

# Regionalization of Flow-Duration Curves through Catchment Classification with Streamflow Signatures and Physiographic–Climate Indices

Laura Boscarello<sup>1</sup>; Giovanni Ravazzani<sup>2</sup>; Alessio Cislighi<sup>3</sup>; and Marco Mancini<sup>4</sup>

**Abstract:** This study addresses the estimation of flow-duration curves (FDC) in ungauged sites through the catchment classification. Forty-six catchments in the Upper Po river basin (Italy) were analyzed and classified through two different frameworks: the first scheme consists of the application of two clustering methods in a series considering six streamflow signatures, and the second one treats indexes of climate, physiography, soil, and land-use with the same clustering procedure. Catchments have been classified into three homogeneous groups: the first one is characterized by the lowest runoff and flash-flood events, the second one includes maximum runoff, and the third one shows intermediate behaviour. The estimation of FDCs was done using a lognormal distribution, whereas the regionalization was constructed applying a stepwise multiple linear regression, followed by a leave-one-out cross-validation. The results show great performance improvement when the regionalization model is found by taking account of the three different hydrological classes, with a mean absolute percentage error that decreases from 11% for the single region case to 7% in the three homogeneous regions case.

## Introduction

Flow-duration curves (FDCs) are the most popular tools used by hydrologists in various water-related applications because they clearly provide concise and detailed information about river streamflow variability. They represent the percentage of time that streamflow is equalled or exceeded for a given catchment (Mendicino and Senatore 2013).

The lack of gauged sites and the limited availability of streamflow observations creates a relevant problem for estimation of FDCs that is usually solved through the application of regionalization models (Castellarin et al. 2004; Ganora et al. 2009). Regionalization requires understanding and knowledge of hydrological processes and spatio-temporal heterogeneity of climatic and landscapes characteristics (Li et al. 2010; Cheng et al. 2012; Ye et al. 2012). A simple nearest-neighbour approach to define similarity among catchments is not enough in some cases because it is unable to work effectively if there are nearby gauged catchments showing different intrinsic characteristics (Boscarello et al. 2013).

Hydrologic similarity between catchments is the basis for catchment classification. The fundamental idea is to simplify

hydrologic systems in groups and subgroups and to recognize salient characteristics that are emblematic to develop models. During the past decades, many studies emphasized the necessity of simplification in hydrological modelling, and in this context, some attempts were made also in developing a catchment classification framework. The two commonly-suggested strategies have focused on spatial proximity (e.g., Oudin et al. 2008; Patil and Stieglitz 2014) and physical proximity using different algorithms to link hydrological responses to physical and climatic attributes (e.g., Coopersmith et al. 2012; Ssegane et al. 2012; Razavi and Coulibaly 2013). Some previous studies showed that the spatial or geographical proximity does not always translate into hydrological similarity (e.g., Acreman and Sinclair 1986; Boscarello et al. 2013). Thus, they considered that the physical (climate and landscape) properties are the dominant controls on the hydrologic behaviour of a catchment and therefore sufficient to group together catchments that are hydrologically similar (Sawicz et al. 2011). However, other studies (Castellarin et al. 2011; Di Prinzio et al. 2011) showed that an explicit quantitative assessment of such relationships (a mapping) is required and has to be tested to fully allow information transfer and to achieve a generalization of the relationship between catchments attributes, climate, and hydrologic behaviour.

Different procedures were proposed in literature within the topic of catchment classification, but the large majority of regionalization studies used clustering techniques (e.g., Yadav et al. 2007; Razavi and Coulibaly 2013). The idea of cluster analysis is to identify clusters in such a way that catchments' similarity within a group is maximized, while similarities between different groups are minimized. There are several types of clustering techniques, such as exclusive, overlapping, hierarchical, probabilistic, and artificial neural-network-based ones. One of the most promising methods is SOM (self organizing map), a popular and widespread artificial neural-network model (Vesanto and Alhoniemi 2000; Ley et al. 2011; Toth 2013).

The main objective of the present paper is to classify catchments within the Upper Po river basin, in Italy, and to demonstrate how a good catchment classification framework can help in

<sup>1</sup>Doctoral Researcher, Dept. of Civil and Environmental Engineering, Politecnico di Milano, 20133 Milan, Italy. E-mail: lauraanna.boscarello@polimi.it

<sup>2</sup>Assistant Professor, Dept. of Civil and Environmental Engineering, Politecnico di Milano, 20133 Milan, Italy (corresponding author). E-mail: giovanni.ravazzani@polimi.it

<sup>3</sup>Student, Politecnico di Milano, 20133 Milan, Italy. E-mail: alessio.cislighi@mail.polimi.it

<sup>4</sup>Professor, Dept. of Civil and Environmental Engineering, Politecnico di Milano, 20133 Milan, Italy. E-mail: marco.mancini@polimi.it

Note. This manuscript was submitted on October 20, 2014; approved on August 21, 2015; discussions must be submitted for individual papers.

defining regionalization model for FDCs. Some regionalization studies are available for this area. Viglione et al. (2006) presented a regional analysis of the mean annual flow, Vezza et al. (2010) applied a regional regression approach to estimate low flow in ungauged catchments and Ganora et al. (2009) presented an approach to estimate dimensionless flow duration curves including 95 basins located in northwestern Italy and Switzerland.

This paper provides a complete catchment classification that separately includes representative and non-correlated streamflow indexes and climate, physiographic, land-use, and soil characteristics. In fact, climate information are generally supposed to be predominant in describing streamflow characteristics, but as confirmed by Winter (2001) and Yadav et al. (2007), physiographic and land-use information are also relevant descriptors of catchment behaviour. The purpose of this paper is to produce two classifications using the same clustering technique with the same reliability. Furthermore, this paper proposes a validation procedure for the classification based on climatic and physiographic catchment descriptors in contrast to a classification based on streamflow regime. Toth (2013) used the leave-one-out cross-validation, however this method is suggested when the training size approaches the samples (i.e.,  $n - 1$  tends to  $n$ ) (Sylvain and Celisse 2010). The validation procedure applied in the present paper is based on the classic technique that is typically used for the validation of hydrological models: an independent set of catchments not used in the clustering training are supposed as ungauged and then used to validate the robustness of the classification scheme.

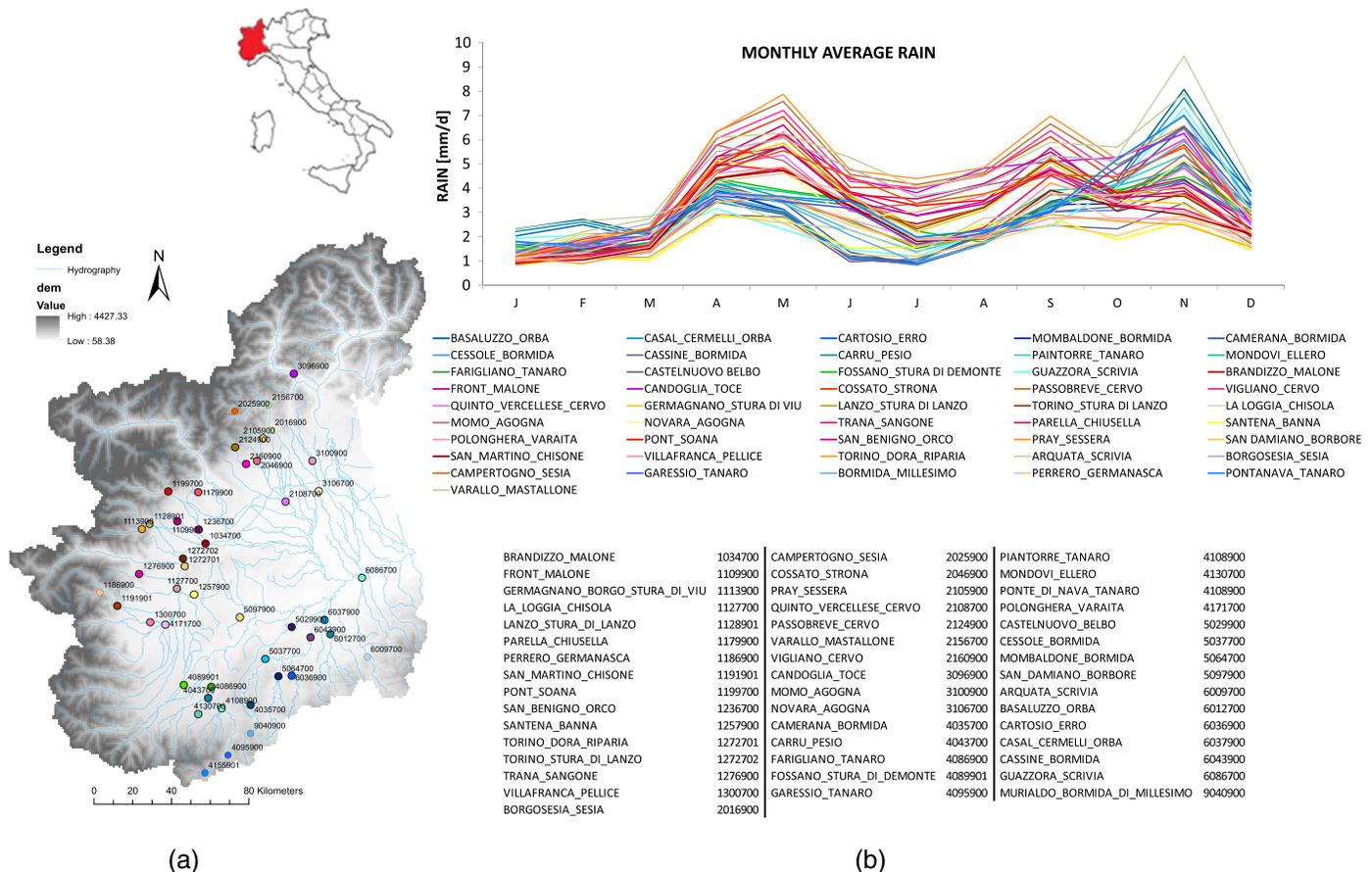
## Study Area and Data

This study considers a set of 46 small-sized to medium-sized gauged catchments in Upper Po river basin, in northwest Italy, as shown in Fig. 1(a). The catchments of interest include several rivers, which are tributaries of the Po river. Among the 46 catchments are 28 upstream catchments, 11 of which are double, triple, or quadruple nested.

