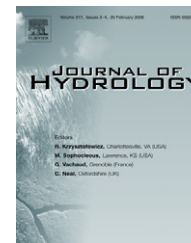




available at www.sciencedirect.com



journal homepage: www.elsevier.com/locate/jhydrol



Baseflow index regionalization analysis in a mediterranean area and data scarcity context: Role of the catchment permeability index

Antonia Longobardi *, Paolo Villani

Department of Civil Engineering, University of Salerno, Via Ponte Don Melillo, 84084 Fisciano (SA), Italy

Received 29 November 2007; received in revised form 25 February 2008; accepted 5 March 2008

KEYWORDS

Baseflow index;
Hydrograph separation;
Regionalization;
Mediterranean area

Summary Low flow characteristics are affected by different physiographic factors such as climate, topography, geology and soils, and regional regression prediction models, to estimate low flow indexes at ungauged sites, mainly rely on these factors. The paper focuses on the baseflow index, one of the most important low flow characteristics for a catchment, and presents: (i) the analysis of baseflow separation algorithms for BFI evaluation and (ii) a regional approach to predict the BFI at ungauged sites in a Mediterranean region, for which only very poor data are available. The prediction of baseflow contribution to total streamflow is based on the introduction of a permeability index, at the catchment scale, and regional linear regression equations simply relate the latter to the BFI. For the studied area geological features have been found to be the major factor affecting baseflow and the permeability index estimation for a particular catchment, in an apparently over-simplified schematization, essentially reflects catchment lithology. As a matter of fact, an accurate catchment geology spatial variability description reduces the average long-term BFI index prediction error from 23% to 14% and above all increases the explained variance from 23% to 68%.

© 2008 Elsevier B.V. All rights reserved.

Introduction

Low flow hydrological features are crucial for efficient development and integrated water resources management and a lot of effort has been made by the scientific community to deal with low flow parameters estimation in

ungauged sites. The statistical multiple linear regression model is one of the most popular approach in these cases. Within a European context, the first significant statistical low flow estimation procedure was proposed by the Institute of Hydrology (1980) and aimed at finding statistical relationships between low flow indexes and catchment characteristics for prediction in ungauged basins, followed a few years later by one of the FRIEND project (Gustard et al., 1989) aimed to improve understanding of

* Corresponding author. Tel.: +39 089 963408; fax: +39 089 964099.
E-mail address: alongobardi@unisa.it (A. Longobardi).

hydrological variability across different regions. Both studies showed the importance of geology, hydrogeology and soil properties in estimating low flow characteristics.

Among others, the BFI index, calculated as the long-term ratio of baseflow volume to total streamflow volume, is one of the most important low flow indexes. Many studies (Vogel and Kroll, 1990; Vogel and Kroll, 1992; Ponce and Shetty, 1995a,b; Nathan et al., 1996; Lacey and Grayson, 1998; Haberlandt et al., 2001; Mwakilila et al., 2002) have demonstrated that it is related to a number of climatic and topographic parameters, to vegetation and soil types, besides catchment geology, but that the latter plays the role of the dominating variable. Variables describing geology features are hard to establish. For this reason, and also depending on the available data quantity and quality, catchment geology, to be used for BFI prediction in ungauged basins, has been accounted for in different fashions. Frequently, soil classes systems, geology–vegetation groups, or combined hydrogeology and soil indexes have been used to this purpose (Gustard et al., 1989; Boorman et al., 1995; Lacey and Grayson, 1998). Very recently, Schneider et al. (2007) have proposed a reclassification of the Soil Geographical Database of Europe (SGDBE) adopting the well-known HOST system developed in the UK (Boorman et al., 1995), to predict the BFI index, in a European context. They have shown that the SGDBE is sufficient for hydrological classification but that the variability of BFI explained by soil classes tends to decrease from Northern to Southern Europe, probably because factors such as climate, vegetation and geomorphology, which are not used to differentiate HOST classes, have a greater influence especially in Mediterranean catchments. Indeed such an approach would require a good knowledge of soil and geology properties and would likely perform poorly in the case of dearth of appropriate data.

In this study, we introduce a permeability index P1, as an alternative variable accounting for geology features, that can be easily derived also in a scarcity data context and particularly suited for typical Mediterranean environment. Initially defined on the base of a hydro-geomorphological classification, successfully used for flood prediction in ungauged sites, it is later computed on the base of an apparently over-simplified scheme which only account for lithological and hydrogeological characteristics of the studied region. It will be shown that the corresponding procedure to compute the permeability index does not require extensive soil surveys, being particularly suited for very poorly gauged sites. The introduced permeability index is further proposed as an independent variable in a regional regression model to predict BFI at ungauged sites. In order to define the most reliable regional relationship between those variables, BFI has been derived by different techniques of baseflow separation, from data analysis to digital filtering algorithms and moreover to a conceptual model approach, and the results have been compared.

In summary, the present paper focuses on the following points:

- application of four baseflow separation algorithms and results comparison to provide the most reliable set of regional model parameters;

- catchment geology – soil and land cover maps analysis and alternative – objective definitions of a catchment permeability index to be used as independent variable in regional prediction equations;
- assessment of linear regression regional relationships between BFI index and catchment permeability index to estimate average baseflow contribution to total streamflow in ungauged catchments and comparison with more regional relationships.

Comparison of BFI indexes derived from baseflow separation procedures

Hydrograph separation has been defined in the past as ‘‘one of the most desperate analysis techniques in use in hydrology’’ (Hewlett and Hibbert, 1967). Indeed the procedures available to this purpose are still, to a large extent, arbitrary (Nathan and McMahon, 1990; Chapman and Maxwell, 1996; Chapman, 1999; Eckhardt, 2005) but provide a repeatable methodology to derive objective measures or indexes related to a particular streamflow source. BFI values used in this paper have been estimated according to the definition of the Institute of Hydrology (1980), that is, baseflow index is the volume of baseflow divided by the volume of total streamflow. Since the baseflow time series are derived by hydrograph separation procedures, the choice of a particular technique and its more or less arbitrary result may affect the BFI evaluation. To account for this influence on the present quantitative study, we applied and compared four different procedures, selected from three main categories of separation algorithms: empirical, filter based and model-based techniques. Comparisons are aimed at finding correlation between filters to evaluate the most likely BFI value for a particular catchment and to assess the stability of each method in the BFI evaluation; a stable solution would be desirable when BFI has to be estimated from short streamflow records and this would be a frequent case for the studied region.

The first procedure used is the empirical definition proposed by Wundt (1958) in one of the early regionalization approaches: it is based on measured daily flow statistics and defines a long-term measure of groundwater outflow as the mean monthly minimum streamflow (MMSF). To compare catchments of different drainage area the mean monthly minimum streamflow has been divided by the mean annual streamflow.

The second procedure has an empirical basis as well and is represented by the Smoothed Minima Technique (SMT) developed by the Institute of Hydrology (1980), and already applied for low flow regimes analysis in the areas of Central Italy (Casadei, 1995).

Among filtering techniques, the Lyne and Hollick (1979) method is used in this study as the third method for baseflow separation (RDF), as it appears to be the first proposed and a widely used algorithm. Filtering techniques, that act as a low-pass filter, filtering out the high frequency quickflow component of streamflow from the low frequency baseflow component of streamflow, are recommended for providing reproducible results (Lyne and Hollick, 1979; Chapman, 1991; Boughton, 1993; Chapman and Maxwell, 1996). The Lyne and Hollick filter equation predicts the quickflow q_q component at time step t as:

$$q_q(t) = \alpha q_q(t-1) + \frac{1+\alpha}{2} [q(t) - q(t-1)] \quad (1)$$

subject to the restriction $q_q > 0$ and the baseflow component q_b at time step t as the difference between total streamflow q and quickflow q_q :

$$q_b(t) = q(t) - q_q(t) \quad (2)$$

subject to the restriction $q_b \leq q$, where α is the filter parameter affecting the degree of attenuation. Accordingly to Nathan and McMahon (1990), the value of the filter that yield the most acceptable results, in term of baseflow separation, is in the range 0.9–0.95. The filter is passed three times over the data, forward, backward and forward again, for a larger smoothing effect, as suggested by Nathan and McMahon (1990). A few high-attenuation filter passes over the data to minimize phase distortion and baseflow attenuation was, moreover, suggested by Spongberg (2000).

