( $\Lambda_L, \Lambda_H$  are the mean annual numbers of independent flood events of the ordinary and extraordinary components, respectively, and  $r_L, r_H$  are the ratios between the average contributing area of the ordinary and extraordinary components, respectively, and the total basin area  $A$ ). All these parameters are strictly related to runoff generation mechanisms and were estimated by Fiorentino et al. (2011) by means of a numerical, at-site evaluation procedure based on maximizing a likelihood function evaluated on AMFS.

Using such estimates of  $\Lambda_q$  and the a priori estimates of  $k$ ,  $\Lambda_p, E[i_{A,\tau}]$  the relationship

$$f_A = \left[ E[i_{A,\tau}^k] \log \left( \frac{\Lambda_p}{\Lambda_q} \right) \right]^{1/k} \quad (1)$$

provides the space-time average hydrologic loss per unit of basin area  $f_A$  of the IF distribution.

**Table 3.** Parameters of the IF and TCIF distributions based on information other than AMFS.

No.	$A$ (km <sup>2</sup> )	$q_0$ (m <sup>3</sup> s <sup>-1</sup> )	$\tau_A$ (h)	$p_1$ (mm h <sup>-1</sup> )	$n$	$E[i_{A,\tau}]$ (mm h <sup>-1</sup> )	$\varepsilon$	$\Lambda_p$	$k$	$\varepsilon'$	$\varepsilon_L$	$\varepsilon_H$
1	58	0.2	2.6	24.27	0.29	2.01	0.39	44.6	0.80	0.5	0.5	0.50
2	56	0.3	2.6	34.50	0.30	2.90	0.39	44.6	0.80	0.5	0.5	0.50
3	455	2.8	7.3	23.29	0.27	0.81	0.39	44.6	0.80	0.5	0.5	0.50
4	57	0.5	2.6	24.57	0.30	2.06	0.39	44.6	0.80	0.5	0.5	0.50
5	233	2.2	5.2	23.33	0.27	1.07	0.39	44.6	0.80	0.5	0.5	0.50
6	92	1.2	3.3	24.08	0.29	1.65	0.39	44.6	0.80	0.5	0.5	0.50
7	539	8.0	8.0	23.86	0.28	0.79	0.39	44.6	0.80	0.5	0.5	0.50
8	715	7.0	9.2	24.13	0.28	0.72	0.39	44.6	0.80	0.5	0.5	0.50
9	263	1.4	5.6	24.13	0.26	1.02	0.39	44.6	0.80	0.5	0.5	0.50
10	124	1.2	3.8	24.13	0.26	1.42	0.39	44.6	0.80	0.5	0.5	0.50
11	1111	26.0	11.5	24.13	0.26	0.78	0.39	21.0	0.80	0.5	0.5	0.50
12	176	3.0	4.6	24.13	0.26	1.59	0.39	21.0	0.80	0.5	0.5	0.50
13	462	5.0	4.3	22.20	0.28	1.46	0.33	21.0	0.80	0.5	0.5	0.50
14	1657	10.0	7.1	23.52	0.28	1.08	0.33	21.0	0.80	0.5	0.5	0.50
15	42	1.5	2.9	21	0.31	2.20	0.39	21.0	0.80	0.0	0	0.50
16	853	25.0	4.8	20.41	0.31	1.27	0.33	21.0	0.80	0.0	0.0	0.50
17	1382	25.0	6.0	21.48	0.31	1.16	0.33	21.0	0.80	0.5	0.5	0.50
18	511	10.0	8.9	21.56	0.36	1.05	0.39	21.0	0.80	0.0	0	0.50
19	1140	45.0	5.6	23.13	0.4	1.53	0.33	21.0	0.80	0.0	0	0.50
20	232	15.0	2.4	21.56	0.36	2.03	0.33	32.0	0.80	0.0	0	0.50
21	1339	52.0	5.5	24.27	0.4	1.05	0.28	20.0	0.53	0.0	0	0.50
22	520	22.4	4.7	22.13	0.46	1.17	0.28	20.0	0.53	0.0	0	0.50
23	285	8.4	3.7	22.27	0.55	1.55	0.28	20.0	0.53	0.0	0	0.50
24	32	1.2	2.8	31.02	0.5	2.65	0.28	20.0	0.53	0.0	0	0.50
25	79	3.4	3	32.66	0.59	4.14	0.32	10.0	0.53	0.0	0	0.50
26	46	2.3	3.0	33.22	0.52	2.77	0.28	20.0	0.53	0.0	0	0.50
27	41	1.8	3.0	33.22	0.47	2.63	0.28	20.0	0.53	0.0	0	0.50
28	182	8.8	3.8	29.84	0.45	1.87	0.28	20.0	0.53	0.0	0	0.50
29	116	7.1	3.9	37.92	0.54	3.97	0.32	10.0	0.53	0.0	0	0.50
30	15	1.0	1.3	39.58	0.63	7.50	0.32	10.0	0.53	0.0	0	0.50
31	113	5.3	4.6	28.84	0.43	1.65	0.28	20.0	0.53	0.0	0	0.50
32	280	12.4	3.7	27.79	0.46	1.37	0.28	34.0	0.53	0.0	0	0.50
33	43	2.7	1.3	26.61	0.47	2.59	0.28	34.0	0.53	0.0	0	0.50

Analogously, for the TCIF distribution, the evaluation of  $f_{A,L}$  and  $f_{A,H}$ , using estimates of  $k$ ,  $\Lambda_p$ ,  $E[i_{A,\tau}]$ ,  $\Lambda_L$ ,  $\Lambda_H$  is provided by:

$$f_{A,L} = \left[ E[i_{A,\tau}^k] \log \left( \frac{\Lambda_p}{\Lambda_L + \Lambda_H} \right) \right]^{1/k} \text{ and}$$

$$f_{A,H} = \left[ E[i_{A,\tau}^k] \log \left( \frac{\Lambda_p}{\Lambda_H} \right) \right]^{1/k} \quad (2)$$

The at-site parameter estimates are reported, for all basins, in Table 4 with the best-fit distribution selected by means of a maximum likelihood criterion. Results show that the IF distribution was selected against TCIF for 19 out of 33 basins. From the analysis of data in Tables 3 and 4 it is possible to observe that the presence of two components is quite homogeneous throughout the study area and independent of climate, observed skewness of AMFS and geographical position of the basin. This consideration raises the important

problem of a priori identification of basin hydrological behaviour from observable physical features.

This issue is faced in Sect. 3 where the basins' behaviour is investigated on a regional scale providing the identification of regional relationships for parameters  $f_A$ ,  $f_{A,L}$ ,  $f_{A,H}$ ,  $r$ ,  $r_L$ ,  $r_H$ , and a criterion for establishing whether the distribution is of the IF or TCIF type. With this aim, we considered the regional variability of the quantity  $F_{A,H} = f_{A,H} \cdot \tau_A$  corresponding to the cumulated infiltration in the lag-time of the river basin. It represents a capacitive threshold of the extraordinary ( $H$ ) component and its evaluation was extended to the entire set of basins by at-site calibration of the TCIF model. It is worth noting that the parameter  $f_{A,H}$ , estimated for AMFS classified as IF distributed, is expected to reach high values with respect to the mean rainfall intensity. Such a condition produces as result that the frequency of the events belonging to the second component tends to zero. These

**Table 4.** At-site parameter estimates and selected probability distribution.

No.	$r$	$r_L$	$r_H$	$f_A$ (mm h <sup>-1</sup> )	$f_{A,L}$ (mm h <sup>-1</sup> )	$f_{A,H}$ (mm h <sup>-1</sup> )	$F_{A,H}$ (mm)	$\Lambda_q$	$\Lambda_L$	$\Lambda_H$	Best fit
1	–	0.03	0.59	–	6.04	10.37	26.95	–	2.38	0.74	TCIF
2	0.19	–	–	7.78	[7.76]	[27.08]	[70.41]	3.9	[3.86]	[0.06]	IF
3	0.12	–	–	1.69	[1.69]	[4.82]	[35.17]	6.1	[5.65]	[0.45]	IF
4	–	0.1	0.72	–	5.31	11.23	29.2	–	3.61	0.61	TCIF
5	–	0.09	0.37	–	2.29	7.96	41.38	–	5.64	0.18	TCIF
6	0.11	–	–	3.12	[3.12]	[15.47]	[51.03]	7.1	[7.04]	[0.06]	IF
7	0.62	–	–	1.76	[1.76]	[17.28]	[138.20]	5.5	[5.52]	[0.00]	IF
8	–	0.41	0.99	–	1.01	3.86	35.53	–	9.86	0.66	TCIF
9	–	0.06	0.99	–	2.93	5.93	33.23	–	2.93	0.49	TCIF
10	0.2	–	–	3.33	[3.33]	[11.55]	[43.88]	5	[4.88]	[0.12]	IF
11	0.9	–	–	1.12	[1.12]	[15.80]	[181.70]	4.8	[4.82]	[0.00]	IF
12	0.16	–	–	1.99	[1.98]	[32.15]	[147.88]	5.6	[5.62]	[0.00]	IF
13	–	0.15	0.99	–	2.0106	5.0826	21.85518	–	3.98	1.04	TCIF
14	0.38	–	–	2.03	[2.02]	[8.44]	[59.90]	3.4	[3.35]	[0.07]	IF
15	0.23	–	–	0	[0.00]	[44.50]	[129.06]	21	[21.00]	[0.00]	IF
16	–	0.25	0.86	–	0.99	8	38.38	–	8.35	0.17	TCIF
17	0.2	–	–	1.32	[1.32]	[9.04]	[54.26]	6.2	[6.13]	[0.07]	IF
18	0.19	–	–	0	[0.00]	[8.16]	[72061]	21	[20.93]	[0.07]	IF
19	–	0.16	0.85	–	0	8.75	49.02	–	20.58	0.24	TCIF
20	0.3	–	–	0	[0.00]	[4.04]	[9.70]	32	[27.29]	[4.71]	IF
21	0.12	–	–	0	[0.00]	[65.02]	[357.59]	15	[20.00]	[0.00]	IF
22	–	0.2	0.6	–	0.01	9.67	45.47	–	17.32	0.3	TCIF
23	0.08	–	–	1.18	[1.18]	[96.12]	[355.63]	6.1	[6.12]	[0.00]	IF
24	–	0.03	0.14	–	1.65	40.49	113.37	–	6.84	0.06	TCIF
25	–	0.08	0.43	–	1.41	21.78	65.34	–	4.25	0.37	TCIF
26	0.04	–	–	0.56	[0.56]	[171.99]	[515.97]	11.1	[11.10]	[0.00]	IF
27	0.05	–	–	0	[0.00]	[163.47]	[490.41]	20	[20.00]	[0.00]	IF
28	0.23	–	–	0.32	[0.32]	[116.14]	[441.33]	11.7	[11.70]	[0.00]	IF
29	0.09	–	–	0.19	[0.19]	[220.86]	[861.35]	7.6	[7.62]	[0.00]	IF
30	0.05	–	–	0	[0.00]	[417.53]	[542.79]	10	[10.00]	[0.00]	IF
31	–	0.12	0.79	–	1.55	11.67	53.7	–	4.9	0.42	TCIF
32	0.21	–	–	0.19	[0.19]	[92.58]	[342.53]	21	[21.02]	[0.00]	IF
33	0.08	–	–	0	[0.00]	[174.45]	[226.79]	34	[33.99]	[0.00]	IF

at-site estimates of  $F_{A,H}$  are reported in Table 4 for all 33 river basins, where all values referred to the H-component evaluated on basins showing only a single ordinary component (because their AMFS are IF distributed) are given in brackets.