The area of the basins ranges from 56 to 1,588 km<sup>2</sup> with six catchments with area greater than 1,000 km<sup>2</sup>. The minimum catchment elevation ranges from about 75 to 1,440 m. Almost all catchments are rural with little urbanization, with 65% of the total area wooded.

Mean annual precipitation ranges from 680 to 1,700 mm. The catchments in the southeast part of the studied region have the highest precipitation rates in November, while the other basins show heavy precipitation in May and September [Fig. 1(b)]. Average annual temperature is 8°C with a minimum in January and a maximum in August. Soil conductivity is low in the highlands and increases in the valleys where sandy river deposits are present.

For the classification, hourly runoff, precipitation, temperature, and radiation data were available from 2000 to 2009; these time series data cover a wide range of diverse annual and seasonal precipitation and runoff events, e.g., the extreme flood event of October 2000 and the drought of summer 2003. Precipitation data were interpolated with an inverse distance technique starting from 486 stations on the whole Upper Po region. Further available data



include digital maps including a digital elevation model, land cover maps, and soil texture maps with 1-km spatial resolution.

## Methods

### Streamflow Signatures

The first step of this work consists of choosing hydrological signatures that provide information about the catchments' functional behaviour and that are suitable to cluster catchments based on hydrologic responses. The signatures are derived from the available time series of streamflow and precipitation, choosing from among the wide choice of indexes reported in literature (Yadav et al. 2007; Clausen and Biggs 2000; Olden and Poff 2003; Sawicz et al. 2011). The chosen indexes are (1) median of daily flows, (2) annual runoff coefficient, (3) (4) slope of the flow duration curve, (5) streamflow elasticity, (6) baseflow index, and (7) Q10 variable. These indexes try to represent all parts of the hydrograph. In particular, the median of daily flows and the runoff coefficient represent the magnitude of average flow events. The Q10 variable represents the magnitude of high flow events, the baseflow index represents the magnitude of low flow events, the slope of the flow duration curve represents the rate of change in flow events, and the streamflow elasticity represents the sensitivity of catchment streamflow response with

respect to changes in precipitation. The six indexes used are reported in Table 1.

### Morpho-Climatic Catchment Properties

The choice of the most representative morpho-climatic catchment characteristics is difficult and, in most of cases, is based on empirical remarks (Sawicz et al. 2011; Yadav et al. 2007; Sanborn and Bledsoe 2006). The first step has been the calculation of many physical catchment properties proposed by the literature (Sawicz et al. 2011; Yadav et al. 2007; Sanborn and Bledsoe 2006; Ley et al. 2011; Toth 2013). It is important to choose a limited set of variables for cluster analysis because redundant variables can collectively bias the solution. To reduce the number of indexes, the ones with the highest values of Pearson's correlation coefficient with the hydrological signatures were selected, and the ones with the smallest variability within each hydrological class as determined after classification with hydrological signatures were discarded. The 15 selected indexes are reported in Table 2.

The indexes were divided in four categories: (1) climate, (2) land use, (3) physiographic characteristics, and (4) soil characteristics. The climate features were derived from meteorological data that include precipitation, temperature, and potential evapotranspiration. Potential evapotranspiration is calculated with the Priestley

**Table 1.** Streamflow Signatures Used for Classification

Metric	Description	Equation	Unit	Range
$Q_{50/A}$	Median of daily flows	$Q_{50}/A$	$m^3/s/km^2$	0.0008–0.0298
$R_{QP}$	Runoff coefficient	$Q/P$	—	0.131–1.023
$Q_{10/50}$	Annual mean of the flow exceeded 10% of the time	$Q_{10}/Q_{50}$	—	2.2–28.6
BFI	Baseflow index	$Q_B/Q$	—	0.592–0.881
$S_{FDC}$	Slope of flow duration curve	$S_{FDC} = \ln(Q_{33\%}) - \ln(Q_{66\%})/0.66 - 0.33$	—	1.3–7.1
$E_{QP}$	Streamflow elasticity	$E_{QP} = \text{median}(dQ/dP \cdot P/Q)$	—	0.62–16.66

Note:  $Q_{50}$  = 50th percentile of the total flow duration curve ( $m^3/s$ );  $A$  = basin area ( $km^2$ );  $Q$  = average streamflow ( $mm$ );  $P$  = average annual precipitation ( $mm$ );  $Q_{10}$  = flow exceeded 10% of the time ( $m^3/s$ );  $Q_{50}$  = annual median flow ( $m^3/s$ );  $Q_B$  = average baseflow ( $mm$ );  $Q_{33\%}$  and  $Q_{66\%}$  = 33rd and the 66th streamflow percentiles ( $m^3/s$ ), respectively;  $dQ$  = difference between the previous year's streamflow and the current year's streamflow ( $mm$ );  $dP$  = same for precipitation ( $mm$ ).

**Table 2.** Morpho-Climatic Catchment Properties Used for Classification

Metric	Description	Units	Range
Climate			
CR	Ratio of average annual Precipitation to Evapotranspiration	—	1.10–3.42
RWD	Ratio of precipitation in wettest month to that of the driest	—	19.9–164
MMA	Median annual maximum 1-h precipitation	$mm/h$	7.2–51.3
T-mo	Average monthly temperature	$^{\circ}C$	3.5–12.5
MAP	Mean cumulated annual precipitation	$mm$	685–1,756
NP	Number of days with precipitation	$days/year$	125–265
D.std	Standard deviation of annual precipitation	$mm$	227–680
Land use			
Agr	Percentage of arable land within the watershed	—	0.02–0.87
Wood	Percentage of woodland within the watershed	—	0.07–0.98
Physiographic			
Slope	Watershed steepness	$^{\circ}$	0.9–17
H.5	5 percentile elevation	$m$	104–1,017
H.95	95 percentile elevation	$m$	335–3,006
Soil			
High.p	Percentage of soil with high permeability within the watershed	—	0.02–0.84
Low.p	Percentage of soil with low permeability within the watershed	—	0–0.625
FC	Watershed mean field capacity	—	0.16–0.76

Taylor equation (Priestley and Taylor 1972), which needs only temperature and solar radiation data as inputs.

Land use characteristics were calculated from Corine Land Cover Program (<http://www.eea.europa.eu/publications/COR0-landcover>). Physiographic properties were derived from digital elevation model (DEM), and soil characteristics are calculated from the soil texture map. More than one-third of indexes are climatic because they exert a major control on runoff generation (Yadav et al. 2007; Ley et al. 2011; Sawicz et al. 2011). The other catchment characteristics that contained low values of Pearson's correlation coefficient, such as area and flow path length, were discarded because they have no influence on the chosen stream-flow indexes (Ley et al. 2011; Yadav et al. 2007).

### Self Organizing Map (SOM)

A SOM is an unsupervised learning algorithm based on artificial neural networks to produce a low-dimensional representation of a high-dimensional input data set. A SOM consists of centroids called neurons, which are organized on a regular grid. Each neuron is represented by a prototype vector with the same dimension as the input vector. The prototype vector with the lowest quantization error, which is the Euclidean distance between input vector and prototype vectors, is called best-matching unit (BMU) and serves as the basis for assigning input vectors (catchments) to neurons. There are only few parameters to set before training a SOM: (1) the distance measure (in this case Euclidean), (2) a neighbourhood function (here the common Gaussian function is used), and (3) the normalization of the input vector so that each index is ensured of having the same importance independent of its values (in this case the logistic normalization was chosen so that all values range from 0 to 1).

According to the sequential approach, an input vector is randomly picked and the Euclidean distance between the input vector and each of the neurons is computed to find the BMU. The prototypes of the winning node and of its neighbouring nodes change to further reduce the distance between the prototype and the input vector. These steps are repeated until a fixed number of iterations is reached. As an alternative to the sequential training approach, the batch-training algorithm *SOM-Toolbox for Matlab 5* was used (Vesanto et al. 2000; Herbst and Casper 2008). Instead of using one data vector at a time, the whole data set is presented to the map before any adjustments are made. This approach speeds up the training process, does not need a learning-rate factor, and makes the SOM independent from a random choice of dataset at the start of the training (Ley et al. 2011).