The problem of identifying baseflow hydrograph on a continuous basis has also been dealt with physically based runoff models that explicitly account for baseflow modeling modules. There exist a large number of proposed models in the literature which can be potentially used to this aim. Among these we focus our attention on a flexible stochastic

shot-noise model proposed by Murrone et al. (1997) to model daily streamflow series for poorly gauged watersheds. The model deterministic response function is assumed to be linear and to describe the functioning of four linear parallel conceptual reservoirs, each related to a conceptual source of discharge: c_3 groundwater component, c_2 deep subsurface flow component, c_1 shallow subsurface flow component and c_0 surface runoff component. The system response to a rainfall event is the outcome of a linear combination of the single reservoir response functions

$$h(t) = c_0 \delta(0) + c_1 \frac{1}{K_1} \exp\left(-\frac{t}{K_1}\right) + c_2 \frac{1}{K_2} \exp\left(-\frac{t}{K_2}\right) + c_3 \frac{1}{K_3} \exp\left(-\frac{t}{K_3}\right) \quad t \geq 0 \quad (3)$$

where c_i ($i = 0, 1, 2, 3$) represents the recharge coefficients and K_i the linear reservoir response times. For daily data used in this study, the surface runoff component c_0 delay time is smaller than the time scale of aggregation and can thus be modeled as a random uncorrelated linear channel process rather than a linear reservoir process, with $\delta(0)$ the Dirac impulse, for $t = 0$. This approach has been applied as the fourth method (conceptual filter method CFM) for

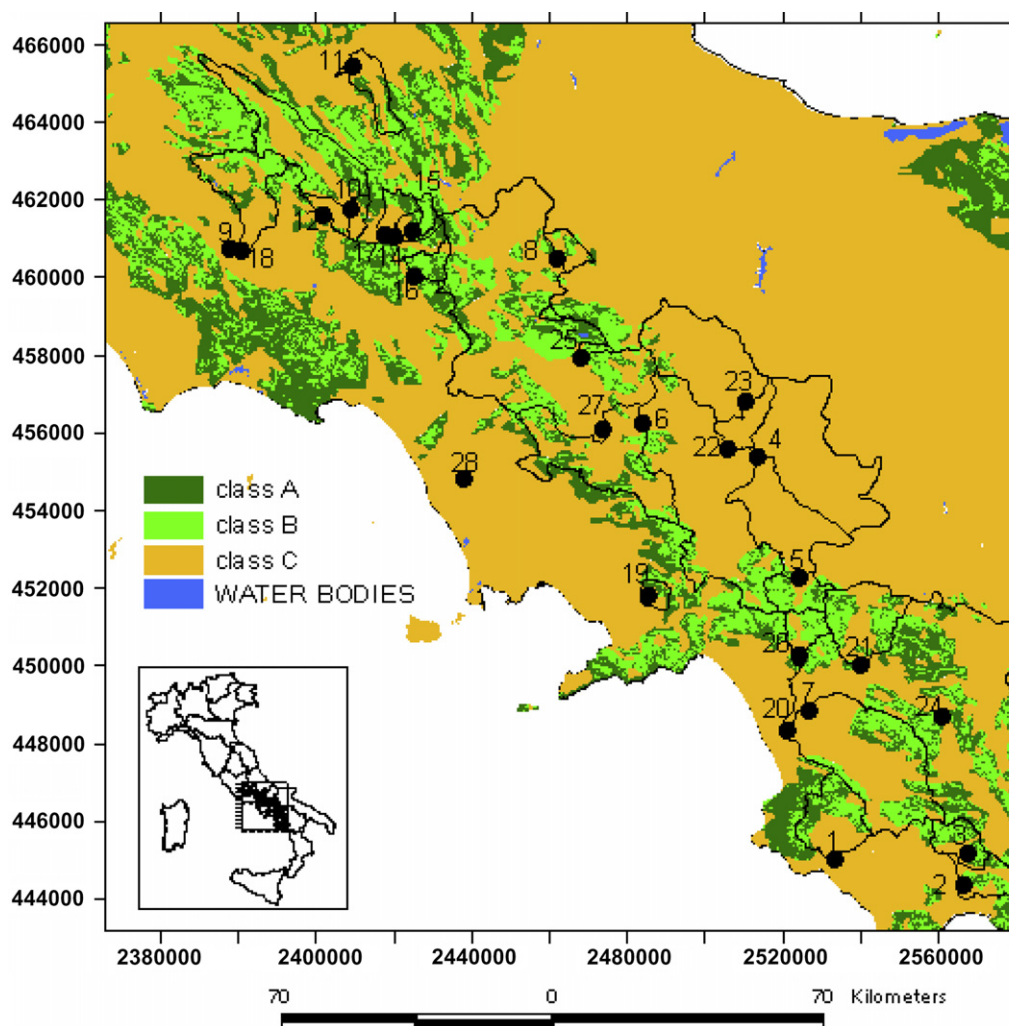


Figure 1 The study area, stations location and hydro-geomorphological classes.

baseflow separation. To set up a repeatable procedure to derive the BFI index related to the baseflow discharge, the latter has been assumed as the sum of c_2 deep subsurface flow and c_3 groundwater processes.

Hydrologic data consist of daily streamflow time series for 28 sites mainly located within the Campania region, Southern Italy, covering a region of about 25,000 km². The location of the gauging stations is shown in Fig. 1. The source of streamflow data is the Servizio Idrografico e Mareografico Italiano SIMI. The record length ranges from 6 to 65 years. The size of the basins ranges between 10 and 5000 km². More catchment data are presented in Table 1 that will be discussed later.

Hydrograph separation performances are assessed in terms of average annual BFI and standard deviation and in terms of π , the mean percentage by which estimated baseflow deviates from minimum annual streamflow, also used by Furey and Gupta (2001) in assessing the filter they proposed. Small values of standard deviation indicate a stable estimate of BFI, that is desirable when only short streamflow time series are available for baseflow index estimation. The Institute of Hydrology (1980) report a standard deviation of 0.04 for BFI stable estimation. Large values of π suggest that estimated baseflow time series are not representative of true baseflow. Although none of the presented techniques would confidently determine baseflow

contribution without field observation, we assume that the most reliable technique is the procedure which exhibits the lower standard deviation and π values. The mentioned statistics are illustrated in Table 2 for each of the hydrograph separation algorithms we applied. SMT and RDF show the smaller standard deviation values, in both cases around 0.07, whereas MMSF and CFM appear to be the less stable techniques with values of 0.13 and 0.11, respectively, for the standard deviation. The mean percentage by which estimated baseflow deviates from minimum annual streamflow is, on average, 25% for SMT, 32% for RDF and 30% for CFM (it is not possible to evaluate π for MMSF technique because it only gives one value of baseflow per year). It is, moreover, systematically higher for RDF compared to SMT. CFM has the benefit to be a conceptual approach, predicting BFI on a physical processes base rather than empirically, but its performance is likely to be affected by model parameters estimation. Based on these comments, in spite of its empirical basis, the Smoothed Minima Technique is indicated as the most reliable procedure for our study area.