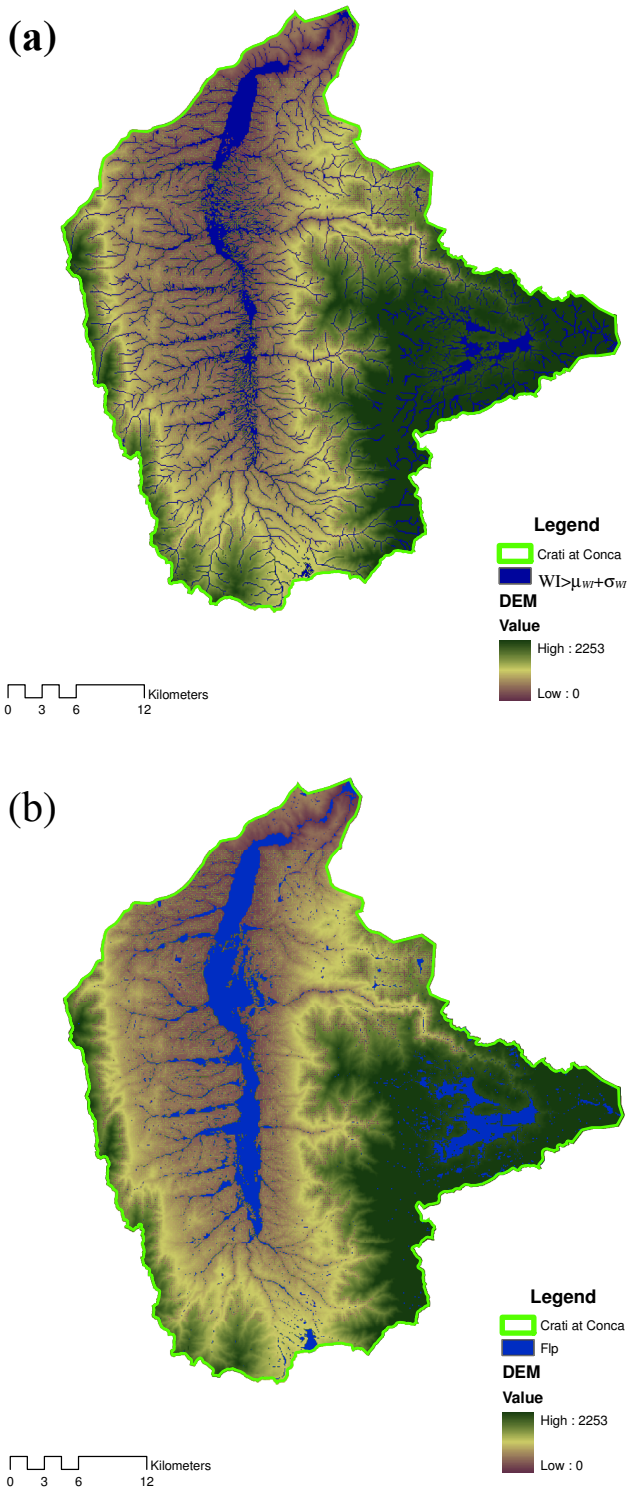
### 3 Regional relationships of model parameters

The spatial variability of parameters  $f_A$ ,  $f_{A,L}$ ,  $f_{A,H}$ ,  $F_{A,H}$ ,  $r$ ,  $r_L$ ,  $r_H$  was analyzed, on a regional scale, by exploring their statistical dependence on the geomorphological descriptors previously introduced. In order to provide reliable relationships, two different criteria were used to identify homogeneous groups of basins, based on climatic conditions and geographical position, for both the IF and TCIF model. In the first case, the basins are partitioned between dry ( $I < 0$ ) and humid ( $I > 0$ ), while in the second case basins are parti-

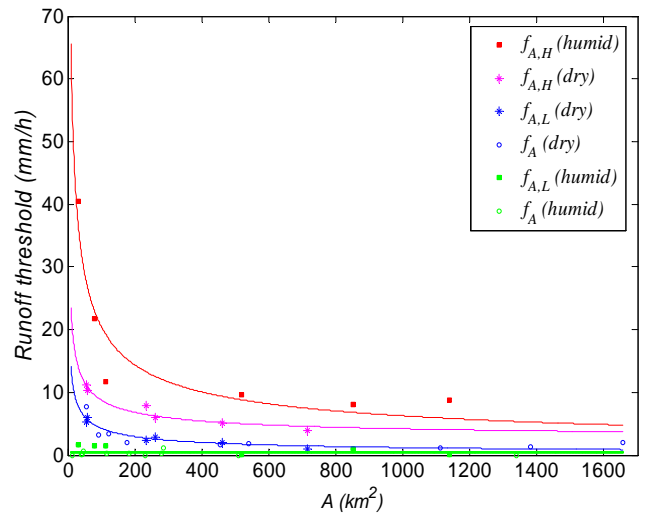
tioned according to the geographical region of origin that correspond to the Puglia (Northern sub-region), Basilicata (Central sub-region) and Calabria Region (Southern sub-region, see Fig. 1). The best-fit regressions were selected according to the adjusted coefficients of determination  $R^2$  while the regional predictive capacity was tested with a jack-knife procedure.

#### 3.1 Runoff thresholds

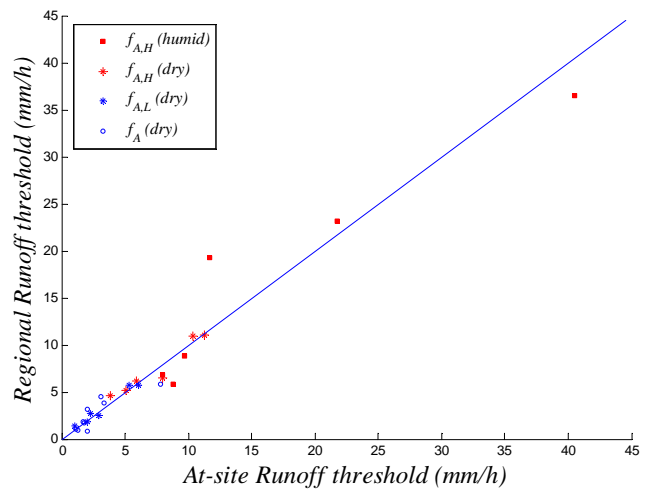
The main purpose of the regional analysis is to provide an estimate of flood peaks with a given return period in ungauged river sites. In this framework, regional relationships for the parameters  $f_{A,L}$ ,  $f_{A,H}$ ,  $r_L$ ,  $r_H$ , of the TCIF distribution and for the parameters  $f_A$  and  $r$  of the IF distribution are needed. Once the value of  $f_A$ ,  $f_{A,L}$ ,  $f_{A,H}$  is available, the corresponding values of  $\Lambda_q$ ,  $\Lambda_L$ ,  $\Lambda_H$  can be evaluated by means of Eqs. (A9) and (A16) reported in Appendix A.



**Fig. 2.** Comparison between (a) the portion of the basin with  $WI > \mu_{WI} + \sigma_{WI}$  and (b) the floodprone areas defined using the procedure proposed by Manfreda et al. (2011).



**Fig. 3.** Scaling relationship of the runoff thresholds for all river basins.



**Fig. 4.** Comparison between the at-site and regional estimates of different runoff thresholds for all river basins.

The regional behaviour of the  $f_A$ ,  $f_{A,L}$  and  $f_{A,H}$  parameters was already analyzed by Fiorentino et al. (2011) and is described in Fig. 3, in which four different scaling relationships can be recognized depending on climatic conditions. Starting from top to bottom of Fig. 3, one finds: (i)  $f_{A,H}$  in humid basins (red squares); (ii)  $f_{A,H}$  in dry basins (magenta stars); (iii)  $f_{A,L}$  (blue stars) and  $f_A$  (blue circles) in dry basins; (iv)  $f_{A,L}$  (green squares) and  $f_A$  values (green circles) in humid basins. The continuous lines represent power laws with exponent 0.5 (red, magenta, blue) and 0 (green). Results show that the runoff threshold of the low component ( $f_{A,L}$ ) of the TCIF distributions has the same behaviour observed for the single runoff threshold of the IF model both in dry and humid climates. This important observation suggests that the single threshold of the IF model is referred to the low component of runoff.

Thus, for the purpose of this paper, we have performed a best fit regression analysis on the three patterns above numbered as (i), (ii) and (iii), by assuming a power law with an exponent equal to 0.5, which provides:

for humid basins – H-type:

$$f_{A,H} = 208.52A^{-0.5} - 0.34; \quad R^2 = 0.89 \quad (3)$$

for dry basins – H-type:

$$f_{A,H} = 67.75A^{-0.5} + 2.04; \quad R^2 = 0.93 \quad (4)$$

for dry basins – L-type:

$$f_A = f_{A,L} = 45.32A^{-0.5} - 0.25; \quad R^2 = 0.82 \quad (5)$$

In Fig. 4, for these three cases, we report the comparison between the “at-site” parameter values and the corresponding regression estimates which we conventionally call “regional” values.

Regarding the fourth pattern in Fig. 3, the regional relationship is:

for humid basins – L- type:

$$f_A = f_{A,L} = 0.45 [\text{mmh}^{-1}]. \quad (6)$$

Thus, for determining the lower threshold,  $f_{A,L}$ , it is possible to adopt, as regional relationship valid for humid river basins, the constant value of infiltration rate reported in Eq. (6), and for dry river basins the value of infiltration rate derived from Eq. (5). For the higher threshold,  $f_{A,H}$ , it is possible to exploit respectively in humid and dry river basins the Eqs. (3) and (4).

These results could be quite satisfactory for a regional application. Nevertheless, we also considered that the basin classification, based on the climatic index  $I$ , may produce uncertainty when  $I$  is close to zero. In this case, a small bias in the evaluation of  $I$  could lead to a large difference in the regional prediction of the runoff thresholds. In order to avoid this discontinuity for small  $I$ , Eqs. (3) and (4) (as well as Eqs. 5 and 6) can be replaced by a single equation taking the variable  $I$  into account to be used in a small range of  $I$  beneath 0. The results of this analysis comprise the following regional relationships:

$$f_{A,H} = 71.60A^{-0.5} + 16.81IA^{-0.5} + 2 \\ R^2 = 0.93, \text{ in the range } -0.4 \leq I \leq 0.2, \quad (7)$$

$$f_A = f_{A,L} = 24.49A^{-0.5} - 56.24IA^{-0.5} + 0.74 \\ R^2 = 0.93, \text{ in the range } -0.2 \leq I \leq 0.2. \quad (8)$$

### 3.2 Average contributing area ratio

In order to identify regional relationships for estimating  $r$ ,  $r_L$ ,  $r_H$ , and  $F_{A,H}$ , we used several available descriptors and performed all the linear multiple regressions obtainable, adopting not more than three descriptors. We also considered different grouping criteria; in the following we report only the regressions which showed the best coefficient of determination  $R^2$  for the described model parameter.

It is noteworthy that we obtained a good fit for the  $r_H$  values including all 19 river basins following a TCIF distribution independently from climate and geographical location. The regional relationship is

$$r_H = 2.51C_{\text{runoff}} - 7.56P(\text{WI} > \mu_{\text{WI}} + \sigma_{\text{WI}}) \\ - 2.22IP(\text{WI} > \mu_{\text{WI}} + \sigma_{\text{WI}}) \quad R^2 = 0.925, \quad (9)$$

where the descriptors selected are: the climatic index  $I$ , the runoff coefficient proposed by De Smedt et al. (2000),  $C_{\text{runoff}}$ , and  $P(\text{WI} > \mu_{\text{WI}} + \sigma_{\text{WI}})$  which is the probability that the wetness index WI (Beven and Kirkby, 1979) exceeds the sum of its mean  $\mu_{\text{WI}}$  and standard deviation  $\sigma_{\text{WI}}$ .

With regard to the expected values  $r$  and  $r_L$  of the partial contributing area, in analogy with results obtained for  $f_A$  and  $f_{A,L}$ , we grouped the  $r$  values associated to the single component of the IF model with the  $r_L$  values of the ordinary component of the TCIF model. In this case, we obtained a better result grouping basins by geographical regions. As a consequence, we obtained the following relationships,

in Puglia:

$$\log r = \log r_L = 31.46\text{Sl} - 86.35IP(\text{WI} > \mu_{\text{WI}} + \sigma_{\text{WI}}) \\ - 3.50\psi - 2.6515, \quad R^2 = 0.86, \quad (10)$$

where Sl is the mean slope of the river basin,  $\psi$  is a permeability index defined as  $\psi = 1 - (\psi_{v,\text{low}} + 0.1\psi_{\text{low}})$  where  $\psi_{v,\text{low}}$  and  $\psi_{\text{low}}$  are the fractions of the total basin area with bedrock characterized respectively by very low ( $\psi_{v,\text{low}}$ ) and low permeability ( $\psi_{\text{low}}$ );

in Basilicata:

$$r = r_L = +0.00078A - 0.00147L + 0.24, \quad R^2 = 0.80 \quad (11)$$

where  $L$  is the total length of the stream network;

in Calabria:

$$\log r = \log r_L = -10.46A^{-0.5} + 0.02\text{Fst} \\ + 56.71\text{Flp}(1 - \psi) - 2.88, \quad R^2 = 0.75 \quad (12)$$

where Fst is the percentage of forest in basin area and Flp is the fraction of flood plain areas evaluated as in Manfreda et al. (2011).