Concerning the dimensions of the map, Vesanto and Alhoniemi (2000) report that the ideal number of neurons is  $5\sqrt{n}$ , where  $n$  is the number of input vectors. However, no definitive method exists to choose the appropriate number of partitions in a data set. In order to guarantee the correct analysis, it must be ensured that the mapping has been correctly chosen. Hence two measures were used in this study to quantify the goodness of the map: topographic error and quantization error. Topographic error is the most simple of the topology-preservation measures. It is computed as follows. For all data samples, the respective best and second-best matching units are determined. If these are not adjacent on the map lattice, this is considered an error. The total error is then normalized to a range from 0 to 1, where 0 means perfect topology preservation. The quantization error is computed by determining the average distance of the sample vectors to the cluster centroids by which they are represented. In case of the SOM, the cluster centroids are the prototype vectors. The optimal map dimension was set so as to minimize both the quantization error and the topologic error.

### Hierarchical Clustering

In this study, the inner squared distance Ward's method was used because it was identified by some authors as the method that outperforms other algorithms in terms of separation to give relatively dense clusters with small within-group variance (Hannah et al. 2005; Köplin et al. 2012). A tree-like diagram, called a dendrogram, usually represents the results of the hierarchical clustering. The lengths of the dendrogram's limbs denote the proximity of points, so data can be clustered by cutting the dendrogram at a desired level.

Principal component analysis was considered in a first step of the work but although the computational efficiency is a strong argument in favor of model-free methods, these algorithms tend to select very large subsets of input variables with a high risk of redundancy, as already noted by Galelli and Castelletti (2013).

### Catchment Classification Procedure

A two-level approach was applied, first considering only the six hydrological signatures, and then considering the 15 physical indexes. Both classifications were made following the same scheme. First SOM was applied by choosing the most appropriate dimension of the grid. Catchments, which are grouped together on the same neuron, are considered to behave similarly and build the smallest unit of clustering. In the second level, hierarchical clustering refines the results of SOM considering all neurons on the SOM's map, and the final clusters are determined.

The two-level approach is considered to perform better than a single-level one because of computational cost and noise reduction (Vesanto and Alhoniemi 2000). The choice in this study tried to emphasize the advantages and reduce the limitations of both methods. In particular, as pointed out by Ley et al. (2011), SOM is numerical, nonparametric, and insensitive against a small portion of missing data and is thus a perfect tool to represent graded relationships and provide visualizations of structure in high dimensional data sets as it does not need any assumptions about data distribution. Meanwhile, the hierarchical clustering is affected by problems such as sensitivity to noise and outliers, difficulty handling different sized clusters, irregular shapes, and tendency to break large clusters (Tan et al. 2006).

There is not a rigorous method to define the most-appropriate number of classes in an unsupervised clustering, and in most cases, it depends on the cluster resolution desired by the user. In this study, as suggested in Sanborn and Bledsoe (2006), the interpretability of results was used as the primary decision factor. Considering the number of input data and the geographic, climatic, and morphologic differences among the catchments in the study area, three classes are sufficient to represent variability: two classes have distinctly opposite characteristics, and the other class has intermediate properties. If the number of classes is too large, it can cause confusion in interpreting the results.

In order to assess overlap between the two classifications, the similarity index (SI) described in Ssegane et al. (2012) and Razavi and Coulibaly (2013) was computed

$$SI(\gamma, \theta) = 1/2[1 - (|\gamma \setminus \theta| + |\theta \setminus \gamma| - 2|\gamma \cap \theta|)/(|\gamma| + |\theta|)] \quad (1)$$

where set  $\gamma$  = one of the homogeneous reference clusters classified using streamflow signatures, whereas set  $\theta$  = one of the clusters classified using physical indexes.

In order to validate the classification based on climatic and physiographic catchment descriptors against a classification based on streamflow regime, 8 of the 46 catchments were excluded from the classification analysis. The choice of the catchments used for

validation was guided by the principle put forward in Wallner et al. (2013), when the distributions of the hydrological indexes are similar between calibration and validation, the validation improves its reliability.

### Regionalization of Flow-Duration Curves (FDCs)

There are several methods to regionalize and to estimate FDCs at ungauged sites (Castellarin et al. 2004; Booker and Snelder 2012; Ganora et al. 2009). Many regionalization techniques were applied to Italian basins (Claps and Fiorentino 1997; Castellarin et al. 2004, 2007; Ganora et al. 2009; Viola et al. 2011; Mendicino and Senatore 2013). In this paper, the presented method is a statistical approach with lognormal distribution as developed by Fennessey and Vogel (1990). The regionalization model was found out first for all the 46 catchments together and, second, using the three hydrological classes found by classification, in order to build a specific regional model for each class.

In this study, the FDCs were constructed by considering the entire data set of observations because this allows the streamflow series to be extended, which solves the problem of gap filling (Castellarin et al. 2004, 2007).

In the following step, a two-parameter lognormal frequency distribution was adopted to represent the distribution of daily streamflow (Fennessey and Vogel 1990):

$$p = P(Z > z_p) = 1 - (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{z_p} \exp\left(-\frac{1}{2}x^2\right) dx \quad (2a)$$

where

$$z_p = \frac{\ln(q_p) - \mu}{\sigma} \quad (2b)$$

is the  $p$ th percentile of the random variable  $Z$ , which is normally distributed with zero mean and unit variance;  $q_p$  =  $p$ th percentile of the log-normally distributed random variable with parameters  $\mu$  and  $\sigma$ .

The optimal parameters  $\hat{\mu}$  and  $\hat{\sigma}$  were estimated by minimizing the sum of the squared residuals between the exceedance probabilities of the observed daily streamflows and the probabilities estimated from Eq. (2a) within the range  $0.33 \leq p \leq 0.99$

$$\sum_{p_i=0.3}^{0.99} [p_i - F_i(z)]^2 = \min \quad (3)$$

where  $p_i$  = Weibull Probability; and  $F_i(z)$  = cumulative density function of the standardized normal distribution.

The goodness-of-fit between the empirical and estimated FDCs was quantified by calculating the Nash-Sutcliffe index criterion (Nash and Sutcliffe 1970), as suggested by Castellarin et al. (2004). The mean of the resulting Nash-Sutcliffe indexes is equal to 0.984, with a standard deviation of 0.015.

To predict regional FDCs in ungauged sites, the development of two regional predictive models for  $\hat{\mu}$  and  $\hat{\sigma}$  is required. A multiple linear regression was performed to develop a model of the form

$$\hat{\theta} = a_0 + a_1 \ln(X_1) + a_2 \ln(X_2) + \dots + a_n \ln(X_n) + \vartheta \quad (4)$$

where  $\hat{\theta}$  = variable to be estimated;  $\hat{\mu}$  or  $\hat{\sigma}$  and  $X_i$  = explicative variables of the models (climatic, topographic indexes);  $a_i$  = parameters of the model; and  $\vartheta$  = model residual.

The optimal subset of explanatory variables was determined using a multiple linear regression and least-squares algorithm. Starting from a constant model, a further explanatory variable

was added to the model at each step. Then it is necessary to choose the set of variables that maximize the adjusted R-square statistic

$$R_{adj}^2 = 1 - \frac{\sum_i (y_{i,obs} - y_{i,mod})^2 / (n - np)}{\sum_i (y_{i,obs} - \bar{y}_{obs})^2 / (n - 1)} \quad (5)$$

where  $y_{i,obs}$  = observed value;  $\bar{y}_{obs}$  = mean of observed value;  $y_{i,mod}$  = value estimated from regression;  $n$  = number of data; and  $np$  = number of parameters used in the model.

In order to test reliability and robustness of the regional model, a cross-validation was performed according to a leave-one-out procedure. A jack-knife cross-validation procedure was used (Efron 1982). This approach consists in leaving out all data associated with each gauging station and then estimating the FDCs for the left-out gauging station using data from all remaining gauging stations (Brath et al. 2001; Castellarin et al. 2004, 2007; Booker and Snelder 2012).

The regression model's structure remains the same at each step, but the coefficients  $a_i$  are modified. In this way, it is possible to compare performances achieved with parameters derived from the regional model and from the leave-one-out cross-validation, assuming that each station is ungauged in turn, and thus evaluating the reliability of the model.

The  $N$  empirical FDCs were compared to the ones retrieved from the regional model by calculating an index of reliability

$$\varepsilon_{s,j} = \frac{\hat{q}_{s,j} - q_{s,j}}{q_{s,j}} \quad (6a)$$

where  $j$  = duration;  $s$  = station;  $\hat{q}_{s,j}$  = cross-validated estimate; and  $q_{s,j}$  = empirical value. Then the mean and the standard deviation of  $\varepsilon_{s,j}$  are calculated

$$\bar{\varepsilon}_s = \frac{1}{N} \sum_{j=1}^N \varepsilon_{s,j} \quad (6b)$$

$$\sigma_{\varepsilon,s} = \sqrt{\frac{1}{N-1} \sum_{j=1}^N (\varepsilon_{s,j} - \bar{\varepsilon}_s)^2} \quad (6c)$$

where  $N$  = number of  $j$ 's duration. The mean and the standard deviation of  $\bar{\varepsilon}_s$  were calculated and used as the reliability indexes.

For evaluating the performance of model, four indexes were computed: mean absolute error (MAE), mean absolute percentage error (MAPE), average deviation (AD), and efficiency index ( $E_s$ ), which are presented in the following equations:

$$\text{MAE} = \frac{\sum_{j=1}^N |\hat{q}_{s,j} - q_{s,j}|}{N} \quad (7)$$

$$\text{MAPE} = \frac{\text{MAE}}{\bar{q}_s} \quad (8)$$

$$\text{AD} = \frac{1}{n} \sum_{j=1}^n \frac{\hat{q}_{s,j} - q_{s,j}}{q_{s,j}} \quad (9)$$

$$E_s = 1 - \frac{\sum_{j=1}^N (\hat{q}_{s,j} - q_{s,j})^2}{\sum_{j=1}^N (q_{s,j} - \bar{q}_{s,j})^2} \quad (10)$$

where  $\hat{q}_{s,j}$ ,  $q_{s,j}$ , and  $\bar{q}_{s,j}$  = respectively, the estimated, observed streamflow, and mean of observed streamflow for site  $s$  and for duration  $j = 1, \dots, N$ .