Long-term BFI indexes, the ratio between the whole separated baseflow and streamflow series volumes, derived from the previously illustrated approaches are also compared in Table 3 and graphically in Fig. 2. The results show that the correlation between the hydrograph separation methods is rather large and, in particular, the stronger cor-

Table 1 Some details of catchments used in the study: reference map number, drainage area, mean annual discharge, mean annual precipitation, mean topographic catchment elevation, mean topographic catchment slope and percentage of forested area

Catchment	Map reference	Area (km ²)	Q_{mean} (m ³ /s)	MAP (mm)	H_{mean} (m)	p_{mean}	FOR
Alento @ Casalvelino	1	270	4.33	1254	328	10.4	52
Bussento @ Caselle in Pittari	2	125	5.19	1398	667	12.7	62
Bussento @ Sicili	3	255	6.48	1589	779	13.4	65
Calore Irpino @ Apice	4	532	8.64	1119	611	8.2	30
Calore Irpino @ Montella	5	130	2.11	1417	1004	13.7	71
Calore Irpino @ Solopaca	6	2991	33.25	1107	546	7.4	22
Calore Lucano @ Persano	7	805	21.93	1362	658	10.4	40
Carpino @ Carpinone	8	69	1.71	1077	898	8.9	55
Cosa @ Ceccano	9	312	1.61	1417	728	9.4	36
Fibreno @ Brocco	10	47	9.39	1369	491	8.3	32
Giovenco @ Pescina	11	138	1.00	859	1298	13.2	26
Liri @ Isola Liri	12	586	28.09	1251	1024	12.1	52
Liri @ Sora	13	480	14.88	1107	1060	12.3	57
Melfa @ Atina	14	86	3.96	1572	1397	21.1	49
Melfa @ Picinisco	15	38	0.91	1492	1074	15.8	78
Rapido @ S. Elia Fiumerapido	16	73	1.75	1463	816	12.4	60
Rio Mollo @ Settignano	17	69	0.71	1378	946	13.0	26
Sacco @ Ceccano	18	912	11.21	1268	439	7.3	25
Sarno @ S. Valentino Torio	19	46	8.64	1032	240	9.8	11
Sele @ Albanella	20	3216	55.61	1197	684	10.0	39
Sele @ Contursi	21	322	10.12	1336	711	12.4	46
Tammara @ Paduli	22	675	9.26	992	597	6.1	17
Tammara @ Pago Veiano	23	558	6.02	1102	633	6.1	20
Tanagro @ Polla	24	660	9.72	1288	792	9.4	39
Torano @ Piedimonte Matese	25	18	3.00	1283	949	15.1	78
Tuscano @ Olevano sul Tusciano	26	102	3.60	1633	961	15.8	85
Volturno @ Amorosi	27	2029	36.69	1387	590	9.4	45
Volturno @ Canello Arnone	28	5586	80.45	1183	534	8.2	31

Table 2 Comparison of annual BFI index mean and standard deviation and π percentage by which estimated baseflow deviates from annual minimum streamflow

Catchment	MMSF		SMT			RDF			CFM		
	Mean	SD	Mean	SD	π	Mean	SD	π	Mean	SD	π
Alento @ Casalvelino	0.08	0.063	0.295	0.102	18.41	0.389	0.106	37.84	0.401	0.121	31.91
Bussento @ Caselle in Pittari	0.47	0.113	0.661	0.054	26.76	0.751	0.056	34.36	0.657	0.110	40.12
Bussento @ Sicili	0.59	0.077	0.629	0.074	14.17	0.748	0.053	24.82	0.600	0.136	13.17
Calore Irpino @ Apice	0.28	0.174	0.492	0.073	18.15	0.579	0.088	13.34	0.567	0.109	20.52
Calore Irpino @ Montella	0.19	0.151	0.526	0.105	5.98	0.590	0.097	7.94	0.562	0.112	21.48
Calore Irpino @ Solopaca	0.14	0.088	0.393	0.088	38.19	0.477	0.093	49.70	0.495	0.113	46.75
Calore Lucano @ Persano	0.17	0.083	0.348	0.086	28.70	0.407	0.092	21.80	0.411	0.097	33.89
Carpino @ Carpinone	0.45	0.247	0.642	0.061	27.99	0.718	0.056	35.59	0.718	0.099	20.05
Cosa @ Ceccano	0.13	0.110	0.433	0.107	30.53	0.507	0.088	45.86	0.511	0.082	40.03
Fibreno @ Brocco	0.77	0.199	0.838	0.016	9.54	0.914	0.026	15.09	0.722	0.109	>> 100
Giovenco @ Pescina	0.61	0.142	0.662	0.048	11.41	0.772	0.045	22.67	0.641	0.111	16.69
Liri @ Isola Liri	0.57	0.124	0.664	0.052	20.34	0.771	0.058	29.88	0.654	0.142	21.53
Liri @ Sora	0.39	0.169	0.588	0.077	31.42	0.676	0.077	48.43	0.610	0.107	38.62
Melfa @ Atina	0.30	0.210	0.614	0.067	28.21	0.637	0.083	35.47	0.670	0.177	35.69
Melfa @ Picinisco	0.43	0.253	0.626	0.071	43.84	0.690	0.072	52.23	0.656	0.137	50.39
Rapido @ S. Elia Fiumerapido	0.79	0.160	0.797	0.034	11.40	0.907	0.039	19.28	0.703	0.091	4.84
Rio Mollo @ Settignano	0.05	0.047	0.227	0.081	66.56	0.284	0.095	<<100	0.324	0.103	79.47
Sacco @ Ceccano	0.11	0.064	0.308	0.090	50.14	0.412	0.082	60.57	0.375	0.106	51.50
Sarno @ S. Valentino Torio	0.77	0.203	0.821	0.023	16.65	0.903	0.034	21.65	0.793	0.061	19.62
Sele @ Albanella	0.28	0.171	0.500	0.098	34.07	0.583	0.094	43.70	0.483	0.140	35.84
Sele @ Contursi	0.49	0.105	0.572	0.112	19.77	0.657	0.203	38.63	0.534	0.122	8.75
Tammaro @ Paduli	0.02	0.021	0.368	0.113	12.88	0.436	0.093	33.10	0.490	0.117	40.15
Tammaro @ Pago Veiano	0.04	0.066	0.352	0.097	1.34	0.437	0.074	11.78	0.442	0.145	29.84
Tanagro @ Polla	0.23	0.084	0.439	0.066	42.87	0.547	0.070	51.66	0.460	0.092	43.34
Torano@ Piedimonte Matese	0.76	0.249	0.811	0.036	10.63	0.891	0.040	19.43	0.664	0.179	6.42
Tusciano @ Olevano sul Tusciano	0.50	0.069	0.694	0.054	24.49	0.781	0.054	32.90	0.717	0.114	25.32
Volturno @ Amorosi	0.34	0.175	0.490	0.105	18.31	0.602	0.104	28.51	0.534	0.089	16.32
Volturno @ Cancellone Arnone	0.32	0.128	0.504	0.095	20.260	0.611	0.080	33.430	0.513	0.119	26.89

relation ($R^2 = 0.98$) has been found between BFI from SMT and BFI from RDF. Overall, the linear regression least square fit relationships have a 45° slope coefficient and a practically negligible intercept, except that for MMSF. BFI from MMSF appears to be biased of about 30% compared to BFI from SMT: values from MMSF are systematically lower. BFI from CFM also appears to be biased compared to BFI from SMT, with systematically higher values.