It is important to point out that Eqs. (3), (4), (7) and (9) were obtained using data from basins that show evidence of



**Table 5.** Regional estimates of parameters  $r, r_L, r_H, f_a, f_{a,L}, f_{a,H}$ .

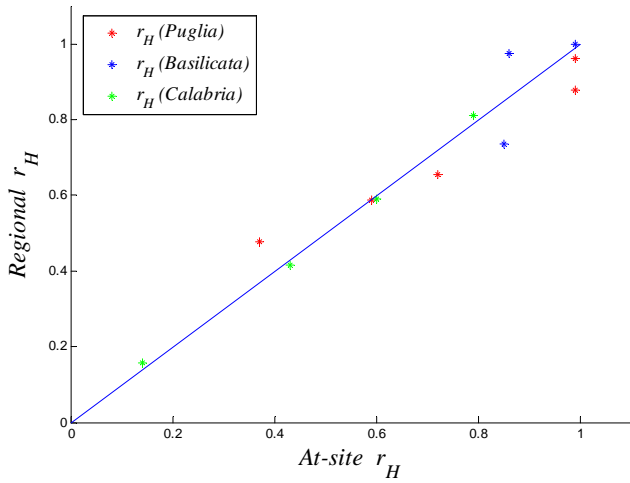
No.	Regional						Regional-Jack Knife					
	$r, r_L$	$r_H$	$f_a, f_{a,L}$ (mm h <sup>-1</sup> )	$F_{a,H}$ (mm)	$f_{a,H}$ (mm h <sup>-1</sup> )	Parent	$r, r_L$	$r_H$	$f_a, f_{a,L}$ (mm h <sup>-1</sup> )	$F_{a,H}$ (mm)	$f_{a,H}$ (mm h <sup>-1</sup> )	Parent
1	0.07	0.59	5.70	30.15	10.94	TCIF	0.09	0.59	5.95	31.20	11.46	TCIF
2	0.14	0.56	5.80	39.04	11.097	TCIF	0.12	0.56	5.80	34.95	11.10	TCIF
3	0.15	0.37	1.87	27.45	5.2198	TCIF	0.17	0.37	1.87	23.95	5.22	TCIF
4	0.13	0.65	5.75	46.08	11.017	TCIF	0.13	0.64	5.75	49.58	10.81	TCIF
5	0.05	0.48	2.71	36.73	6.4821	TCIF	0.03	0.56	2.71	36.02	6.14	TCIF
6	0.09	0.75	4.47	49.85	9.107	TCIF	0.08	0.75	4.47	49.57	9.11	TCIF
7	0.54	1.00	1.70	104.54	13.068	IF	0.51	1.00	1.70	55.84	4.96	TCIF
8	0.41	0.88	1.44	42.00	4.5773	TCIF	0.40	0.85	1.46	83.71	9.10	IF
9	0.06	0.96	2.54	52.45	6.2212	TCIF	0.05	0.96	2.62	57.09	6.29	TCIF
10	0.19	1.00	3.82	51.03	8.1277	TCIF	0.18	1.00	3.82	52.23	8.13	TCIF
11	0.84	1.00	1.10	200.64	17.447	IF	0.79	1.00	1.10	245.49	21.35	IF
12	0.20	0.87	3.16	123.17	26.775	IF	0.25	0.87	3.16	112.87	24.54	IF
13	0.22	1.00	1.85	31.61	5.1956	TCIF	0.24	1.00	0.00	36.42	5.24	TCIF
14	0.38	1.00	0.86	60.67	3.708	TCIF	0.38	1.00	0.00	60.80	3.71	TCIF
15	0.22	0.51	0.00	123.04	42.427	IF	0.21	0.51	0.47	109.74	37.84	IF
16	0.26	0.97	0.00	21.47	6.8007	TCIF	0.26	1.00	0.42	14.09	6.28	TCIF
17	0.19	1.00	0.96	40.16	3.8661	TCIF	0.17	1.00	0.00	34.62	3.87	TCIF
18	0.19	0.73	0.00	78.17	8.7827	IF	0.19	0.73	0.47	82.58	9.28	IF
19	0.15	0.74	0.00	60.60	5.8369	TCIF	0.15	0.72	0.47	73.65	4.39	TCIF
20	0.25	0.68	0.00	19.08	13.351	TCIF	0.23	0.68	0.47	22.61	13.35	TCIF
21	0.17	0.86	0.45	310.63	56.478	IF	0.19	0.86	0.47	274.99	50.00	IF
22	0.18	0.59	0.45	132.48	8.8054	TCIF	0.17	0.59	0.47	162.59	8.51	TCIF
23	0.09	0.23	0.45	386.78	104.54	IF	0.09	0.23	0.40	434.96	117.56	IF
24	0.03	0.16	0.45	63.00	36.523	TCIF	0.03	0.19	0.38	45.56	26.74	TCIF
25	0.11	0.42	0.45	139.15	23.122	TCIF	0.13	0.41	0.39	201.23	67.08	IF
26	0.06	0.54	0.45	351.67	117.22	IF	0.06	0.54	0.44	280.21	93.41	IF
27	0.05	0.73	0.45	233.37	77.789	IF	0.05	0.73	0.47	208.75	69.58	IF
28	0.11	0.89	0.45	243.94	64.195	IF	0.10	0.89	0.46	226.39	59.58	IF
29	0.14	0.31	0.45	823.55	211.17	IF	0.15	0.31	0.46	802.21	205.70	IF
30	0.04	0.04	0.45	650.70	500.54	IF	0.02	0.02	0.47	726.36	558.74	IF
31	0.10	0.81	0.45	151.94	19.277	TCIF	0.10	0.82	0.38	202.09	43.93	IF
32	0.15	0.86	0.45	398.07	107.59	IF	0.14	0.86	0.46	427.23	115.47	IF
33	0.07	0.61	0.45	103.02	31.64	TCIF	0.07	0.61	0.47	83.25	31.64	TCIF

the extraordinary component (because their AMFS are TCIF distributed), while Eqs. (5), (6), (8), (10), (11) and (12) exploit data from all the river basins considered, by merging the single component of the IF distribution with the ordinary component of the TCIF distribution.

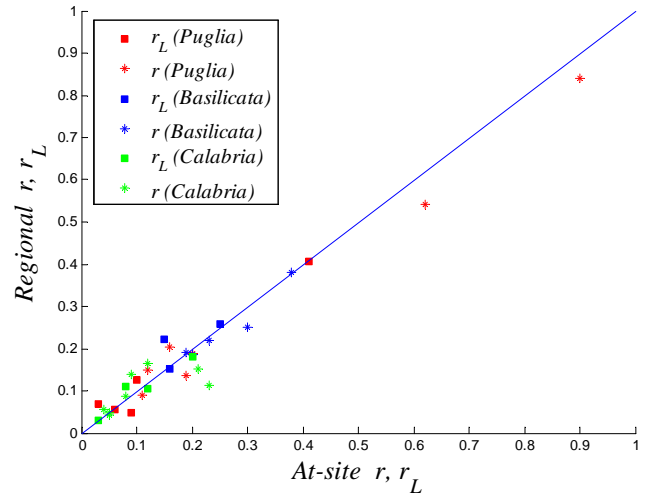
The regressions obtained show that physical parameters controlling the mean extent of the contributing area are mainly related to geomorphological characteristics of the river basin. This confirms that the contributing area (described by the parameter  $r$ ) is strongly controlled by the topographic characteristics of a basin that are described here by the indices derived from the maps of wetness index.

### 3.3 Cumulated infiltration and a priori identification of basin behaviour

Finally, we wish to report the results of the analysis performed on the parameter  $F_{A,H}$ . In this case the aim is to recognize the basin behaviour as characterized by only one or two components. Thus, it comprehends the at-site values estimated in all 33 river basins. Considering that  $f_{A,H}$  and  $\tau_A$  are respectively inversely and directly proportional to the root square of the basin area, their product  $F_{A,H}$ , should be independent from the basin area. On a regional scale, different values may be expected according to local values of  $f_1$  and  $\tau_1$  (which are space-time average infiltration and lag-time, respectively, per unit basin area). Nevertheless a certain regional homogeneity could be expected in areas



**Fig. 5.** Comparison between the at-site and regional estimates of  $r_H$  in river basins showing the extraordinary component.



**Fig. 6.** Comparison between the at-site and regional estimates of  $r$  and  $r_L$  for all river basins.

where the scaling law of lag-time has a constant coefficient  $\tau_1$ . Even in this case the regional relationships were identified according to the geographical criterion,

in Puglia:

$$\log F_{A,H} = 9.59 C_{\text{runoff}} - 3.87 P(\text{WI} > \mu_{\text{WI}}) \psi - 14.30 \text{Flp}(1 - \psi) + 1.49 \quad R^2 = 0.80 \quad (13)$$

in Basilicata:

$$F_{A,H} = 3640.5 P(\text{WI} > \sigma_{\text{WI}}) + 8425 \psi P(\text{WI} > \mu_{\text{WI}} + \sigma_{\text{WI}}) - 2313.3 \quad R^2 = 0.91 \quad (14)$$

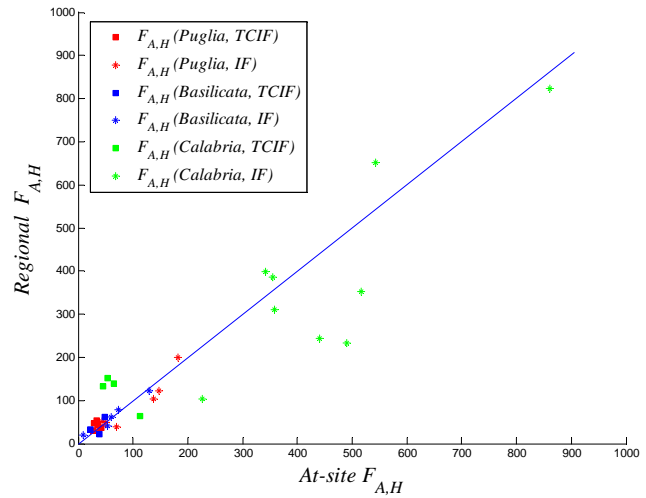
in Calabria:

$$\log F_{A,H} = 0.076 CN + 30.26 \text{NDVI} + 0.14 \text{Gr} - 13.38 \quad R^2 = 0.57 \quad (15)$$

in which the followings descriptors are used:  $CN$  SCS-Curve Number,  $\text{NDVI}$ , annual Normalized Difference Vegetation Index,  $\text{Gr}$ , percentage of grass cover in basin area.

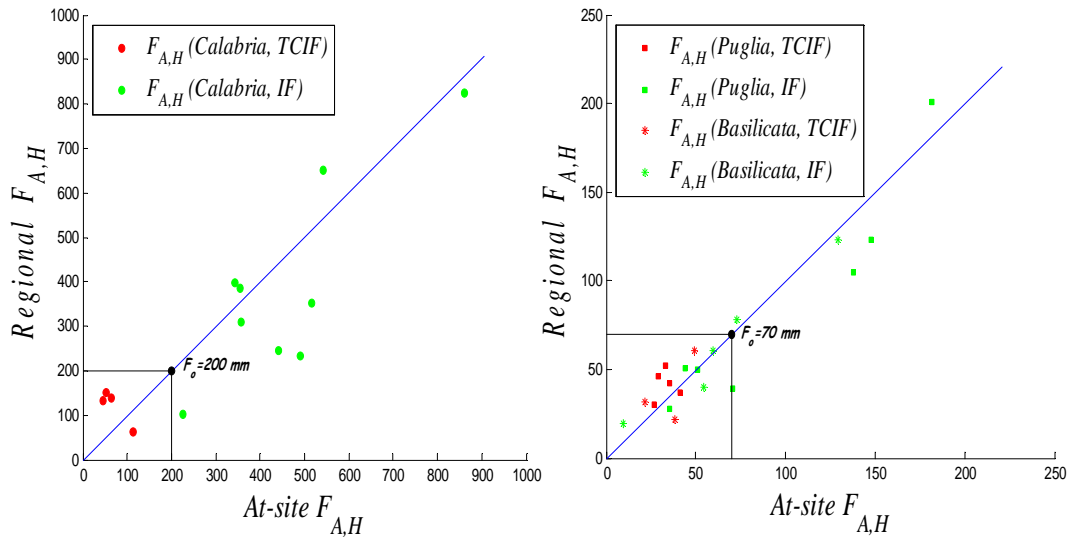
In Table 5 the regional estimates of  $r_L$ ,  $r_H$ ,  $f_{A,L}$ ,  $f_{A,H}$ , and  $F_{A,H}$  for 33 river basins by means of Eqs. (3)–(15), are reported. In Figs. 5, 6, 7 and 8, the comparison between the at-site and regional estimates of  $r_H$ ,  $r$ ,  $r_L$ ,  $F_{A,H}$  is displayed.