## Results and Discussion

### Clustering with Streamflow Signatures

SOM and hierarchical clustering were applied to the six streamflow attributes identified for each of the 38 catchments. The SOM output is represented in Fig. 2(a). The map is composed of a  $8 \times 4$  grid for a total of 32 neurons; 12 of the 32 neurons are empty, meaning that SOM has therefore created 20 clusters. The distribution of the values of the six indexes is represented in Fig. 2(b) together with the U-matrix, which represents the distance between neighbouring neurons. The results of the hierarchical clustering application on SOM are shown in Fig. 3. The SOM map [Fig. 3(a)] is subdivided in the three classes identified by the dendrogram tree [Fig. 3(b)].

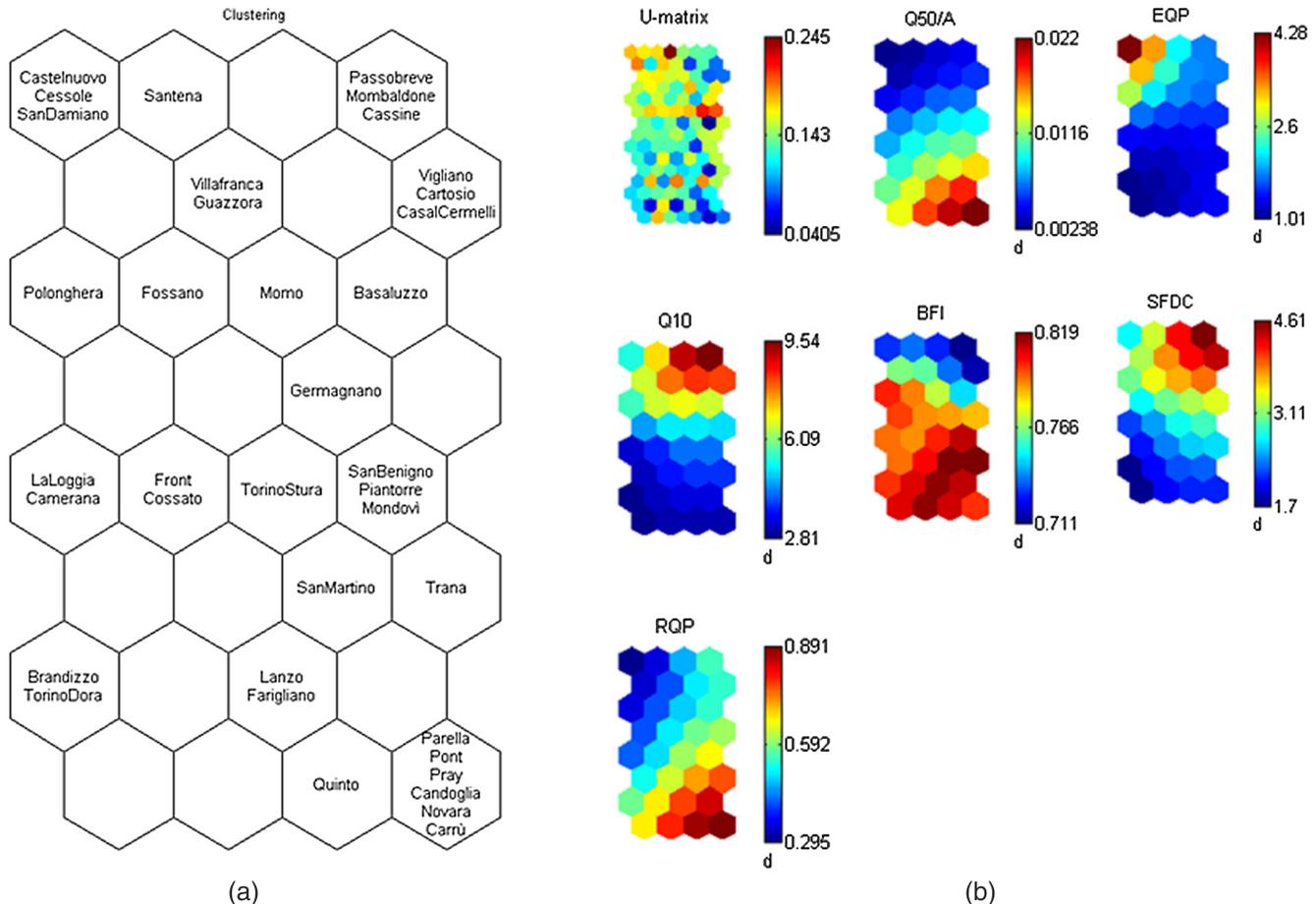
Results regarding the distribution of signature values in each cluster are shown in Fig. 4 as histograms of the centroids of each class. For each of the signatures, the normalized median value for each class (C1, C2, and C3) is shown. Class C1 is the one with the lowest values for runoff coefficient, median flow, and baseflow index but shows the highest values of both slope of the flow duration curve and Q10, so it corresponds to the class with the lowest runoff and characterized by flash-flood events. Class C3 is characterized by the highest values of runoff coefficient, baseflow, and median flow, while the other three signatures display the lowest values, meaning that this is the class with the maximum runoff. Class C2 shows intermediate behaviour for all signatures. Fig. 5 represents locations on the map of the 38 catchments and their membership in classes. It is evident that there is not a strong subdivision of

catchments' classification from a geographic point of view, meaning that not only climate influences runoff regime. However, precipitation exerts an influence on the classification. For example, most of the catchments in Class C1 are located in the south of the region that is the zone with the lowest precipitation rate, in agreement with the lowest value of median flow. On the contrary, catchments belonging to Class C3 are mostly located in the northwest part of the region that has the highest values for the precipitation rate, as the high value of median flow attests. Classes C1 and C3 appear therefore to reflect the dissimilarity of Apennine and Alpine regimes, respectively.

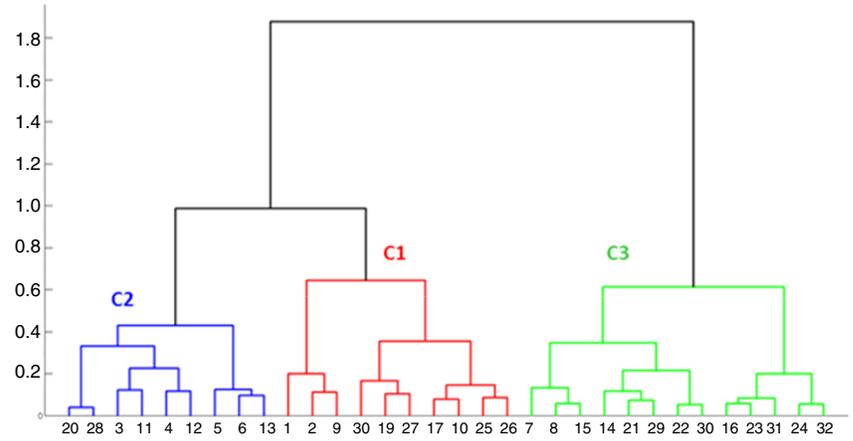
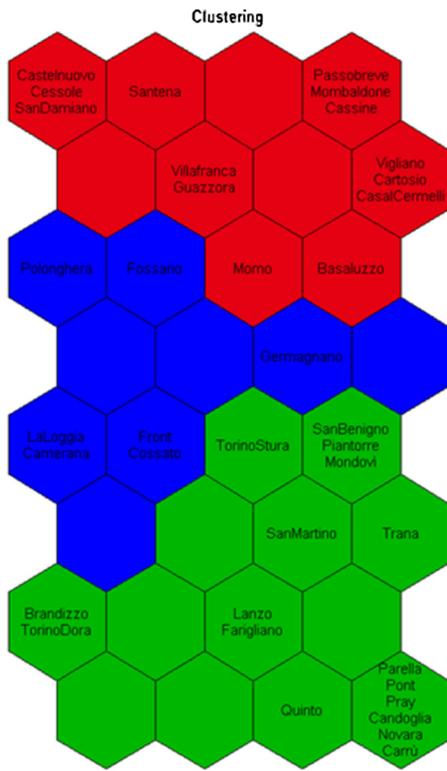
### Clustering with Physical Properties

An important feature of a cluster analysis is the ability to discriminate between different groups, not only based on streamflow signatures, but also on physical and climate characteristics, so to find corresponding clusters between one classification and the other.

In this regard, a new classification of the 38 catchments was carried out, based on the 15 physiographic and climatic indexes. As already done with streamflow signatures, first the SOM algorithm was applied by finding an optimal grid of  $5 \times 2$  neurons. Fig. 6 shows results of SOM application. A first grouping of basins is visible in the SOM map of 10 neurons, even if the U-matrix shown in Fig. 6(b) highlights strong distances between the upper and lower parts of the map. Indeed, by applying hierarchical clustering on the SOM map, it is possible to find three classes, identified with A, B, and C, as shown in Fig. 7.