On the role of the catchment permeability index in a mediterranean environment: the case study

The studied region

The study region is a complex relief area, with inland highlands running north-west to south-east and wide and flat plains facing the Tirrenian Sea, where all the river channels included in this analysis flow into. The geology is rather variable: it includes marly clayey impermeable complex in the north-east area, fissured calcareous and dolomitic complex in the central area, representing the most important regional aquifer with the highest potential infiltration coefficient, and alluvial complex along the coastline. About 30% of the study area can be consid-

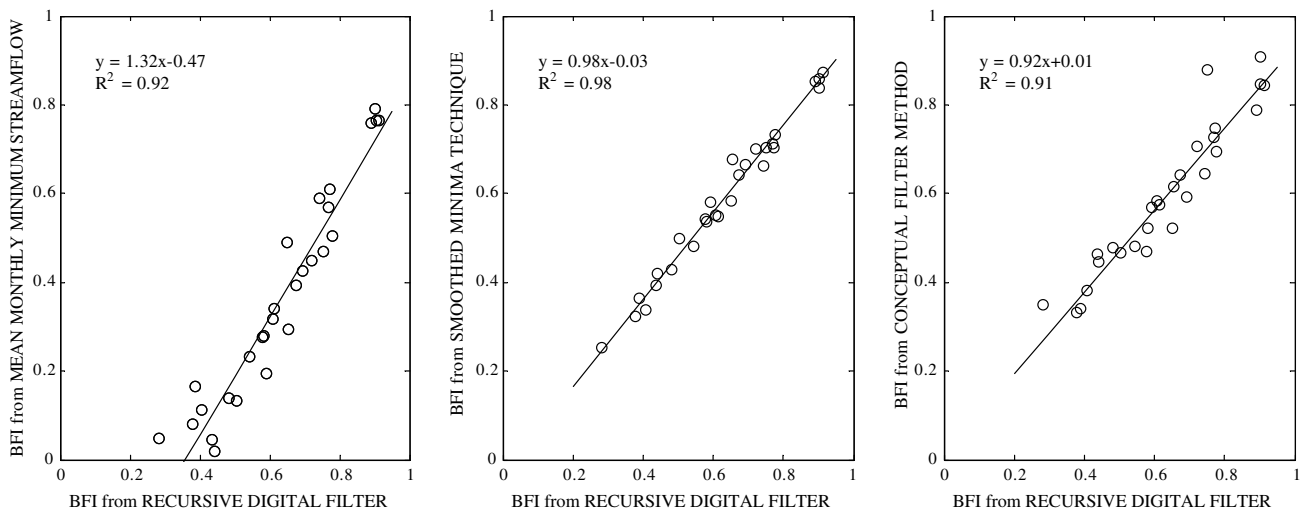
ered as permeable and about 32% is covered by forest. The following physiographic characteristics have been computed for each catchment: drainage area (km^2), mean annual precipitation (mm), mean topographic catchment elevation (m), mean topographic catchment slope (%), land use (%) and geology (%). Land use is represented by the percentage of forested area and geology is represented by the introduced permeability index P1. The computation of this index will be later illustrated. Some details of the catchments are given in Table 1.

Hydro-geomorphological classification and BFI prediction

With regard to the infiltration process, geological and morphological features of the studied region make possible the identification of three main hydro-geomorphological groups that can be used in water balance assessment (Celico et al., 1989). Groups result from the combination of three variables. These are geology, mean topographic catchment slope (p_{mean}) and forested percentage area (FOR). Celico et al. (1989) suggested to use geology data to replace insufficient soil properties data, as it is known that soil is the result of the rock weathering processes and its properties are highly related to the substrate

Table 3 Long-term BFI indexes comparison

Catchment	BFI			
	MMSF	SMT	RDF	CFM
Alento @ Casalvelino	0.08	0.32	0.38	0.33
Bussento @ Caselle in Pittari	0.47	0.70	0.75	0.88
Bussento @ Sicili	0.59	0.66	0.74	0.64
Calore Irpino @ Apice	0.28	0.54	0.58	0.52
Calore Irpino @ Montella	0.19	0.58	0.59	0.57
Calore Irpino @ Solopaca	0.14	0.43	0.48	0.48
Calore Lucano @ Persano	0.17	0.36	0.39	0.34
Carpino @ Carpinone	0.45	0.70	0.72	0.71
Cosa @ Ceccano	0.13	0.50	0.50	0.46
Fibreno @ Brocco	0.77	0.87	0.91	0.84
Giovenco @ Pescina	0.61	0.70	0.77	0.75
Liri @ Isola Liri	0.57	0.71	0.77	0.73
Liri @ Sora	0.39	0.64	0.67	0.64
Melfa @ Atina	0.30	0.68	0.65	0.61
Melfa @ Picinisco	0.43	0.66	0.69	0.59
Rapido @ S. Elia Fiumerapido	0.79	0.84	0.90	0.85
Rio Mollo @ Settignano	0.05	0.25	0.28	0.35
Sacco @ Ceccano	0.11	0.34	0.41	0.38
Sarno @ S. Valentino Torio	0.77	0.86	0.90	0.91
Sele @ Albanella	0.28	0.54	0.58	0.47
Sele @ Contursi	0.49	0.58	0.65	0.52
Tammaro @ Paduli	0.02	0.42	0.44	0.45
Tammaro @ Pago Veiano	0.04	0.39	0.44	0.46
Tanagro @ Polla	0.23	0.48	0.54	0.48
Torano@ Piedimonte Matese	0.76	0.85	0.89	0.79
Tuscano @ Olevano sul Tusciano	0.50	0.73	0.78	0.69
Volturno @ Amorosi	0.34	0.55	0.61	0.57
Volturno @ Canello Arnone	0.32	0.55	0.61	0.58

**Figure 2** Relationships between BFI values estimated by MMSF, SMT, RDF and CFM approaches.

lithology. Highly fractured carbonatic deposits are one of the main features of the studied region; they are regarded as the pervious catchment areas. Hydro-geomorphological groups represent classes of soils–geology–vegetation having the same runoff potential under similar

storm and cover conditions and are listed below in decreasing permeability order:

- (1) bare calcareous and dolomitic complex areas with mean catchment slope above 10% (class A);

- (2) forested calcareous and dolomitic complex areas with mean catchment slope above 15% (class B);
- (3) impervious areas with mean catchment slope below 10% (class C).

Extensive soil surveys are not needed for the identification of the mentioned classes: soils – geology types are derived from 1:100.000 digital geological map, land cover is derived by the Corine Land Cover system and mean topographic catchment slope is derived from DEMs. This classification, illustrated in a GIS map in Fig. 1 for the studied region, has been successfully used for flood prediction in ungauged basins in the same region we investigated (VALutazione delle Plene in Campania; Rossi and Villani, 1994). A runoff coefficient was associated to each class, with class A being the lowest production runoff class and class C being the highest runoff production class. Forest covering carbonatic areas causes, in this region, a reduction of the carbonatic high infiltration capacity of about 40–50% thus land cover, besides catchment geology, was indicated as a main variable for regional flood prediction.

The same hydro-geomorphological classification is initially proposed in this study for regional BFI prediction, assigning a BFI value to each class. This approach is similar to other studies that documented high values of BFI explained variance: Boorman et al. (1995) found the 79% of explained variance in UK; Schneider et al. (2007) found the 68% in England and Wales, whereas outside this region the explained variance decreases, especially for Mediterranean basins. The BFI for a particular catchment can be computed as

$$BFI = BFI_A \frac{A_A}{A} + BFI_B \frac{A_B}{A} + BFI_C \frac{A_C}{A} + BFI_{min}, \quad (4)$$

where BFI_i is the BFI associated to the class i , BFI_{min} is a constant value added to account for the observed BFI lower limit of about 30%, A is the catchment drainage area and A_i is the catchment area which belongs to the class i . BFI_i are estimated as the least square regression coefficients based on the discharge data of the studied catchments. A regional low flow index estimation approach based on the hydro-geomorphological classification performed very poorly for the studied area: explained variance only amount to 17%, with underestimated large BFI values and overestimated small BFI values.