In particular, Fig. 8 reveals that river basins with AMFS following the TCIF model in Puglia (no. 1, 4, 5, 8, 9) and Basilicata (no. 13, 16, 19) are characterized by values of  $F_{A,H}$  always less than 70 mm, while  $F_{A,H}$  is always less than 200 mm for TCIF basins of Calabria (no. 22, 24, 25, 31). On the other hand, most of the river basins that follow the IF model are characterized by values of  $F_{A,H}$  greater than 70 mm in Puglia and Basilicata and 200 mm in Calabria. Thus, we adopted the value of  $F_{A,H}$ , to be estimated in ungauged basins by means of Eqs. (13), (14) and (15), as an



**Fig. 7.** Comparison between the at-site and regional estimates of  $F_{A,H}$  for all river basins.

index for the a priori classification of the basin behaviour. The limits just introduced can be used to identify the parent distribution, which will be considered TCIF if  $F_{A,H} < F_0$ . ( $F_0 = 70$  in Puglia and Basilicata and  $F_0 = 200$  mm in Calabria). In this case, we suggest using the regional Eqs. (3), (4) or (7) (depending on the climatic index  $I$  value) for the estimation of threshold  $f_{A,H}$ . In the second case, for basins with  $F_{A,H} > F_0$  we suggest using an IF distribution with regional estimates of  $f_A$  from Eqs. (5), (6) or (8). If  $F_{A,H} > F_0$  one may also use a TCIF distribution whereas the  $f_{A,H}$  value should be estimated from Eqs. (13), (14) or (15) as  $f_{A,H} = F_{A,H}/\tau_A$ . In fact, the IF and TCIF distributions provide indistinguishable results for high values of  $f_{A,H}$  because the TCIF distribution tends towards the IF distribution.



**Fig. 8.** Comparison between the at-site and regional estimates of  $F_{A,H}$  for all river basins.

In Table 5 we report the indication of the parent distribution following the above described criterion. As one can see, the identification is consistent with the at-site estimate in 25 out of 33 cases. Nevertheless it can be seen that the identification of a TCIF model, which is observed in 8 basins, always produces a quantile overestimation with respect to the IF model (then, safe side errors). In Table 6 we report the flood-quantile estimates for return periods of technical interest of 30, 200 and 500 years obtained by the regional model.

#### 4 Model validation

In order to test the performances of the above described regional model, we performed a jack-knife procedure structured as follows:

1. consider a regression model that relates  $y$  and  $\mathbf{X}$  by a function  $\varphi()$  so that:

$$y = \varphi(\mathbf{X}, \boldsymbol{\beta}), \tag{16}$$

where  $y$  is a scalar quantity while  $\mathbf{X}$  is a vector of length equal to the number  $N_D$  of the considered descriptors. The vector  $\boldsymbol{\beta}$  consists of  $N_D + 1$  unknown coefficients which are recovered by regression analysis at step 2.

2. the regression analysis is performed by estimating the coefficients  $\boldsymbol{\beta}$  from the set of equations

$$y_i = \varphi(\mathbf{X}_i, \boldsymbol{\beta}) \tag{17}$$

with  $i$  varying from 1 to  $N_S$  number of considered gauged sites with  $N_S > N_D$ ;

3. one of these gauged sites, say station  $j$ , and its corresponding  $y$  and  $\mathbf{X}$  values are removed from the set of equations in Eq. (17);
4. the regression analysis is carried out using in Eq. (17) the data of the remaining  $N_S - 1$  gauged sites and finding a new vector of coefficients  $\boldsymbol{\beta}_j$ ;
5. a jack-knife estimate of  $y$  for site  $j$  is then retrieved replacing in Eq. (16) the regression coefficients identified at steps 3 and 4;
6. steps 3–5 are repeated  $N_S$  times, considering in turn one of the gauged sites.

On the right side of Table 5, we report the jack-knife estimates of parameters of the proposed regional model including those for the identification of the parent distribution. The results show that, in this case, we have 21 out of 33 basins with indication consistent with the at-site estimate. Only in three basins, no. 8, 25 and 31, a regional IF distribution is selected instead of a TCIF compared to the at-site best-fit selection.

In Table 6 we display the flood quantile estimates for return periods of technical interest of 30, 200 and 500 years performed with at-site, regional and regional jack-knife procedures.

In Fig. 9 we report for all 33 river basins a Gumbel probability plot representing:

- the Weibull Plotting Positions (blue circles) of the observed AMFS;
- the cdf of the best-fit (IF or TCIF) distribution (continuous blue line) with at-site calibration procedure (Table 5);

**Table 6.** Quantiles of the TCIF and IF models with return periods of 30, 200 and 500 years.

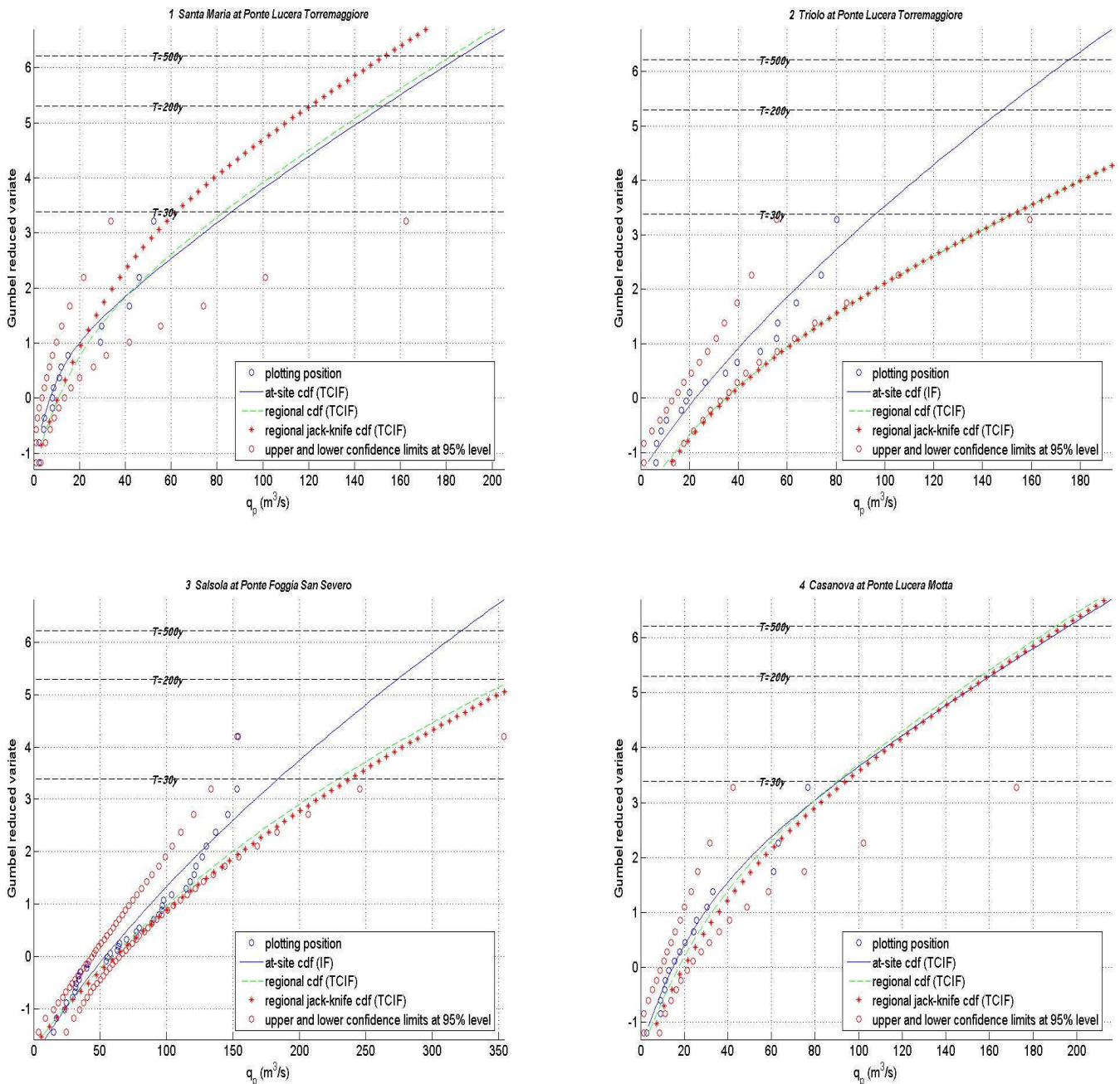
No.	At-Site $Q_T$ (m <sup>3</sup> /s)				Regional $Q_T$ (m <sup>3</sup> /s)				Regional jack-knife $Q_T$ (m <sup>3</sup> /s)			
	$T = 30$ years	$T = 200$ years	$T = 500$ years	Parent	$T = 30$ years	$T = 200$ years	$T = 500$ years	Parent	$T = 30$ years	$T = 200$ years	$T = 500$ years	Parent
1	86.483	152.34	186.62	TCIF	83.094	148.2	182.33	TCIF	62.316	120.38	153.33	TCIF
2	96.383	148.19	175.45	IF	152.86	243.46	290.88	TCIF	152.59	243.32	290.78	TCIF
3	183.76	274.31	322	IF	228.73	361.72	436.94	TCIF	237.85	372.59	447.39	TCIF
4	90.925	160.2	196.44	TCIF	89.997	155.53	190.6	TCIF	94.093	159.4	194.3	TCIF
5	118.6	197.6	245.93	TCIF	138.22	261.77	327.83	TCIF	164.47	299.97	370.92	TCIF
6	73.876	109.42	128.06	IF	114.76	205.14	252.44	TCIF	114.02	205.05	252.4	TCIF
7	534.33	790.17	921.27	IF	460.17	702.29	827.48	IF	467.18	722.64	854.79	TCIF
8	634.5	934	1090.4	TCIF	573.26	860.85	1011.8	TCIF	532.38	793.41	929.69	IF
9	212.14	384.85	472.84	TCIF	196.3	367.58	455.09	TCIF	192.19	363.15	450.55	TCIF
10	113.97	172.33	203.09	IF	151.1	253.75	309.34	TCIF	149.85	253.06	308.87	TCIF
11	1155.3	1703.1	1981.4	IF	1117.5	1658.9	1934.6	IF	1091	1625.9	1898.7	IF
12	154.65	233.37	274.9	IF	179.87	270.95	319.09	IF	208.2	311.92	366.4	IF
13	649.5	1070.4	1286.3	TCIF	652.6	1063.6	1277.3	TCIF	649.59	1057.1	1269.8	TCIF
14	1446.2	2268	2702.5	IF	1949.8	3029	3589.6	TCIF	1972.5	3046.8	3605.5	TCIF
15	74.901	108.21	125.67	IF	70.423	102.61	119.64	IF	69.181	100.85	117.61	IF
16	851.45	1344	1623.8	TCIF	962.45	1506.5	1811	TCIF	997.56	1565.5	1880.1	TCIF
17	970.7	1464.7	1727.1	IF	1610.6	2607.3	3121.8	TCIF	1604.7	2605.1	3120.4	TCIF
18	389.12	561.1	651.69	IF	361.26	530.39	620.33	IF	359.78	528.76	618.17	IF
19	1253.5	2079.4	2574.4	TCIF	1497.4	2481.2	3005.7	TCIF	1721.3	2741.3	3271.7	TCIF
20	475.38	675.35	780.63	IF	431.84	636.42	749.41	TCIF	410.14	610.64	723.28	TCIF
21	1226.3	2074.4	2559.6	IF	1425.9	2473.7	3087.4	IF	1582.2	2746.9	3438	IF
22	1032.1	1978.2	2559.3	TCIF	1015.3	1988	2577.1	TCIF	1014.2	1995.1	2586.5	TCIF
23	248.57	438.95	551.65	IF	283.38	491.94	615.89	IF	293.06	507.75	636	IF
24	30.162	73.097	105.47	TCIF	34.885	85.481	121.39	TCIF	49.421	116.44	157.5	TCIF
25	366.88	790.76	1047.8	TCIF	370.32	773.19	1021.9	TCIF	245.54	446.89	564.49	IF
26	47.325	81.267	101.49	IF	61.223	105.35	131.31	IF	63.969	110.08	136.66	IF
27	49.885	84.756	105.16	IF	46.247	79.63	99.419	IF	45.94	79.139	98.826	IF
28	448.83	775.11	967.61	IF	264.19	456.26	569.21	IF	237.32	409.24	511.72	IF
29	268.7	485.21	615	IF	352.3	640.1	808.9	IF	375.65	680.75	864.79	IF
30	45.209	81.034	102.46	IF	39.063	70.749	89.782	IF	27.034	48.799	61.864	IF
31	335.02	689.85	896.14	TCIF	224.4	551.03	754.86	TCIF	137.32	236.45	295.31	IF
32	565.62	940.68	1159.3	IF	435.88	726.66	898.6	IF	415.55	693.22	856.34	IF
33	84.076	138.34	169.78	IF	144.62	339.05	455.75	TCIF	144.05	338.97	455.72	TCIF

- the regional cdf (dashed green line) evaluated by means of the regional estimates;
- the regional jack-knife cdf (red stars) evaluated by means of the regional relationships reelaborated by means of the jack-knife procedure described above;
- the upper and lower confidence limits at 95% obtained by MonteCarlo analysis. For each river basin, the confidence bands are defined by extracting from the at-site best-fit (IF or TCIF) distribution, a number of samples equal to the size of the AMFS. Each sample is created generating one thousand random elements from the selected distribution. For each sample a plotting position corresponding to non-exceeding probability of 5% and 95% is reported as a confidence band.