**Fig. 2.** (a) SOM map with neuron labels obtained for streamflow signatures classification; (b) component planes for each index and U-matrix



(a)

(b)

**Fig. 3.** (a) SOM map with neurons divided in the three classes based on streamflow signatures; (b) dendrogram on SOM; numbers on the dendrogram correspond to each neuron of the SOM map

As for the hydrologic classification, three classes are still clearly visible for classification with physiographic and climatic characteristics. Classes A and C represent the two extremes, while Class B has an intermediate value for each index, as shown in Fig. 8, where each histogram represents the mean value of each index for each of the three classes.

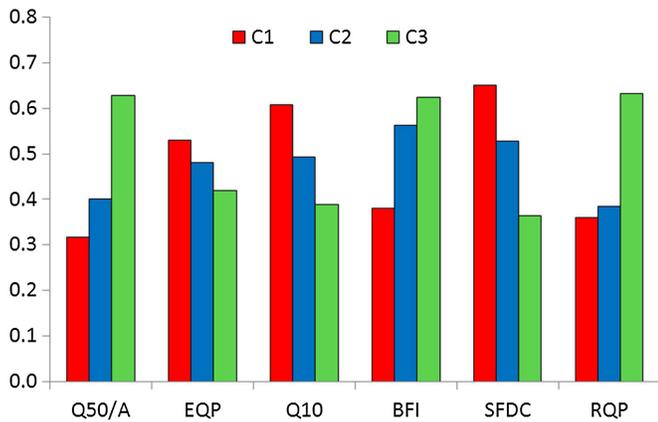
In terms of climatic conditions, Class A is characterized by the lowest values of CR, MAP, D.std and NP while, conversely, it has the maximum values of MMA, RWD and T-mo; it represents those catchments in which rainfall is not uniform during the year but is concentrated in short and heavy precipitation events, as indicated

by the high values of MMA and RWD. The average temperature is quite high. From the physiographic point of view, Class A includes rather flat basins mostly comprised of arable land, and the average altitude is lower in respect to that of the other catchments.

Class C is the opposite of Class A. In fact, it includes catchments with the highest values of CR, MAP, and NP, and the minimum values of MMA and RWD, showing the presence of more-abundant and uniform rainfall throughout the year. The average temperature is generally low, and from the physiographic perspective, Class C represents mountainous catchments with most of their area occupied by woodlands.

Class B collects all catchments with intermediate characteristics, even if it few basins belong to this class.

In Fig. 9, the location of the 38 catchments and their class membership is shown. Again, in this classification scheme, a subdivision between Alpine catchments (Class C) and Apennine catchments (Class A) can be seen.



**Fig. 4.** Median of the six hydrological indexes for each class of streamflow classification

### Comparison between Classifications

Starting from the two classification schemes, a contingency table was built (Table 3) to count the number of catchments with a certain hydrologic behavior that matches a certain class of physiographic characteristics. Correspondence between classes is evident: Cluster C1 is associated with Cluster A, Cluster C3 is associated with Cluster C, and Cluster C2 is associated with Cluster B. The values on the diagonal indicate the number of catchments that belong to the same cluster, while off-diagonal values denote the misclassified catchments. The overlap between the two schemes is 68.5%.

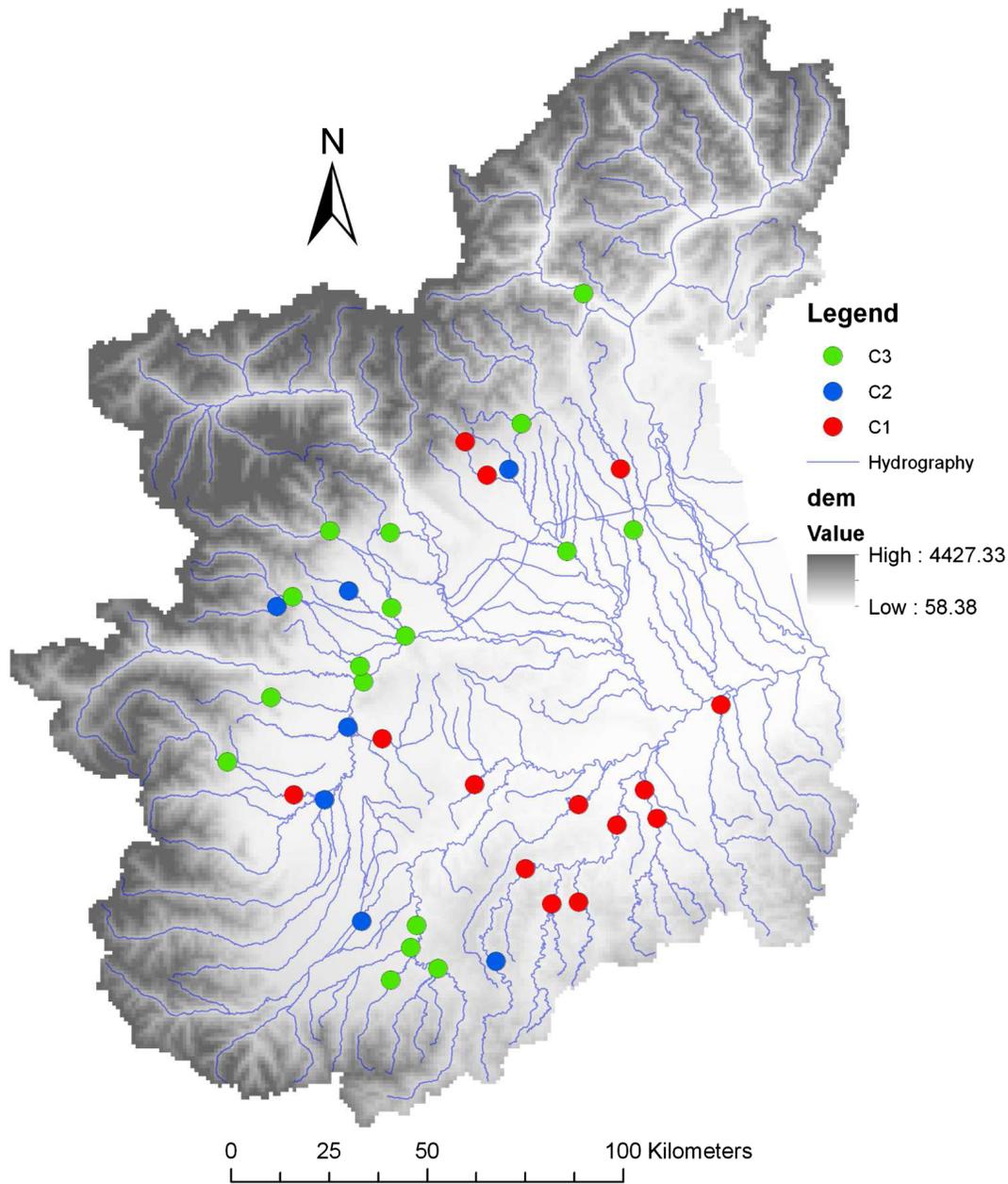


Fig. 5. Classification of catchments based on streamflow signatures

Table 4 presents the SI values between the two largest groups inside of two classifications. Comparing the results of the work developed by Ssegane et al. (2012), the SI values are slightly lower but very satisfactory.

One possible disadvantage of using such classification methods is the wrong allocation of a catchment. The cluster analysis does not create sharper borders and, in unlikely cases, it allocates similar catchments into different clusters. To overcome this problem, Sawicz et al. (2011) and Hall and Minns (1999) used other methods of clustering that can take into account borderline classes and are able to define the membership of one element to a class in terms of probability. In this case, the question is addressed by considering the topological correctness of the SOM and the quantization errors of each element. When creating the SOM map, one of the requirements was that the topologic error must be null in order to ensure topological correctness; examining the quantization error reveals that some catchments have similar quantization

error to different neurons, so that they can be recognized as belonging to two different clusters. With this analysis, the overlapping percentage between the two classifications increases until reaching 74%. Less than 30% of data is misclassified, meaning that some catchments with similar behavior have different physical characteristics. This fact is generally described as process equifinality (Ley et al. 2011; Hellebrand et al. 2011) in the runoff signal: different processes produce the same values of discharge characteristics in different basins. For example, Momo basin fits Classes C1 and A, like most of the catchments in Class C1, while Villafranca basin fits Class C1 but belongs to Class C and none of these catchments are borderline cases.