Flood events occur on a small temporal scale, and to estimate the flood runoff volume it is very important to understand in which measure each portion of the catchment contributes. The BFI is instead the result of a long-term water balance and in this case it would be more important to understand whether a particular portion of the catchment contributes rather than in which measure it contributes. In the Mediterranean region, a typical runoff process is the infiltration excess process and the infiltration phenomenon itself is a threshold phenomenon limited by the soil properties. In this sense, it is plausible to consider that the vegetation has not a relevant role on long-term water balance. In the light of these comments, a simple partition of catchment area into permeable areas and impervious areas can be proposed for BFI regional estimation. This partition is only based on lithological and hydrogeological characteristics of the studied region neglecting the land use

effect. In this simplified schematization, permeable areas A_{perm} , that for the case study correspond to the sum of A_A and A_B , are the only source of baseflow contribution and Eq. (4) becomes

$$BFI = BFI_{perm} \frac{A_{perm}}{A} + BFI_{min} \quad (5)$$

with BFI_{perm} being the BFI value of the permeable complex. The ratio A_{perm}/A is indicated as the catchment permeability index P1 and the regional BFI index is then calculated as

$$BFI = BFI_{perm} \times P1 + BFI_{min} \quad (6)$$

with BFI_{perm} and BFI_{min} being the least square regional regression parameters based on the discharge and catchment features data of the studied area. The goodness-of-fit of this model will be discussed in the following paragraph, but it is worth to notice here that the region partition into permeable areas and impervious areas, make possible (i) the introduction of a permeability index, at the catchment scale, which summarizes, in a single measure, the whole catchment features, with regard to the infiltration process, and (ii) the prediction of the BFI from such measure in a data parsimonious regional regression approach.

Results and discussion

Regional regression and multiple regional regression approaches

Regional BFI prediction in ungauged catchments is based on the introduced permeability index P1 used as independent variable in a simple linear regression model:

$$BFI = BFI_{perm} \times P1 + BFI_{min} = A \times P1 + B \quad (7)$$

Even though the Smoothed Minima Technique has been found to be the most reliable baseflow separation algorithm for the studied area, since one of the goal of the paper was to find the optimal regional regression model for BFI prediction, all of the combinations between the permeability index and separation algorithm have been investigated to find the optimal combination. To this aim regional regression coefficients A and B have been estimated for each of the hydrograph separation procedures and regional prediction equations are illustrated in Fig. 3. Empirical points ($P1$, BFI) have been removed, highlighting the regression lines and R^2 as the measure of the degree of dispersion around them. Prediction equations, derived from different hydrograph separation methods, are almost parallel meaning that the relationship between BFI and P1 does not change for different BFI evaluation. Estimates given by SMT, RDF and CFM are comparable with a comparable explained variance of about 63% for the RDF method, about 68% for the SMT method and about 53% for the CFM method. The use of the MMSF method instead induces underestimated values of regional BFI of about 30–40% compared to other baseflow separation techniques, as it could have been already observed in Fig. 2.

Geological influence on baseflow has been found to be dominant within the studied area. Fig. 4 shows scatter plots and correlation coefficients for BFI, H_{mean} mean topographic catchment elevation, p_{mean} mean topographic catchment

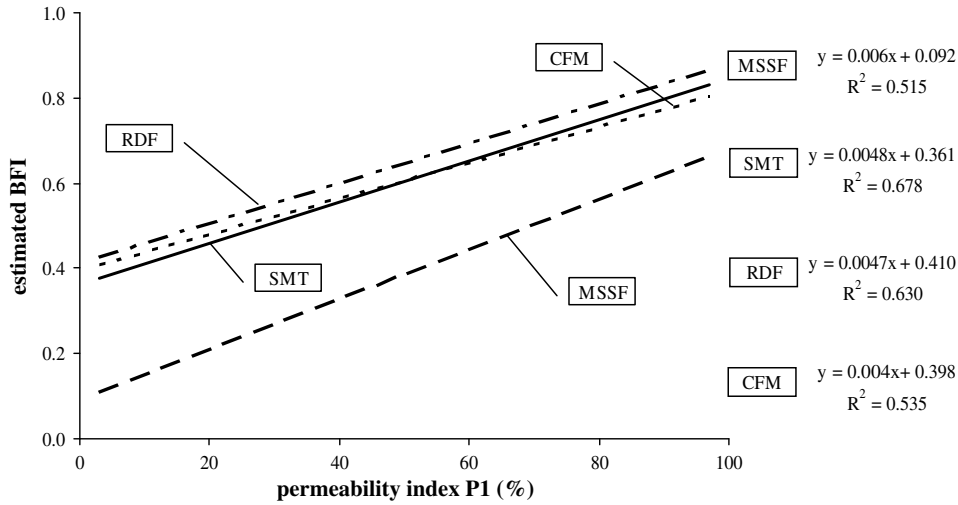


Figure 3 Regional regression prediction equations relating BFI to P1 permeability index.

BFI					
0.67	P1				
0.08	0.25	H _{mean}			
0.17	0.64	0.56	p _{mean}		
0.02	0.41	0.06	0.36	MAP	
0.17	0.39	0.23	0.45	0.46	FOR

Figure 4 Scatter plot and correlation matrix for BFI and basin and climate characteristics (P1 permeability index, H_{mean} mean topographic catchment elevation, p_{mean} mean topographic catchment slope, MAP mean annual precipitation, FOR percentage of forested area).

slope, MAP mean annual precipitation, FOR percentage of forested area and P1 permeability index. Row and column intersection, for a couple of target variables, defines, in the upper part of the matrix, the scatter plot between such variables and, in the lower part of the matrix, their degree of correlation. The BFI appears to be very poorly related to the mentioned physical catchment descriptors except that for the P1 permeability index. The weak correlation coefficient between the BFI and the percentage of forested area, highlights the fact that while for flood prediction land cover data are of great importance this is not the case for low flow indexes prediction in the studied region, underpinning the choice to ignore the land cover effect and simply partition catchment area on the base of lithological features, to cal-

culate the permeability index. The studied region is characterized by highly fractured carbonatic deposits and their features explain some of the stronger correlations illustrated in Fig. 4, such as the positive trend between H_{mean} and p_{mean} (typical for carbonatic relieves) and the positive trend between p_{mean} and FOR, with the percentage of forested area being larger for higher catchment slope, that is for the carbonatic relieves. Fig. 4 also shows a weak correlation between the mean annual precipitation MAP and the BFI proving that climate seems not to represent a key variable in this study.

Even though the database consists of very poor data, the introduction of H_{mean} , p_{mean} , FOR and MAP besides P1, in a multiple linear regression equation

$$\begin{aligned} \text{BFI} &= \text{BFI}_{\text{perm}} \times \text{P1} + \text{BFI}_{\text{min}} \\ &= a_1 \times \text{P1} + a_2 \times H_{\text{mean}} + a_3 \times p_{\text{mean}} + a_4 \times \text{FOR} + a_5 \\ &\quad \times \text{MAP} + b \end{aligned} \quad (8)$$

increases the explained variance from about 68% to about 80%, the latter corrected to account for the larger number of parameters to be estimated compared to the case of a single independent variable, according to the following:

$$R_{\text{cor}}^2 = 1 - (1 - R^2) \frac{k - 1}{k - p} \quad (9)$$

with p is the number of model parameters and k is the sample length. It can be observed that the increase in the explained variance is rather significant but the least square estimate parameters ($a_1 = 0.0066$; $a_2 = -6.28E-05$; $a_3 = -0.0046$; $a_4 = 0.0012$; $a_5 = -0.00047$; $b = 0.9252$) do not appear as physically based as they appear in the case of the simple regression. Despite this, it is however clear that lithological properties are more relevant, in the studied region, as it would have been expected from the correlation matrix inspection. For these reasons we proceed in the following analysis referring to the simple linear regression model (6). For an overall view, comparison between observed and predicted BFI from regional multiple regression and hydro-geomorphological classes is given in Fig. 5.