It is important to observe that the regional-jack knife cdf provides satisfactory results in most of the cases and it is almost always within (or very close to) the confidence limits at 95% for return periods of technical interest. The only clear exception is provided by basin 28, Corace at Grascio, which is not affected by the choice of the parent distribution which is IF in all the estimation procedures (see Table 6). The quantiles underestimation is, in this case mainly due to the underestimation of the parameter  $r$  and hence in the evaluation of descriptors in Eq. (12).

## 5 Conclusions

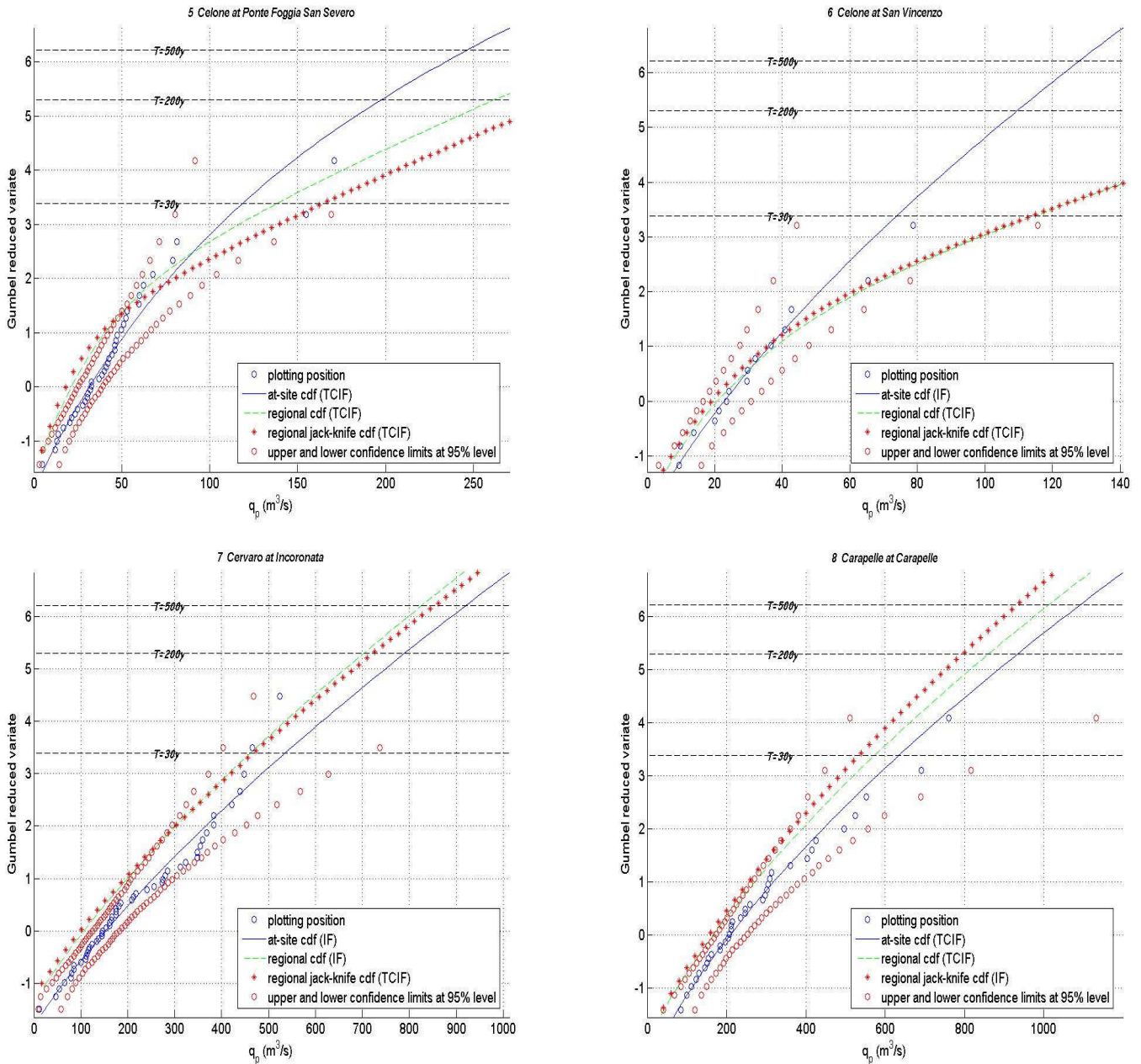
The estimation of flood quantiles in ungauged basins is one of the most challenging topics in hydrological applied research. In many areas of the world, the lack of reliable



**Fig. 9.** Weibull Plotting Positions (blue circles) of the observed AMFS, cdf of the best fit model (continuous blue line), regional cdf (dashed green line) and regional Jack.

discharge measurements and the unavailability of time series of sufficient length still compels the use of statistical analysis of rainfall coupled with deterministic rainfall-runoff models. In this context, the employment of a regional analysis framework performed by the investigation of the most appropriate relationships between model parameters and river basin descriptors related to climate, geomorphology, vegetation coverage, soil type and permeability, may help to solve the open problem of identifying hydrologic similarity and al-

lowing physical information in ungauged river basins to be exploited. We have shown that, within this context, the use of theoretically derived distribution may be beneficial to improve available regional methods. The main result that we have obtained is in the increased capability of investigating the spatial variability of parameters affecting the flood frequency distribution.



**Fig. 9.** Continued.

As a second important result, we mention the possibility to identify significant regional relationships in sub-regions characterized by geographical continuity. The positive side of this is that such subdivision is operatively more reliable than others referring to hydrological factors whose determination may be affected by the estimation procedure or by the choice of the indicator.

As a final remark we have to observe that we exploited a high number of geomorphological descriptors, and in some cases, with similar meaning. This is probably due to the fact that topographical information today has become signif-

icantly reliable, thanks to the availability of DEMs of high quality and high resolution. On the other hand, the proliferation of geomorphological descriptors is also a consequence of the lack of soil type characterization. The pedological datasets within the studied region are inhomogeneous and also with coarse resolution.

We believe that, in progress, the results of regional analysis may be further improved by means of detailed study of basins' hydrological behaviour, for example by allowing for a direct identification of events that arise as a consequence of different mechanisms of runoff generation. Important

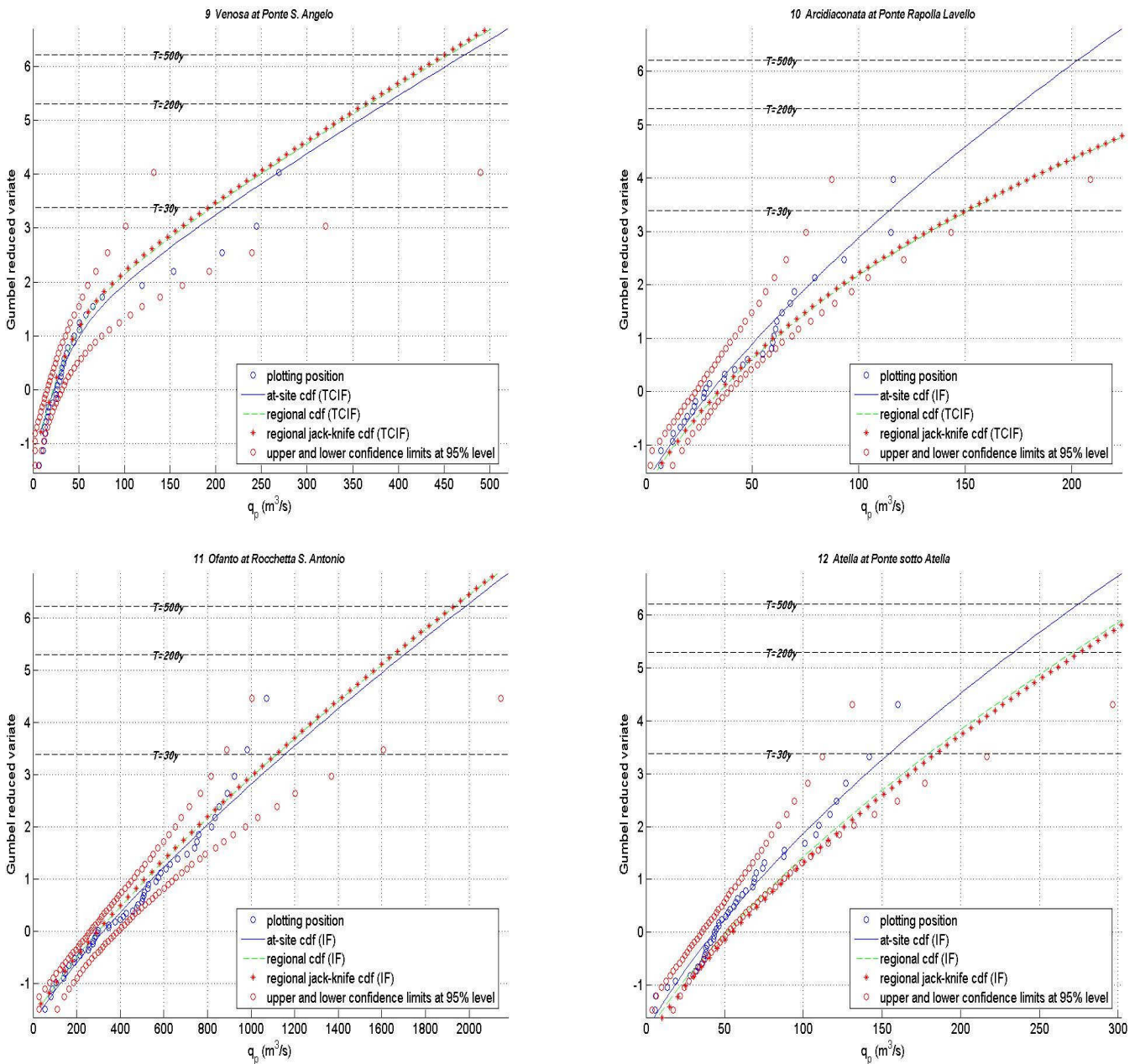


Fig. 9. Continued.

breakthroughs are expected in this direction from the study of soil moisture dynamics and from the observation of hydrological processes.

## Appendix A

### The IF and TCIF distributions

The IF model (Iacobellis and Fiorentino, 2000) assumes that

- the peak of direct streamflow  $Q$  is the product of two random variables strongly correlated, the source area

contributing to runoff peak  $a$  and the runoff peak per unit of contributing area,  $u_a$ .

- the runoff peak per unit area,  $u_a$ , is linearly dependent on the areal net rainfall intensity (which is assumed Weibull distributed with shape parameter  $k$ ) in a time interval equal to lag-time  $\tau_a$  (intended as the lag of direct runoff centroid to effective rainfall centroid) with a constant routing factor  $\xi$
- the lag-time  $\tau_a$  scales with  $a$  according to a power law with exponent 0.5.