The comparison of the two pools of clusters shows that both physiographic and climate characteristics are important in defining classes: if only climate properties were taken into account, the classification would have changed a great deal. For example, from the climatic point of view, Passobreve and Quinto basins (which are

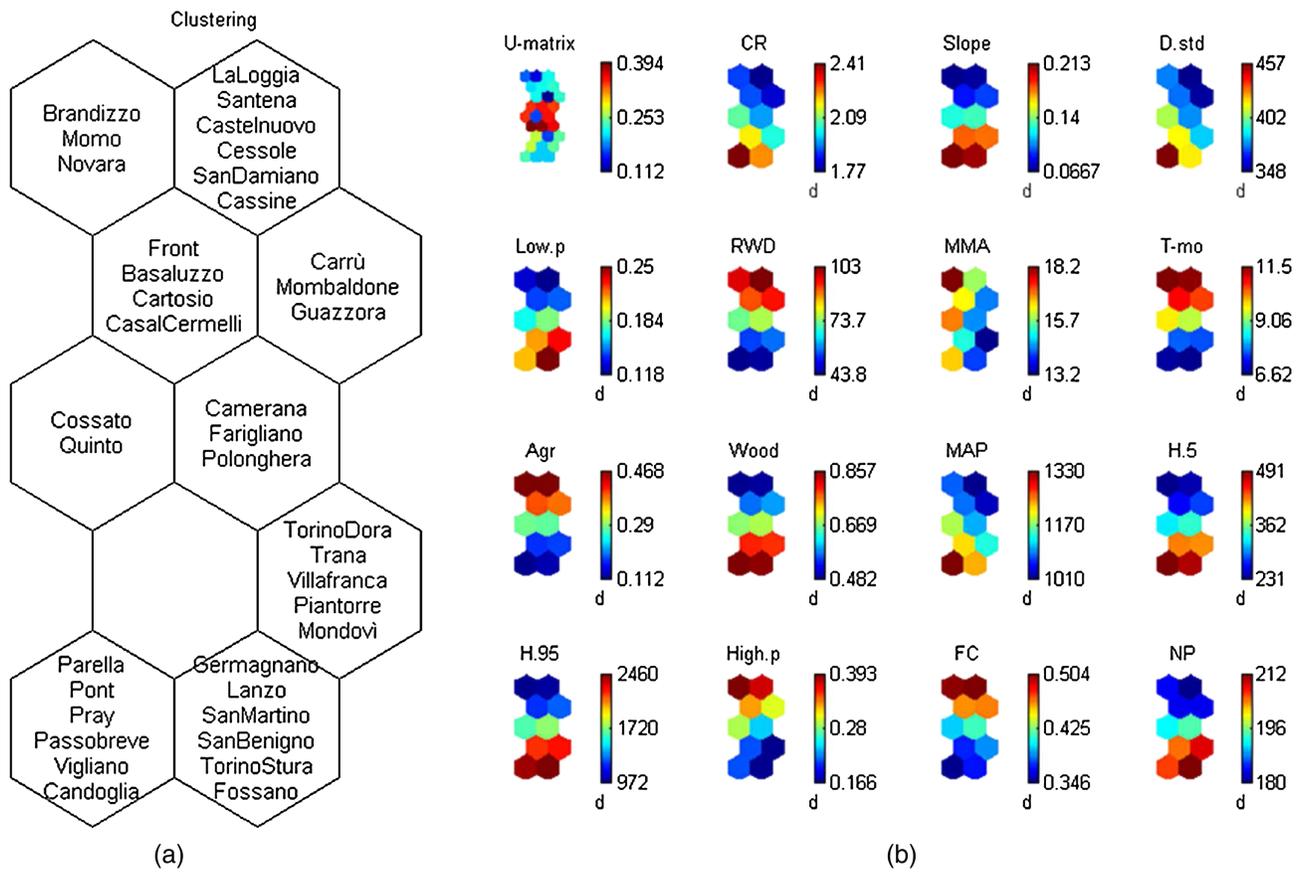


Fig. 6. (a) SOM map with neuron labels obtained for physical properties' classification; (b) component planes for each index and  $U$ -matrix

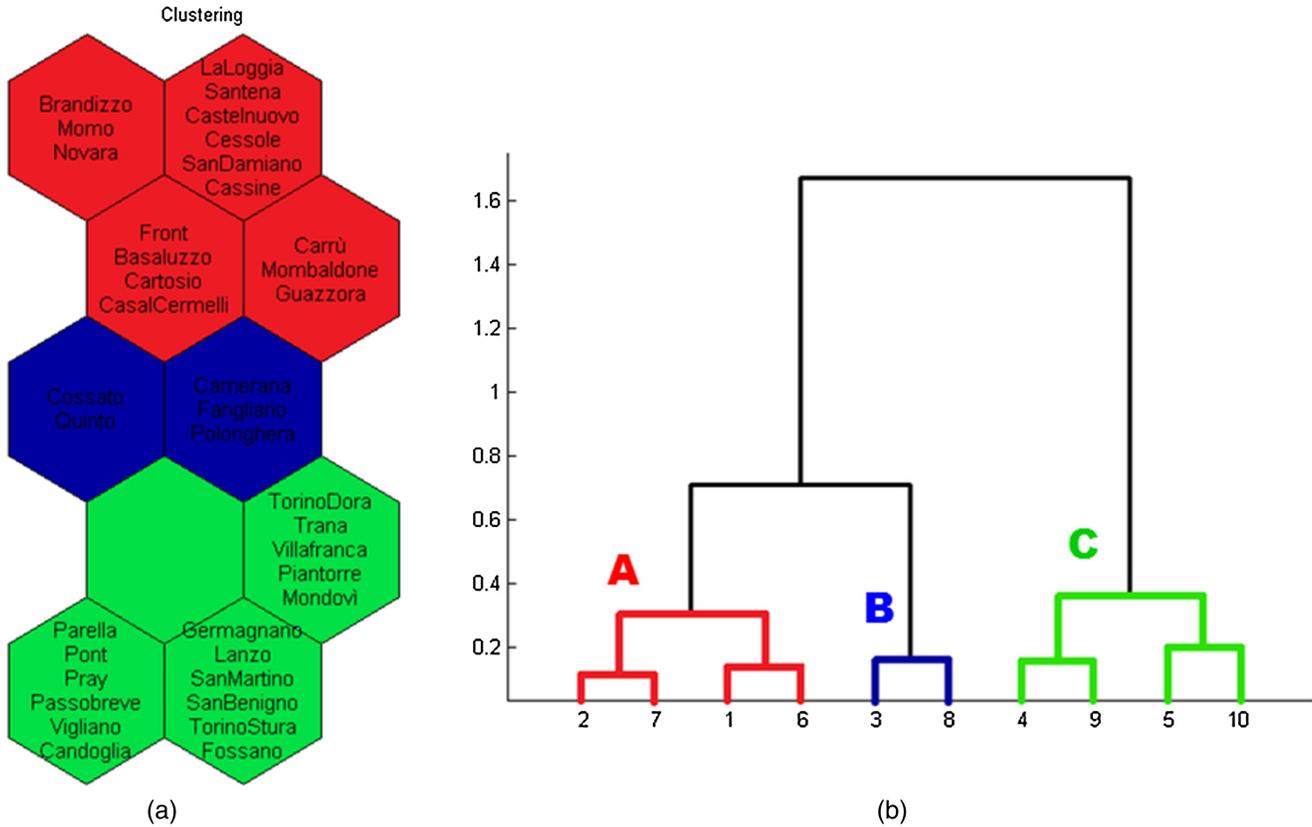


Fig. 7. (a) SOM map with neurons divided in the three classes based on physical properties; (b) dendrogram on SOM; numbers on the dendrogram correspond to each neuron of the SOM map

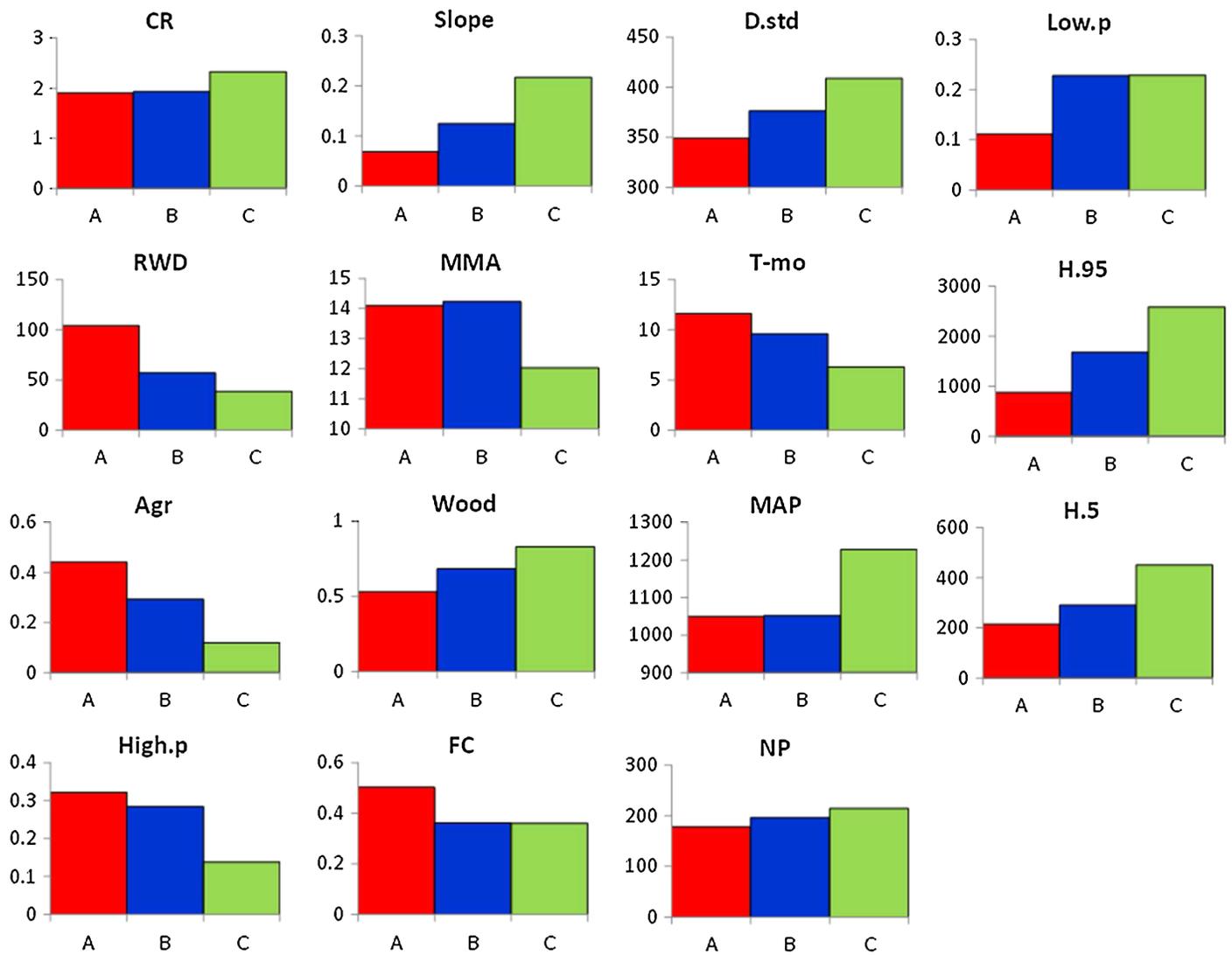


Fig. 8. Median of the 15 physical indexes for each class of physical classification

nested basins) are quite similar, but hydrologic signatures classify them in different classes, meaning that there are other characteristics, such as topography, soil properties, and land use, that influence classification.