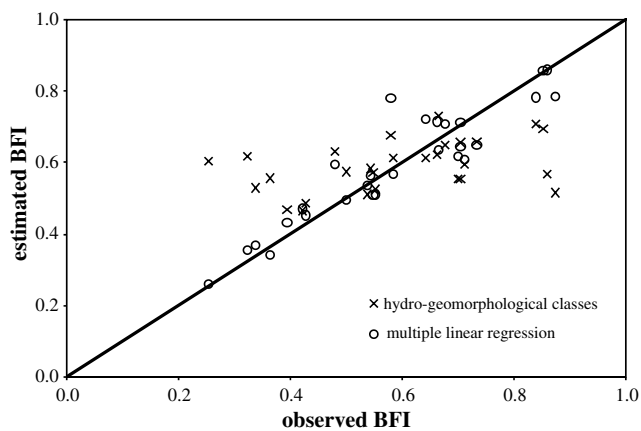


Figure 5 Observed versus predicted BFI derived by the hydro-geomorphological proposed classification (x) and a multiple linear regional regression model (o).

Regional relationships assessment and comparison

Since the permeability index P1 is the only independent variable used in this study for regional BFI prediction, we expect regional regression approaches performances to be affected, to a large extent, by P1 and particular care has to be taken in its computation. In the past, a permeability index was already introduced within a national context, as a variable accounting only for geological features and it is interesting to compare different P evaluation and their performances in regional models assessment.

The earliest permeability index was introduced in the 1930s from the Servizio Idrografico e Mareografico Italiano SIMI, responsible, at that time, for the national rainfall and streamflow gauging network management. It was the result of a geological classification and derived only from old geological maps completed in the 1920s. We refer to this index as P2. Subsequently, a second index was introduced within the national SIVAPI flood forecasting project (Sistema Informativo territoriale per la VALutazione delle Plene in Italia, Gabriele, 1998). It was computed with the same cri-

teria as for P2, but from updated 1:500.000 geological maps. We refer to this index as P3. Estimation from regional regression model using P2 and P3 as independent variables is compared with estimation obtained using the proposed P1 as an independent variable. Table 4 lists the permeability index values for each of the catchment included in the

Table 4 Catchment permeability index, according to different evaluations

Catchment	P1 (%)	P2 (%)	P3 (%)
Alento @ Casalvelino	3	52	52
Bussento @ Caselle in Pittari	70	90	90
Bussento @ Sicili	85	85	85
Calore Irpino @ Apice	26	20	68
Calore Irpino @ Montella	84	67	74
Calore Irpino @ Solopaca	13	11	34
Calore Lucano @ Persano	14	14	14
Carpino @ Carpinone	34	65	65
Cosa @ Ceccano	42	62	62
Fibreno @ Brocco	80	89	89
Giovenco @ Pescina	35	64	64
Liri @ Isola Liri	49	40	49
Liri @ Sora	55	43	43
Melfa @ Atina	97	65	52
Melfa @ Picinisco	68	52	100
Rapido @ S. Elia Fiumerapido	87	68	68
Rio Mollo @ Settignano	10	65	65
Sacco @ Ceccano	10	60	60
Sarno @ S. Valentino Torio	70	63	63
Sele @ Albanella	36	33	33
Sele @ Contursi	47	47	47
Tammaro @ Paduli	8	6	30
Tammaro @ Pago Veiano	10	7	27
Tanagro @ Polla	48	95	95
Torano @ Piedimonte Matese	85	100	100
Tuscano @ Olevano sul Tusciano	78	70	70
Volturno @ Amorosi	39	32	90
Volturno @ Canello Arnone	26	22	43

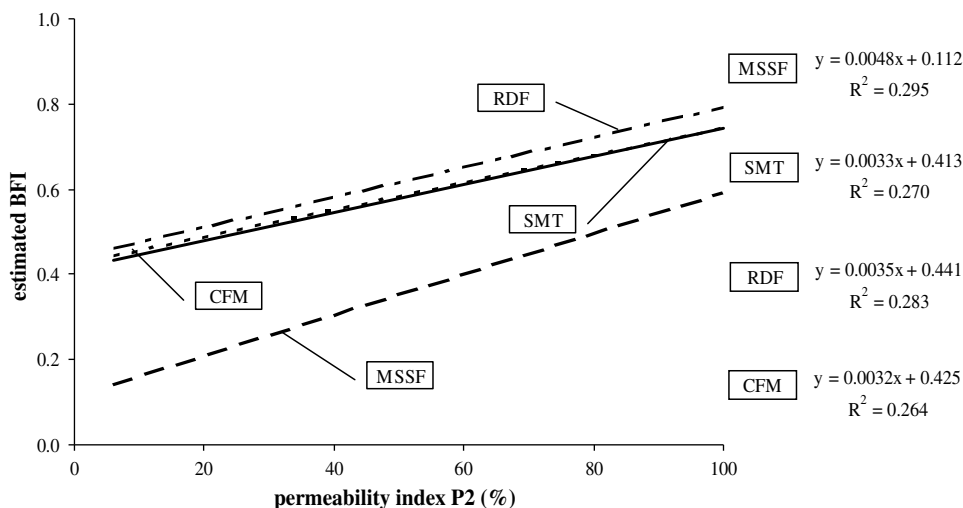


Figure 6 Regional regression prediction equations relating BFI to P2 permeability index.

present analysis. Differences are rather evident between P1, P2 and P3, with P2 and P3 comparable but not identical and P1, on average, smaller than both P2 and P3. Figs. 6 and 7 show empirical regional prediction equations for BFI evaluation in ungauged sites, derived, respectively, for P2 and P3. It can be observed that

1. P2 and P3 have comparable BFI predictive ability, caused by the similarity in their evaluation;
2. P1 has a better BFI predictive ability, compared to P2 and P3, for each of the used hydrograph separation techniques, with the larger least squares LS correlation coefficient R^2 related to the SMT method (0.68);

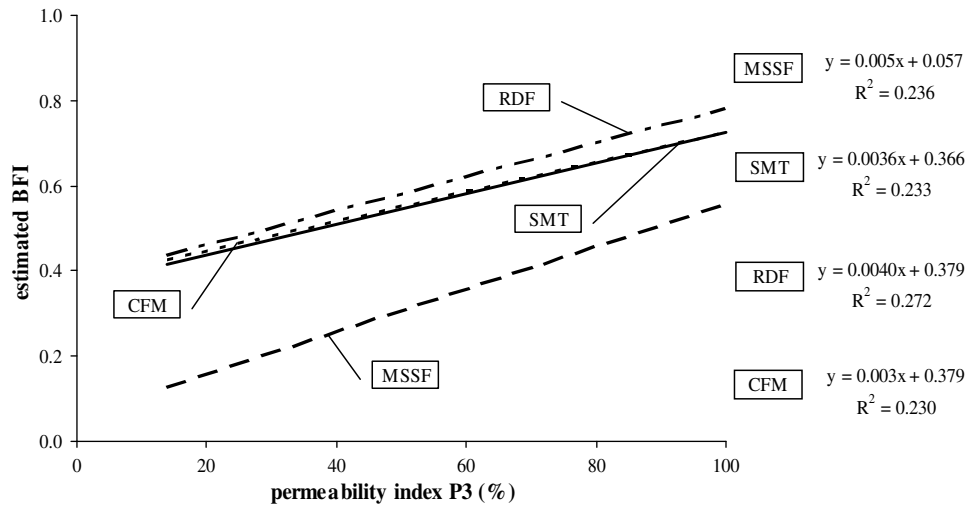


Figure 7 Regional regression prediction equations relating BFI to P3 permeability index.