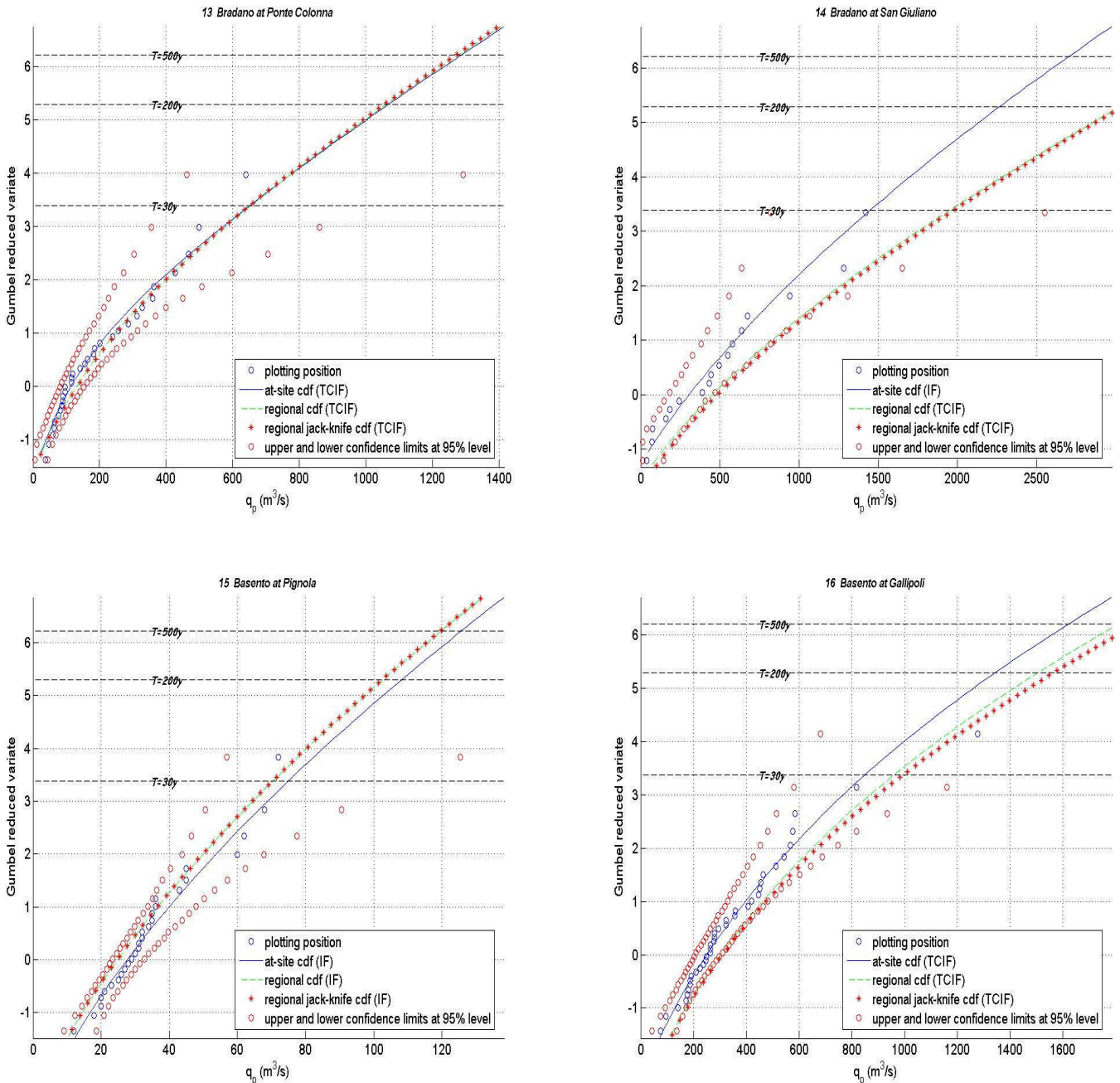


Fig. 9. Continued.

The non exceedance probability distribution of the peak flood  $G_Q(q)$ , is derived integrating (e.g. Eagleson, 1972) the joint probability (pdf) density function  $g(u,a)$  of the runoff peak per unit area,  $u_a$ , and the contributing area,  $a$ , over the appropriate domain where their product is minor than  $Q$ . The function  $g(u_a,a)$  is expressed as:

$$G_Q(q) = \int_0^A \int_0^{\frac{q}{a}} g(u|a) g(a) du da \quad (A1)$$

The distribution of the contributing areas  $g(a)$  is:

$$g(a) = \frac{1}{\alpha \Gamma(\beta)} \left(\frac{a}{\alpha}\right)^{\beta-1} \exp\left(-\frac{a}{\alpha}\right) + \delta(a-A) P_A \quad (A2)$$

in which  $\Gamma(\cdot)$  is the particular function Gamma,  $P_A = \text{prob}[a=A]$  and  $\delta$  is the Dirac function.



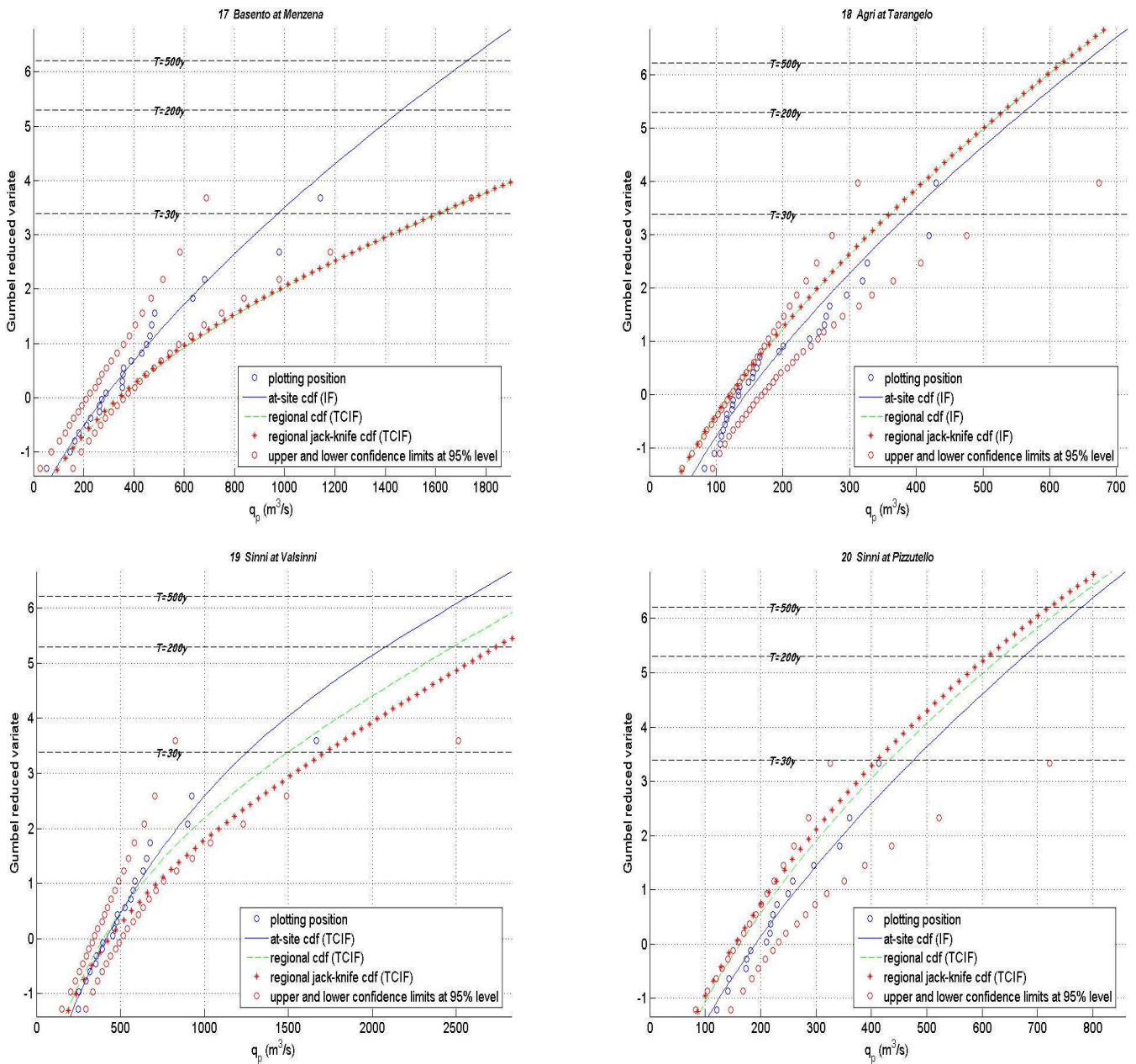


Fig. 9. Continued.

The coefficients  $\alpha$  and  $\beta$  represent the position and shape parameter of the Gamma Distribution, respectively. The position parameter  $\alpha = rA/\beta$  depends on the ratio,  $r$ , of the expected contributing area to the basin area:

$$r = \frac{E[a]}{A} \tag{A3}$$

The runoff peak per unit area,  $u_a$ , is related to the rainfall intensity  $i_{a,\tau}$  in the critical duration,  $\tau_a$ , of area  $a$ , minus the hydrological loss  $f_a$  depending on the same area  $a$ :

$$u_a = \xi(i_{a,\tau} - f_a) \tag{A4}$$

$\xi$  is a constant routing coefficient,  $i_{a,\tau}$  is the mean spatio-temporal rainfall intensity in the duration equal to  $\tau_a$  and in the area  $a$ , and  $f_a$  is the corresponding mean spatio-temporal hydrological loss. The distribution of  $u_a$  is obtained exploiting Eq. (A4) and assuming that the distribution of the  $i_{a,\tau}$ , is Weibull distributed with two parameters,  $\theta$  and  $k$ , in particular the first is a function of the mean of the  $i_{a,\tau}$ :

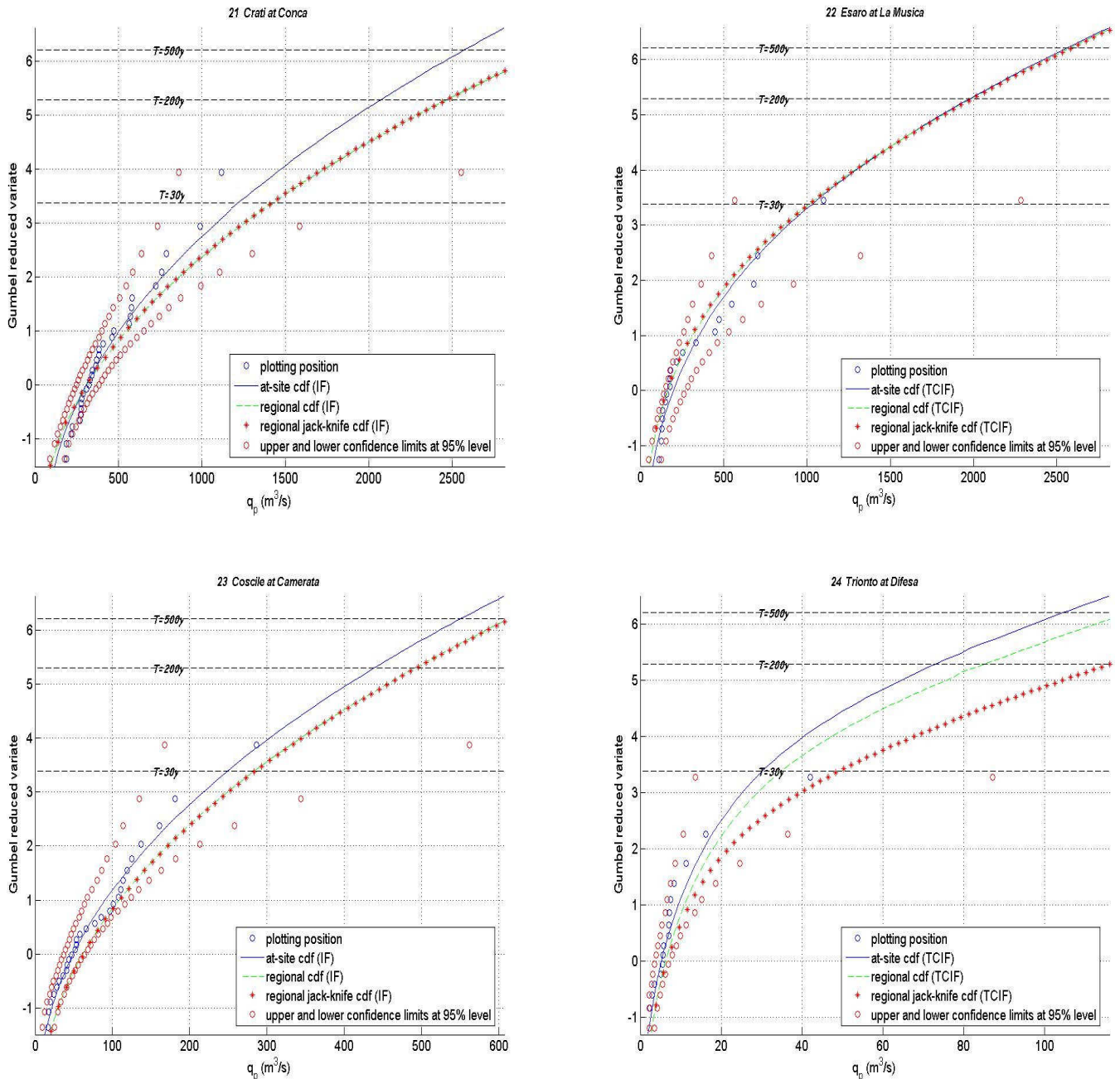


Fig. 9. Continued.

$$\theta_{a,\tau} = E \left[ i_{a,\tau}^k \right] = \left( \frac{E[i_{a,\tau}]}{\Gamma(1+1/k)} \right)^k \quad (A5)$$

in which  $\Gamma(1+1/k)$  is the Gamma function.