Classification of catchments also has to consider scale and heterogeneity of the respective region. The study area is relatively small and there are many nested catchments; however, a remarkable result is that classification shows geographic groups but not continuous regions. In some cases, even if catchments are nested and spatially close, they belong to different classes, meaning that spatial proximity cannot always be an indicator of similarity. This is not always valid, because all the basins in the south of the region show both spatial proximity and similarity. Therefore, an important result emerging from this analysis is that even in a relatively small area, omitting physiographic characteristics could bring errors in classifying catchments.

### Classification Scheme Validation

After computing streamflow signatures for the eight catchments excluded from classification, the Euclidean distances from the centroids of Classes C1, C2, and C3 were calculated, and the class membership was assigned by choosing the minimum distance.

With this procedure, a class was associated to each basin as shown in Fig. 10. Position and classification of the eight basins are shown on the map with triangles. None of the basin belongs to Class C2. They are divided in Class C1 and Class C3, respecting the subdivision into Alpine and Apennine catchments.

To check the validity of the classification scheme, the eight basins were supposed as ungauged, and their membership to Class C1, C2, or C3 was derived by calculating physiographic and climate characteristics and by checking to which class out of A, B, or C that each basin belongs to. Then, knowing the relationship between Class C1 and Class A; Class C2 and Class B; and Class C3 and Class C, it was possible to determine the hydrologic class of all eight basins.

Membership of A, B, or C classes was defined by calculating the Euclidean distance from the centroid of the three classes and the results are reported in Fig. 11, together with position on the map. Comparing the two classifications, a correspondence of 62.5% is found, meaning that five out of eight basins are correctly classified. The three misclassified catchments are (1) Garessio, which passes from effectively Class C3 to Class B, so the exchange is with the intermediate cluster; (2) Perrero; and (3) Ponte di Nava, both of which pass from hydrologic Class C1, to physiographic Class C. In the latter cases, membership in Class C1 is explained by low

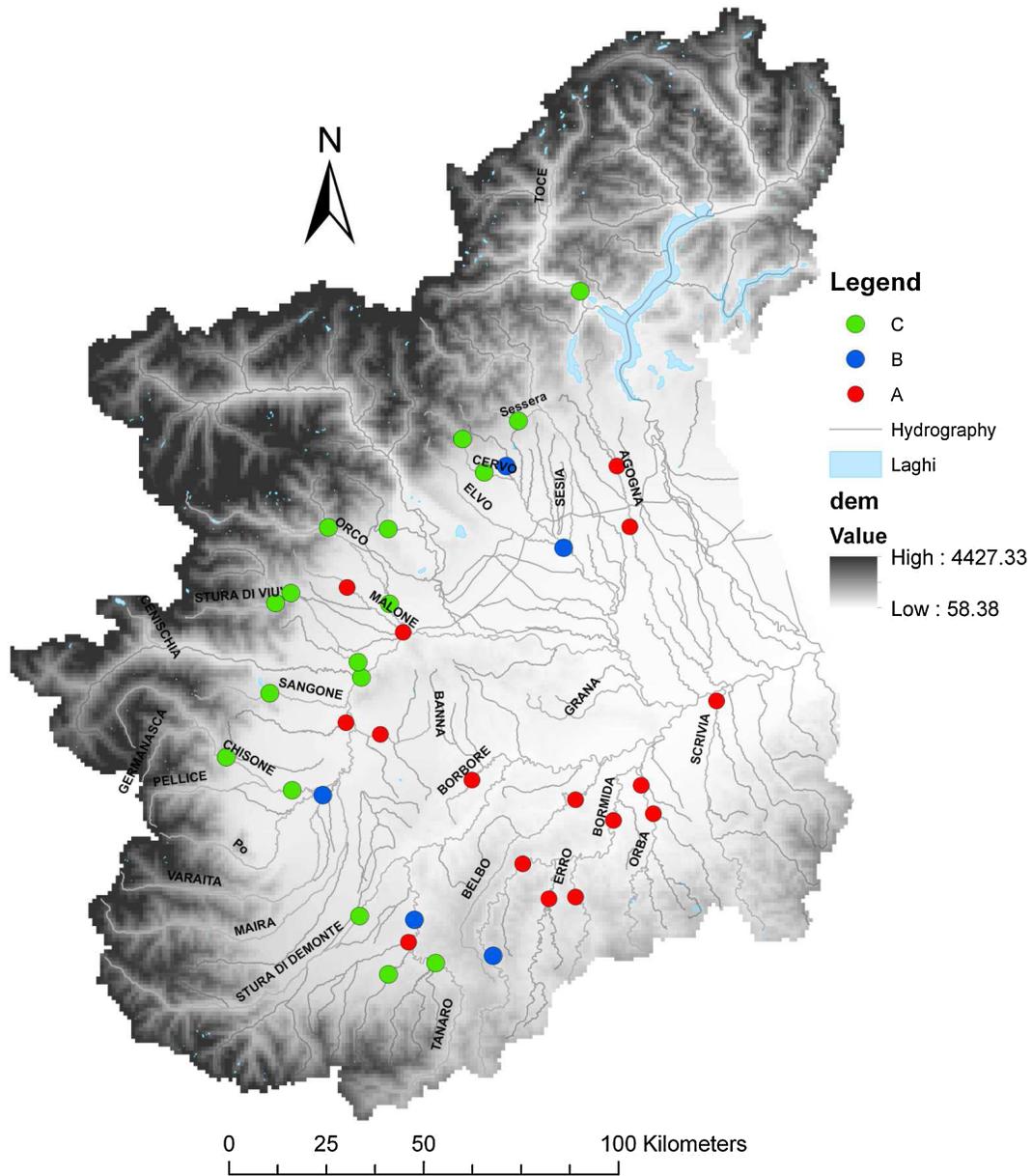


Fig. 9. Classification of catchments based on physical properties

Table 3. Contingency Table between Hydrological and Physical Classification

Streamflow signatures	Physiographic-climate index		
	A	B	C
C1	11	0	3
C2	2	3	2
C3	3	2	12

values of medium flow and by high values of SFDC. The classification into Class C, instead of the expected Class A, is probably due to physiographic characteristics rather than to climate indexes; in fact, both catchments have high average slope, high percentage of soils with low permeability, and relatively high average altitude.

However, the validation is considered to be satisfactory since the percentage of correspondence is quite similar to the one found between the two master classification schemes.

Table 4. Similarity Index between the Hydrological Classes and Physical Groups

Set and similarity index	C1-A	C2-B	C3-C
$\gamma$	14	7	17
$\theta$	16	5	17
$\gamma \setminus \theta$	3	4	5
$\theta \setminus \gamma$	5	2	5
$\gamma \cap \theta$	11	3	12
$ \gamma  +  \theta $	30	12	34
$\gamma \cup \theta$	19	9	22
$n$	38	38	38
SI[-]	0.733	0.500	0.706

#### Application to the Regionalization of Flow-Duration Curves

The analytic FDC presented in Eq. (2a) was extended for application in ungauged sites by estimating regional regression equation.