Table 5 Long-term BFI prediction error (%) for each permeability index and for each baseflow separation procedure

Catchment	P1				P2				P3			
	MMSF	SMT	RDF	CFM	MMSF	SMT	RDF	CFM	MMSF	SMT	RDF	CFM
Alento @ Casalvelino	37	12	17	24	352	64	81	79	296	55	72	70
Bussento @ Caselle in Pittari	8	2	1	21	16	1	1	19	8	2	2	21
Bussento @ Sicili	1	9	16	17	12	1	5	8	18	3	2	5
Calore Irpino @ Apice	12	9	10	3	25	12	11	6	42	12	14	18
Calore Irpino @ Montella	205	36	32	32	125	14	9	12	121	14	9	12
Calore Irpino @ Solopaca	22	2	1	5	20	1	5	4	64	7	14	4
Calore Lucano @ Persano	5	23	18	34	8	60	27	38	24	42	15	26
Carpino @ Carpinone	35	21	25	24	6	7	10	10	15	11	14	14
Cosa @ Ceccano	154	21	13	23	206	31	24	34	174	25	18	28
Fibreno @ Brocco	26	14	15	13	30	18	19	16	35	19	21	18
Giovenco @ Pescina	51	26	25	27	31	14	11	16	38	18	15	19
Liri @ Isola Liri	33	17	16	17	46	24	23	24	47	25	24	24
Liri @ Sora	6	1	2	2	19	12	13	12	31	18	19	17
Melfa @ Atina	125	33	22	31	44	2	7	3	7	10	18	9
Melfa @ Picinisco	16	5	4	15	15	10	12	1	31	12	9	23
Rapido @ S. Elia Fiumerapido	23	9	7	10	45	25	24	24	50	28	27	27
Rio Mollo @ Settignano	213	61	62	26	778	136	148	82	690	125	138	74
Sacco @ Ceccano	33	12	22	15	254	60	82	61	215	52	73	54
Sarno @ S. Valentino Torio	34	18	19	24	46	27	28	31	52	30	31	34
Sele @ Albanella	10	1	2	17	2	4	4	13	20	11	11	5
Sele @ Contursi	25	3	1	14	31	7	3	10	41	13	8	4
Tammaro @ Paduli	599	2	5	3	609	5	3	1	938	13	13	9
Tammaro @ Pago Veiano	242	5	4	5	230	7	11	4	333	12	18	2
Tanagro @ Polla	61	17	23	25	144	42	52	52	128	40	48	48
Torano @ Piedimonte Matese	22	9	10	4	22	11	13	6	27	12	15	8
Tuscano @ Olevano sul Tusciano	10	1	1	5	11	12	12	6	19	15	16	10
Volturno @ Amorosi	6	3	1	2	22	10	5	8	49	21	26	21
Volturno @ Cancellò Arnone	22	12	12	13	31	15	12	15	14	9	5	9

Bold values provide the minimum prediction errors associated to a particular catchment.

Table 6 Regional prediction equations performances (std, bias, RMSE, R^2), jackknife errors estimates (mt, ma, me) and 95% confidence interval for A intercept and B slope prediction equations parameters

Regression equation	std	bias	RMSE	R^2	mt	ma	me	A_{conf}	B_{conf}
P1-MMSF	0.235	0.002	0.235	0.515	0.795	0.725	0.326	± 0.120	± 0.0022
P1-SMT	0.169	-0.002	0.169	0.678	0.149	0.137	0.058	± 0.070	± 0.0013
P1-RDF	0.176	-0.001	0.171	0.63	0.148	0.136	0.058	± 0.076	± 0.0014
P1-CFM	0.165	0.001	0.165	0.535	0.174	0.162	0.066	± 0.082	± 0.0015
P2-MMSF	0.235	-0.001	0.235	0.295	1.231	1.135	0.477	± 0.169	± 0.0028
P2-SMT	0.169	0.001	0.169	0.27	0.248	0.233	0.085	± 0.124	± 0.0021
P2-RDF	0.176	0.000	0.176	0.283	0.240	0.225	0.083	± 0.125	± 0.0021
P2-CFM	0.165	0.002	0.165	0.264	0.226	0.212	0.079	± 0.121	± 0.0020
P3-MMSF	0.235	0.000	0.235	0.236	1.362	1.261	0.512	± 0.228	± 0.0035
P3-SMT	0.169	0.002	0.169	0.233	0.263	0.247	0.091	± 0.164	± 0.0025
P3-RDF	0.176	-0.001	0.176	0.272	0.251	0.234	0.091	± 0.164	± 0.0025
P3-CFM	0.165	0.003	0.165	0.230	0.233	0.218	0.083	± 0.160	± 0.0024

- regardless to the approach used to compute the permeability index and to the baseflow separation technique, a minimum value of about 0.4 for the BFI is expected for the permeability index lower limit; overall, the BFI appears to increase from about 0.4 to about 0.7 over the full range of variation for the permeability index;
- for a particular hydrograph separation technique, the empirical relationships between BFI and the permeability index, for each of the introduced permeability index, have a very similar intercept (A) and slope (B) parameters. As a consequence, regional BFI estimations for ungauged sites derived from regression prediction equations are very similar, except for the degree of dispersion around the regression line, which is very large when using P2 and P3 as the independent regional variables.
- A more accurate geological features' spatial variability description through higher resolution geological map enhances a finer delineation of permeable and impervious areas and the regression model performance, reducing the average long-term BFI index prediction error from 23% to 14%. Regional prediction equations are compared in Table 5 in terms of long-term BFI prediction errors (%), for each of the separation hydrograph technique and for each of the permeability index. The minimum prediction errors are highlighted: it is evident that P1 is more frequently, compared to P2 and P3, associated with the lowest percentage errors. Maximum prediction error is about 64% when using P1, 148% for P2 and 137% for P3, whereas on average it is about 14% for P1, 22% for P2 and 23% for P3.

Regional prediction equations are further compared in terms of more conventional statistics such as variance, bias, root mean square error RMSE besides the correlation coefficient R^2 (Table 6). The same table also provides jackknife errors estimates derived from the application of the jackknife resampling procedure, to assess different sources of error such as mean total true error, mean apparent error, a measure of goodness of fit, and mean expected excess error, a measure of model robustness (Efron, 1982). More measures of statistical error are given in terms of confidence intervals, estimated at 5% significance level, for the intercept (A_{conf}) and the slope (B_{conf}) parameters of the regional prediction equations.

The results of the analysis indicate that, on one hand, the choice of a particular technique for baseflow separation is not extremely relevant and that, on the other hand, more relevance to this study has the computation of the permeability index at the catchment scale, which has the potential to be a BFI indicator even though it is derived from an apparently over-simplified schematization. Among the twelve regional prediction equations, regression relationships using P1 as an independent variable have better performance compared to others. They are characterized by the larger explained variance and the larger goodness-of-fit (larger R^2 values and smaller ma errors), the smaller model parameters estimation uncertainty (narrower confidence intervals) and represent the more robust approaches (smaller ME errors values).

Conclusions

This paper has presented a regional regression approach to predict the BFI index at ungauged sites, based on the introduction of a permeability index P1, within a Mediterranean region and data scarcity context. Two criterion of catchment area partition have been compared to calculate the permeability index: on one hand, a comprehensive soil-vegetation-morphology classification, proposed in the past and successfully used for flood prediction in ungauged sites, and on the other hand, a simple permeable and impervious areas classification. Likely because of the differences between long-term water balance and short temporal scale events water balance, the latter, which is the simpler criterion, has a favourable performance: in the investigated area, where catchment geology has been found to be the most relevant variable for low flow index estimation, the explained variance is about 68%. To estimate the more suitable regional relationship, different algorithms for baseflow separation have been considered. Twelve linear regression equations resulted from the combinations of computed BFI (MMSF, SMT, RDF, CFM) and computed P (P1, P2, P3). They have been compared in terms of conventional statistics (variance, bias, root mean square error RMSE and correlation coefficient R^2) and of jackknife prediction errors.