Besides the three main hypotheses mentioned before, the IF model assumes that both average rainfall intensity ( $E[i_{a,t}]$ ) and average hydrologic loss ( $f_a$ ) scale with contributing area according to the following power law relationships:

$$E[i_{a,\tau}] = E[i_{A,\tau}] (a/A)^{-\varepsilon} \quad f_a = f_A (a/A)^{-\varepsilon'} \quad (A6)$$

where  $E[i_{A,\tau}]$  and  $f_A$  are the average rainfall intensity and the average hydrologic loss respect to the entire basin area  $A$ ;  $\varepsilon'$  is a parameter representative of the space-time average hydrological losses.

Finally, under the hypothesis of Poissonian occurrence of independent annual maximum floods, the cumulative distribution function (CDF) of the annual maximum flood peak

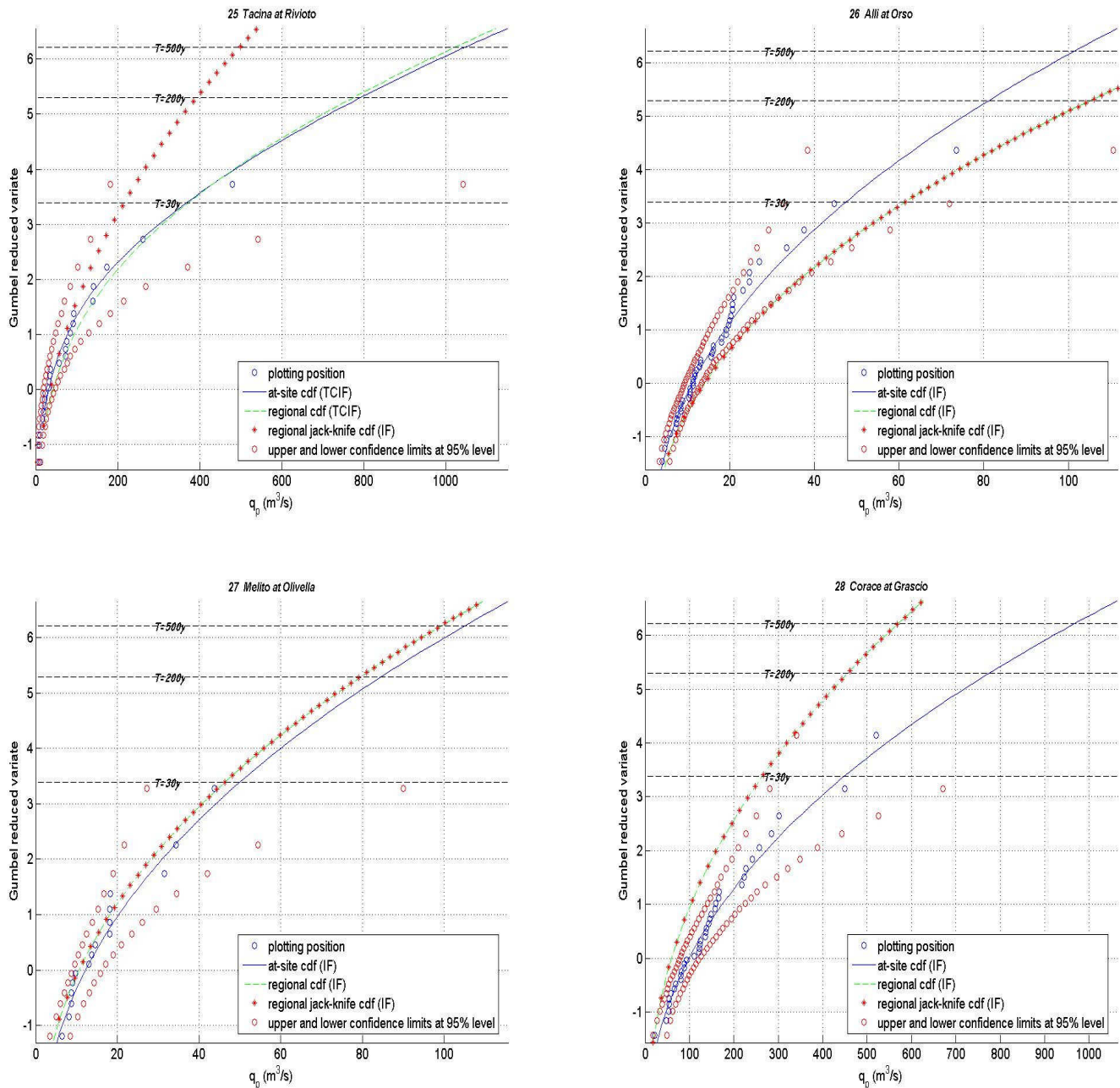


Fig. 9. Continued.

$q_p = Q + q_o$ , is derived, with  $q_o$  the base flow.

$$\text{CDF}_{Q_p}(q_p) = \exp \left\{ -\Lambda_q \left[ \int_0^A g(a) \exp \left( -\frac{((q_p - q_o)/(\xi a) + f_a)^k - (f_a)^k}{(E[i_{a,\tau}] / \Gamma(1 + 1/k))^k} \right) da \right] \right\} \tag{A7}$$

The probability density function (PDF) is

$$\text{PDF}_{Q_p}(q_p) = \text{CDF}_{Q_p}(q_p) \left[ \Lambda_q \left\{ \int_0^A g(a) \frac{k}{(\xi a) (E[i_{a,\tau}] / \Gamma(1 + 1/k))^k} \left( \frac{(q_p - q_o)}{(\xi a)} + f_a \right)^{k-1} \exp \left( -\frac{((q_p - q_o)/(\xi a) + f_a)^k - (f_a)^k}{(E[i_{a,\tau}] / \Gamma(1 + 1/k))^k} \right) da \right\} \right] \tag{A8}$$

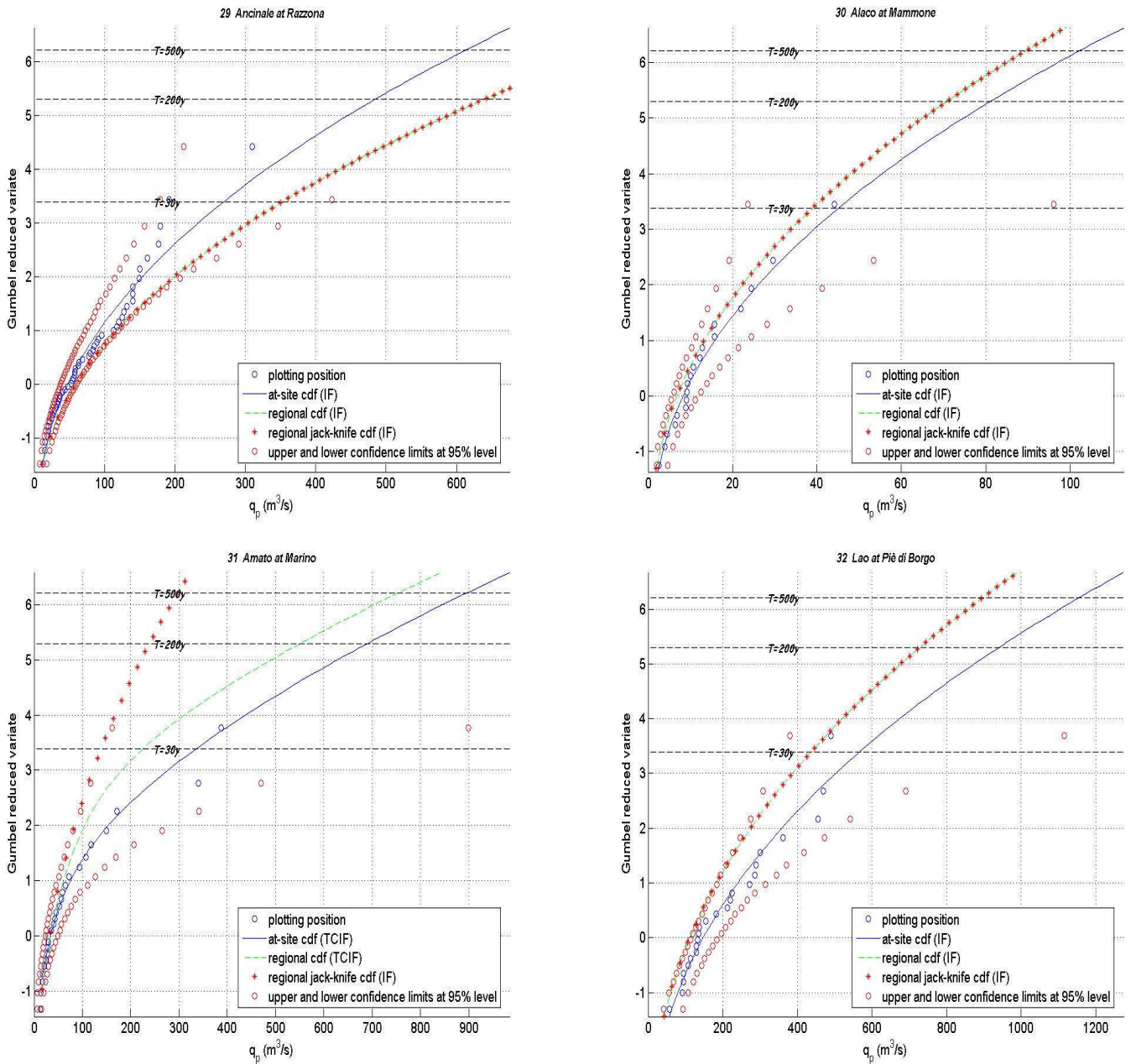


Fig. 9. Continued.

Based on these hypotheses, the following relationship between the mean annual number of rainfall ( $\Lambda_p$ ) and flood events ( $\Lambda_q$ ) holds:

$$\Lambda_q = \Lambda_p \exp\left(-f_a^k / E[i_{a,\tau}^k]\right) \quad (A9)$$

In which  $\Lambda_p$  and  $\Lambda_q$  are, respectively, the mean number of independent rainfall and flood events.

The two-component derived distribution called “Two Component IF” distribution (TCIF) was introduced by Gioia et al. (2008), who generalized the IF distribution considering

that two different response-mechanisms may arise, in both humid and dry basins:

- “L-type” (frequent) response, occurring when a lower threshold  $f_{a,L} = f_{A,L} (a_L/A)^{-\epsilon_L}$  is exceeded, and responsible for ordinary floods likely produced by a relatively small portion of the basin  $a_L$ , producing a peak runoff per unit area  $u_{a,L} = \xi (i_{a,\tau} - f_{a,L})$ ;
- “H-type” (rare) response, occurring when a higher threshold  $f_{a,H} = f_{A,H} (a_H/A)^{-\epsilon_H}$  is exceeded, and providing extraordinary floods mostly characterized by

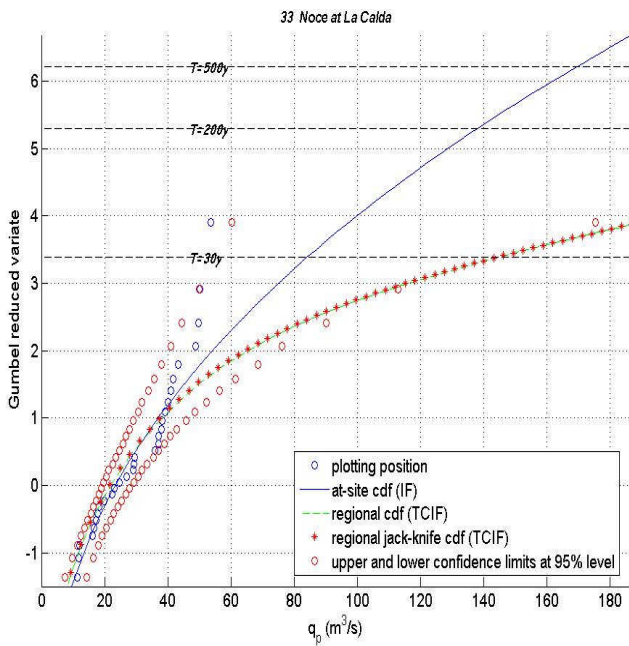


Fig. 9. Continued.

larger contributing areas  $a_H$ , producing a peak runoff per unit area  $u_{a,H} = \xi (i_{a,\tau} - f_{a,H})$ ;

where  $\xi$ , routing coefficient, is independent from the mechanism of runoff generation and two different runoff thresholds are introduced  $f_{a,L}$  and  $f_{a,H}$  (with  $f_{a,H} > f_{a,L}$ ) which scale with the contributing area following the power law relationships:

$$f_{a,L} = f_{A,L} (a_L/A)^{-\varepsilon_L} \tag{A10}$$

$$f_{a,H} = f_{A,H} (a_H/A)^{-\varepsilon_H} \tag{A11}$$

The flood-peak contributing areas  $a_L$  and  $a_H$  are distributed as in the IF model with expected contributing areas related to  $r_L$  and  $r_H$ :

$$r_L = E[a_L]/A \tag{A12}$$

$$r_H = E[a_H]/A \tag{A13}$$

Assuming that L-type and H-type events are independent and that both rates of occurrence are Poisson distributed, the overall process of exceedances is also a Poisson process and the cdf of the annual maximum floods is

$$CDF_{Q_p}(q_p) = \exp \{ -\Lambda_L [G'_{Q,L}(q_p)] - \Lambda_H [G'_{Q,H}(q_p)] \}, \tag{A14}$$

where  $G_{q,L}$  and  $G_{q,H}$  are the peak flow distributions of the direct streamflow  $Q$  corresponding respectively to L-type events and H-type events and are described as:

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

$$G'_{Q,H}(q) = \int_0^A \int_{\frac{q}{a}}^{\infty} g(u|a_H) g(a_H) du da_H, \tag{A15}$$

$\Lambda_L$  and  $\Lambda_H$  are respectively the mean annual number of independent flood events for L-type and for H-type processes and are related to the runoff thresholds by means of the following relationships:

$$\Lambda_q = \Lambda_L + \Lambda_H = \Lambda_p \exp \left( -\frac{f_{A,L}^k}{E[i_{A,\tau}^k]} \right) \text{ and}$$

$$\Lambda_H = \Lambda_p \exp \left( -\frac{f_{A,H}^k}{E[i_{A,\tau}^k]} \right). \tag{A16}$$

The TCIF cumulative distribution function and its probability density function are:

$$CDF_{Q_p}(q_p) = \exp \left\{ -\Lambda_L \left[ \int_0^A g(a_L) \exp \left( -\frac{((q_p - q_0)/(\xi a_L) + f_{a,L})^k - (f_{a,L})^k}{(E[i_{a_L,\tau}]/\Gamma(1+1/k))^k} \right) da_L \right] \right. \\ \left. + \exp \left\{ -\Lambda_H \left[ \int_0^A g(a_H) \exp \left( -\frac{((q_p - q_0)/(\xi a_H) + f_{a,H})^k - (f_{a,H})^k}{(E[i_{a_H,\tau}]/\Gamma(1+1/k))^k} \right) da_H \right] \right\} \right\} \tag{A17}$$

$$PDF_{Q_p}(q_p) = CDF_{Q_p}(q_p) \\ \left[ \Lambda_L \left\{ \int_0^A g(a_L) \frac{k}{(\xi a_L) (E[i_{a_L,\tau}]/\Gamma(1+1/k))^k} \left( \frac{(q_p - q_0)}{(\xi a_L)} + f_{a,L} \right)^{k-1} \exp \left( -\frac{((q_p - q_0)/(\xi a_L) + f_{a,L})^k - (f_{a,L})^k}{(E[i_{a_L,\tau}]/\Gamma(1+1/k))^k} \right) da_L \right\} \right. \\ \left. + \Lambda_H \left\{ \int_0^A g(a_H) \frac{k}{(\xi a_H) (E[i_{a_H,\tau}]/\Gamma(1+1/k))^k} \left( \frac{(q_p - q_0)}{(\xi a_H)} + f_{a,H} \right)^{k-1} \exp \left( -\frac{((q_p - q_0)/(\xi a_H) + f_{a,H})^k - (f_{a,H})^k}{(E[i_{a_H,\tau}]/\Gamma(1+1/k))^k} \right) da_H \right\} \right]. \tag{A18}$$

## Appendix B

### Parameter estimation of the IF and TCIF models

Nine parameters depending on information other than AMFS, are common to the IF and TCIF models and were all available from previous studies (Table 3), they are:

- base flow ( $q_0$ ) estimated as at-site average monthly flow observed in January and February (Fiorentino and Iacobellis, 2001; Claps et al., 2000);
- four parameters strictly related to basin geomorphology ( $A, \tau_A \xi, \beta$ ); basin area  $A$  and lag-time  $\tau_A$ , available in regional studies conducted on basins located in Puglia, Basilicata and Calabria (Iacobellis and Fiorentino, 2000; Fiorentino and Iacobellis, 2001; Fiorentino et al., 2011);  $\beta = 4$  and  $\xi = 0.7$  as discussed in Iacobellis and Fiorentino (2000).
- four rainfall parameters ( $E[i_{A,\tau}], \varepsilon, \Lambda_p, k$ ) estimated, for each basin, exploiting regional frequency analysis of annual maximum rainfall series (AMRS); for the shape parameter  $k$  of the Weibull distribution of rainfall intensity, a regional estimation procedure, based on the PEV distribution (Villani, 1993), was applied to 403 annual maximum daily rainfall series (178 belonging to Puglia and Basilicata and 225 to Calabria) and provided regional values of  $k$  equal to 0.8 in Puglia and Basilicata (Fiorentino and Iacobellis, 2001) and 0.53 in Calabria (Claps et al., 2000);  $\Lambda_p$  estimated as the regional value of the mean annual number of independent rainfall events ( $\Lambda_1$ ) of the ordinary component of the TCEV distribution evaluated in the Italian VAPI reports (e.g. Copertino and Fiorentino, 1994); expected value of space-time average rainfall intensity  $E[i_{A,\tau}]$  evaluated as:

$$E[i_{A,\tau}] = \frac{p_1 \tau_A^{n-1} [1 - \exp(-1.1 \tau_A^{0.25}) + \exp(-1.1 \tau_A^{0.25} - 0.004A)]}{\Lambda_p \sum_{j=0}^{\infty} \frac{(-1)^j \Lambda_p^j}{j!(j+1)^{(1/k+1)}}} \quad (\text{B1})$$

in which the US Weather Bureau areal reduction factor is used,  $p_1$  and  $n$  (Table 3) are the parameters of the intensity-duration-frequency (IDF) curve of the expected annual maximum rainfall intensity.

The regional scaling of  $E[i_{A,\tau}]$ , provides the exponent  $\varepsilon$  for the three regions investigated; in particular for Puglia and Basilicata Fiorentino and Iacobellis (2001) recognized two groups of basins: for basins 1 to 12, 15 and 18  $\varepsilon = 0.39$ , for basins 13–14, 16–17, 19–20,  $\varepsilon = 0.33$ ; for Calabria, Claps et al. (2000) recognized three zones: Tyrrhenian (basins 32 and 33 with  $\varepsilon = 0.28$ ), Central (basins 21 to 24, 26 to 28

and 31 with  $\varepsilon = 0.28$ ) and Ionian (basins 25, 29 to 30, with  $\varepsilon = 0.32$ ).

The loss threshold scaling factors  $\varepsilon', \varepsilon_L, \varepsilon_H$  deserve particular attention. According to Fiorentino and Iacobellis (2001) the exponent of the scaling relationship in Eq. (A6) of the threshold  $f_a$  in humid climates ( $I > 0$ ) is typically  $\varepsilon' = 0$  reflecting the threshold behaviour of a constant infiltration rate, while  $\varepsilon' = 0.5$  is expected in dry climate ( $I < 0$ ) where the threshold  $f_{a,H}$  shows a storage behavior. Fiorentino et al. (2011) analyzed the loss threshold scaling factors  $\varepsilon_L$  and  $\varepsilon_H$  of the TCIF and observed that for all basins, independently from climate, the higher threshold  $f_{A,H}$  scales with exponent  $\varepsilon_H = 0.5$  while the lower threshold scales with exponent  $\varepsilon_L = 0$  in humid basins ( $I > 0$ ) and  $\varepsilon_L = 0.5$  in arid basins ( $I < 0$ ).

*Acknowledgements.* This work was realized with support of PRIN-Cubist-CoFin2007 of the MIUR (Italian Ministry of Instruction, University and Research) and the Calabria Region that funded the project “Statistical methods for the estimate of flood peak events”. We thank the editor and the anonymous reviewers for their useful suggestions.

Edited by: R. Deidda

Reviewed by: two anonymous referees

## References

- Beven, K. J. and Kirkby, M. J.: A physically-based, variable contributing area model of basin hydrology, *Hydrological Sciences Bulletin*, 24, 43–69, 1979.
- 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*, 341–352, 2000.
- Copertino, V. A. and Fiorentino, M.: Valutazione delle piene in Puglia, 177–209, DIFA-Università della Basilicata e CNR-GNDCI, Potenza, 1994 (in Italian).
- 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.
- Eagleson, P. S.: Dynamics of flood frequency, *Water Resour. Res.*, 8(4), 878–898, 1972.
- De Smedt, F., Yongbo, L., and Gebremeskel, S.: Hydrologic modelling on a catchment scale using GIS and remote sensed land use information, edited by: Brebbia, C. A., *Risk Analysis II*, WIT press, Southampton, Boston, 295–304, 2000.
- 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, *Convegno Nazionale: La difesa della montagna*, 2002 (in Italian).
- 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., Gioia, A., Iacobellis, V., and Manfreda, S.: Regional analysis of runoff thresholds behaviour in Southern Italy

- based on theoretically derived distributions, *Adv. Geosci.*, 26, 139–144, doi:10.5194/adgeo-26-139-2011, 2011.
- 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.
- Gioia, A., Iacobellis, V., Manfreda, S., and Fiorentino, M.: Runoff thresholds in derived flood frequency distributions, *Hydrol. Earth Syst. Sci.*, 12, 1295–1307, doi:10.5194/hess-12-1295-2008, 2008.
- 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.
- Gottschalk, 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, 148–162, 1998.
- 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.
- 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(2), 469–482, 2000.
- Iacobellis, V., Fiorentino, M., Gioia, A., and Manfreda, S.: Best Fit and Selection of Theoretical Flood Frequency Distributions Based on Different Runoff Generation Mechanisms, *Water*, 2, 239–256, 2010.
- Klemes, V.: Stochastic models of rainfall-runoff relationship. Statistical Analysis of Rainfall and Runoff, *Proc. Int. Symp. on Rainfall-Runoff Modelling*, Mississippi State University, 139–154, 1982.
- Kurothe, R. S., Goel, N. K., and Mathur, B. S.: Derived flood frequency distribution for negatively correlated rainfall intensity and duration, *Water Resour. Res.*, 33(9), 2103–2107, 1997.
- Manfreda, S., Di Leo, M., and Sole, A.: Detection of Flood Prone Areas using Digital Elevation Models, *J. Hydrol. Eng.*, posted ahead of print 3 January, doi:10.1061/(ASCE)HE.1943-5584.0000367, 2011.
- Manfreda, S.: Runoff generation dynamics within a humid river basin, *Nat. Hazards Earth Syst. Sci.*, 8, 1349–1357, doi:10.5194/nhess-8-1349-2008, 2008.
- Manfreda, S. and Fiorentino, M.: A stochastic approach for the description of the water balance dynamics in a river basin, *Hydrol. Earth Syst. Sci.*, 12, 1189–1200, doi:10.5194/hess-12-1189-2008, 2008.
- Raines, T. H. and Valdes, J. B.: Estimation of flood frequencies for ungauged catchments, *J. Hydraul. Eng.*, 119(10), 1138–1154, 1993.
- Rossi, F., Fiorentino, M., and Versace, P.: Two component extreme value distribution for flood frequency analysis, *Water Resour. Res.*, 20(7), 847–856, 1984.
- Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J. J., Mendiondo, E. M., O'Connell, P. E., Oki, T., Pomeroy, J. W., Schertzer, D., Uhlenbrook, S. and Zehe, E.: IAHS Decade on Predictions in Ungauged Basins (PUB), 2003-2012: Shaping an exciting future for the hydrological sciences, *Hydrol. Sci. J.*, 48(6), 857–880, 2003.
- Thorntwaite, C. W.: An approach toward a rational classification of climate, *Am. Geograph. Rev.*, 38(1), 55–94, 1948.
- Villani, P.: Extreme flood estimation using Power Extreme Value (PEV) distribution, *Proc. IASTED Int. Conf.*, Pittsburgh, 1993.