The choice of a particular baseflow separation technique, excluding the empirical MMSF, which has resulted

to be the less stable procedure for the investigated area, seems not to be extremely relevant to the presented analysis. The well-known Smoothed and Minima Technique has been, however, found to be the most suitable, stable and reliable procedure for the studied region. More relevance has appeared to have the computation of the permeability index: although the estimation of P1 appears to be derived from an over-simplified schematization, it has been shown that a detailed spatial variability description of lithological and hydrogeological features reduces the average long-term BFI index prediction error from 23% to 14% and moreover increases the explained variance from 23% to 68%. It is also interesting to observe that, even though the use of P1 increases model performances, estimated model parameters for the twelve regression model equations are very similar, thus, on average, predicted BFI values at an ungauged site are very similar, regardless to the approach used to define P, but differing in the degree of dispersion around the regression line. This consideration reduces somehow the uncertainties in catchment permeability index definition in the sense that the major influence of catchment features affecting baseflow production has been captured by the simple introduced scheme.

As the consequent findings of the current analysis are certainly related to the investigated database, an extension of the database itself to contiguous regions has been planned. This would improve, refine and systematize the definition of the introduced catchment permeability index P1, taking advantage of the application to different environments and, moreover, this would increase the non-homogeneities in the data, including climatic features that may affect regional relationships estimation. The presented analyses have furthermore the role of a preparatory step toward a comprehensive framework of more low flow indexes estimation which has already been tested on a more limited region (Longobardi and Villani, 2007). In such a framework, measures of environmental minimum flow requirements, flow-duration and flow-frequency curves parameters are related to the BFI evaluation from regional catchments features.

Acknowledgement

The authors gratefully acknowledge funding support provided through the Instruction, University and Research Italian Ministry (MIUR) under the Grant PRIN 2005 2005080490_003. The authors would also thank the anonymous reviewers for their helpful comments, which resulted in a significant improvement of the manuscript.

References

- Boughton, W.C., 1993. A hydrograph-based model for estimating the water yield of ungauged catchments. In: Hydrology and Water Resources Symposium. Institution of Engineers Australia, Newcastle, pp. 317–324.
- Boorman, D.B., Hollis, J.M., Lilly, A., 1995. Hydrology of Soil Types: a Hydrologically-Based Classification of the Soils of the United Kingdom. Institute of Hydrology, Wallingford.
- Casadei, S., 1995. Analisi dell'indice del deflusso di base (BFI) di alcuni bacini dell'Italia Centrale. *Idrotecnica* 3, 177–191.
- Celico, P., De Innocentis, M., Rossi, F., Villani, P., 1989. Influenza dei parametri fisiografici del bacino sul coefficiente di afflusso di piena. GNDCI 1989 Report, pp. 189–208.
- Chapman, T., 1991. Comment on "Evaluation of automated techniques for baseflow and recession analyses" by R.J. Nathan and T.A. McMahon. *Water Resources Research* 27 (7), 1783–1784.
- Chapman, T., 1999. A comparison of algorithms for streamflow recession and baseflow separation. *Hydrological Processes* 13, 701–714.
- Chapman, T., Maxwell, A., 1996. Baseflow separation – comparison of numerical methods with tracer experiments. In: Hydrology and Water Resources Symposium, Institution of Engineers Australia, Hobart, pp. 539–545.
- Eckhardt, K., 2005. How to construct recursive digital filters for baseflow separation? *Hydrological Processes* 19, 507–515.
- Efron, B., 1982. The jackknife, the bootstrap and other resampling plans. Society for Industrial and Applied Mathematics, Philadelphia, USA, 92.
- Furey, P.R., Gupta, V.K., 2001. A physically based filter for separating baseflow from streamflow time series. *Water Resources Research* 37 (11), 2709–2722.
- Gabriele, S., 1998. Evolution of a regional methodology for flood evaluation in Italy. In: Andha, K. (Ed.), *Managing Hydro-Geological Disasters in Vulnerable Environment*. CNR-GNDCI No. 1900, Perugia.
- Gustard, A., Roald, L.A., Demuth, S., Lumadjeng, H.S., Gross, R., 1989. Flow regimes from experimental and network data. *Hydrological Studies*, vol. 1. Institute of Hydrology, Wallingford.
- Haberlandt, U., Klocking, B., Krysanova, V., Becker, A., 2001. Regionalisation of the base flow index from dynamically simulated flow components – a case study in the Elbe River Basin. *Journal of Hydrology* 248, 35–53.
- Hewlett, J.D., Hibbert, A.R., 1967. Factors affecting the response of small watershed to precipitation in humid areas. In: *International Symposium of Forest hydrology*. Pergamon press, pp. 275–290.
- Institute of Hydrology, 1980. *Low Flow Studies*, vols. 1–4. Wallingford, UK.
- Lacey, C.G., Grayson, R.B., 1998. Relating baseflow to catchment properties in south-eastern Australia. *Journal of Hydrology* 204, 231–250.
- Longobardi, A., Villani, P., 2007. Low flows regional statistical analysis within a Southern Italy context. *Geophysical Research Abstracts* 9, 08720.
- Lyne, V.D., Hollick, M., 1979. Stochastic time-variable rainfall runoff modelling. In: *Hydrology and Water Resources Symposium*. Institution of Engineers Australia, Perth, pp. 89–92.
- Murrone, F., Rossi, F., Claps, P., 1997. Conceptually-based Shot Noise modelling of streamflows at short time interval. *Stochastic Hydrology and Hydraulics* 11 (6), 483–510.
- Mwakalila, S., Feyen, J., Wyseure, G., 2002. The influence of physical catchment properties on baseflow in semi-arid environments. *Journal of Arid Environments* 52, 245–258.
- Nathan, R.J., McMahon, T.A., 1990. Evaluation of automated techniques for baseflow and recession analyses. *Water Resources Research* 26 (7), 1465–1473.
- Nathan, R.J., Austin, K., Crawford, D., Jayasuriya, N., 1996. The estimation of monthly yield in ungauged catchments using a lumped conceptual model. *Australian Journal of Water Resources* 1 (2), 65–75.
- Ponce, V.M., Shetty, A.V., 1995a. A conceptual model of catchment water balance. 1. Formulation and calibration. *Journal of Hydrology* 173, 27–40.

- Ponce, V.M., Shetty, A.V., 1995b. A conceptual model of catchment water balance. 2. Application to runoff and baseflow modeling. *Journal of Hydrology* 173, 41–50.
- Rossi, F., Villani, P., 1994. Regional flood estimation methods. In: Rossi, G., Harmancioglu, N., Yevjevich, V. (Eds.), *Coping with Floods*. Kluwer Academic Publisher, pp. 93–217.
- Schneider, M.K., Brunner, F., Hollis, J.M., Stamm, C., 2007. Towards a hydrological classification of European soils: preliminary test of its predictive power for the base flow index using river discharge data. *Hydrology and Earth System Science* 11, 1501–1513.
- Spongberg, M.E., 2000. Spectral analysis of baseflow separation with digital filters. *Water Resources Research* 36 (3), 745–752.
- Vogel, R.M., Kroll, C.N., 1990. Generalized low-flow frequency relationships for ungauged sites in Massachusetts. *Water Resources Bulletin* 26 (2), 241–253.
- Vogel, R.M., Kroll, C.N., 1992. Regional geohydrologic–geomorphic relationships for the estimation of low-flow statistics. *Water Resources Research* 28 (9), 2451–2458.
- Wundt, W., 1958. Die Kleinstwasserführung der Flüsse als Maß für die verfügbaren Grundwassermengen. *Forschungen Zu deutschen Landeskunde* 104, 47–54.