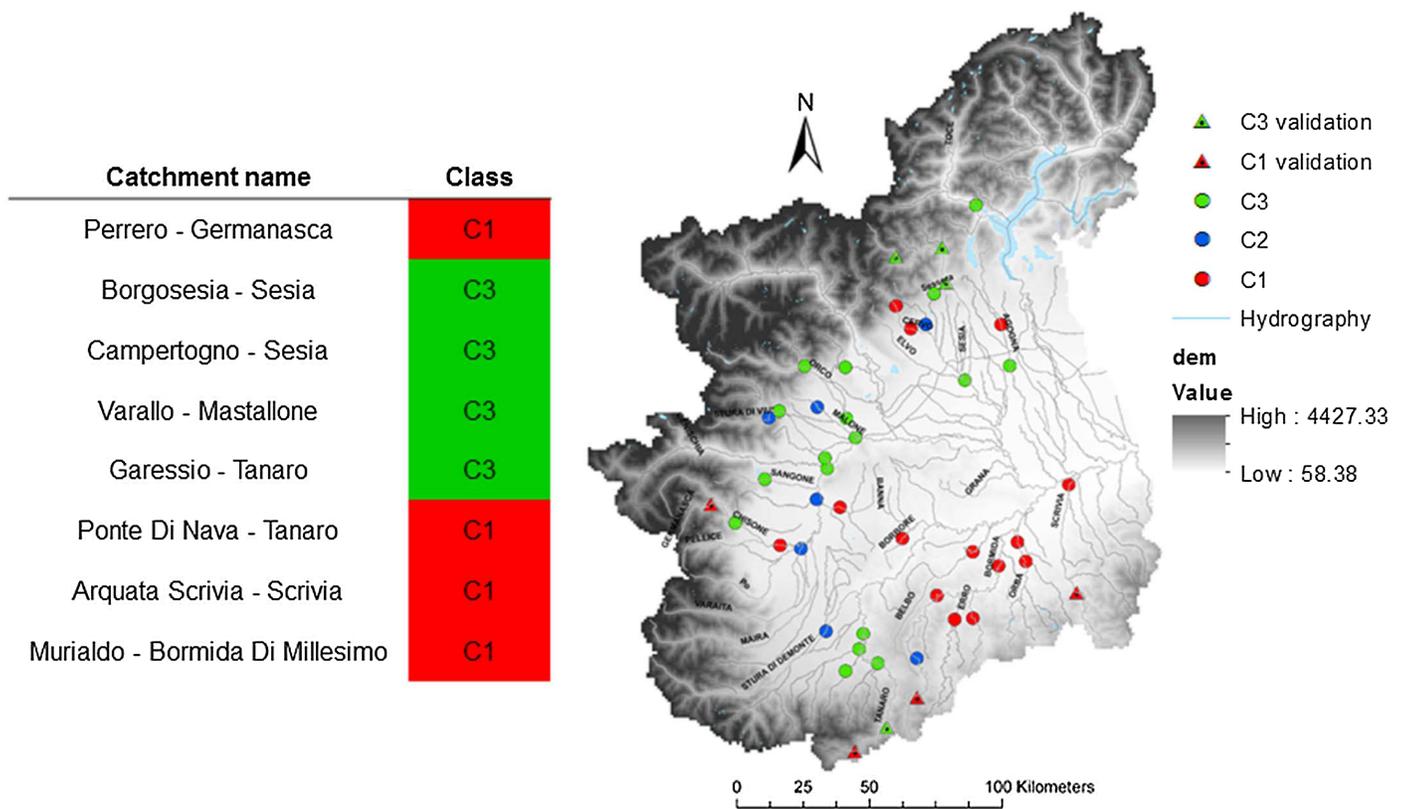


Fig. 10. Classification of the eight validation catchments with streamflow signatures

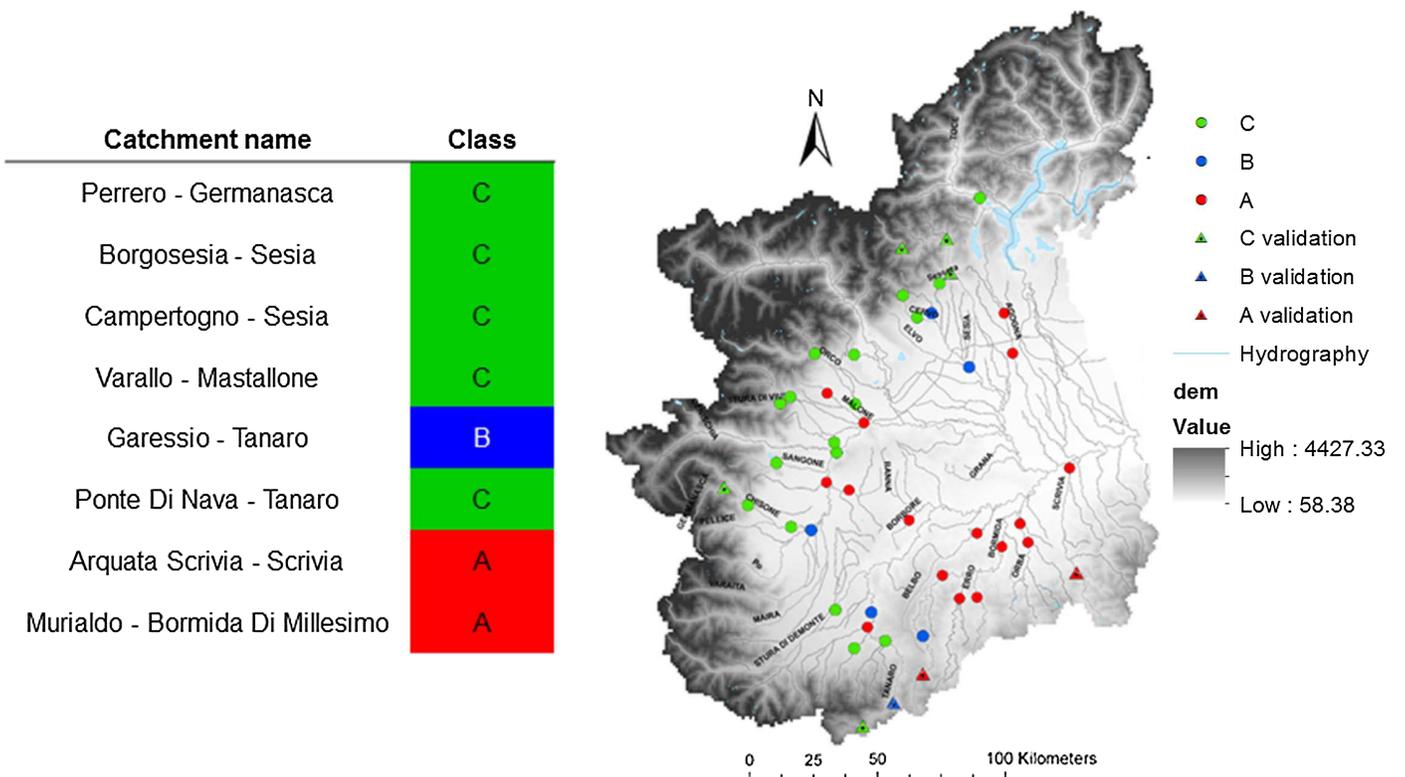


Fig. 11. Classification of the eight validation catchments with physical properties

For each equation, the combination of variable that maximizes adjusted  $R^2$  statistic was chosen.

First, the regression regional model was constructed considering all the catchments together. The identified models are as follows:

$$\hat{\mu} = -14.878 + 1.136 \ln(MMA) + 1.026 \ln(A) + 0.944 \ln(H_{\max} - H_{\text{sta}}) \quad (11a)$$

$$\hat{\sigma} = 4.561 + 0.034 \ln(\text{Low}.p) + 0.432 \ln(\text{Wood}) - 0.432 \ln(H_{\max} - H_{\text{sta}}) \quad (11b)$$

The regional regression model to estimate  $\hat{\mu}$  of the lognormal distribution gives a coefficient of determination  $R^2$  equal to 0.819. The included variables are the median annual maximum of 1-h precipitation,  $MMA$ ; the basin area,  $A$ ; and the difference between the maximum elevation of the basin and the minimum elevation,  $H_{\max} - H_{\text{sta}}$ . The regional regression model for the estimation of  $\hat{\sigma}$  depends on: low permeability,  $\text{Low}.p$ ; the percentage of basin area occupied by wood,  $\text{Wood}$ ; and the difference between the maximum and the minimum elevation,  $H_{\max} - H_{\text{sta}}$ . This regression gives a coefficient of determination  $R^2$  equal to 0.325. The characteristics included in the regression are in agreement with those found by other authors (Mendicino and Senatore 2013; Fennessey and Vogel 1990; Castellarin et al. 2004; Li et al. 2010).

The scatterplot between optimal (i.e., obtained by fitting lognormal distribution to observations) and estimated parameters is shown in Fig. 12. The achieved results using the regression model (black dots) and the ones with cross-validation (white circles) are shown for each site. The cross-validated points overlapping the black dots indicate the stability of the regression model. Performance of the regression model is satisfactory for the mean; however, it is poor for the standard deviation. Indeed, the correlation between optimal standard deviation and the explanatory variables is low.

In order to improve performance of the regression model, the analysis considered the subdivision in hydrological Classes C1, C2, and C3. For each hydrological class, a different pair of regression model for  $\hat{\mu}$  and  $\hat{\sigma}$  was found.

The three couples of models found after regression are as follows:

- For class C1

$$\hat{\mu} = -29.9 + 0.679 \ln(\text{RWD}) + 2.949 \ln(\text{MAP}) + 1.175 \ln(A) \quad (12a)$$

$$\hat{\sigma} = -7.148 - 0.695 \ln(\text{Slope}) + 0.494 \ln(\text{RWD}) + -0.552 \ln(\text{Tmo}) - 1.454 \ln(\text{FC}) + 0.674 \ln(H_{\max} - H_{\text{sta}}) \quad (12b)$$

- For class C2

$$\hat{\mu} = -15.596 + 1.537 \ln(\text{MAP}) + 0.978 \ln(A) \quad (13a)$$

$$\hat{\sigma} = -0.383 - 0.103 \ln(\text{RWD}) + 0.240 \ln(H_{\max} - H_{\text{sta}}) \quad (13b)$$

- For class C3

$$\hat{\mu} = -3.189 + 1.076 \ln(\text{MAP}) - 1.668 \ln(\text{NP}) + 1.055 \ln(A) - 0.184 \ln(A) \quad (14a)$$

$$\hat{\sigma} = -2.897 - 0.084 \ln(\text{Mountain}) - 0.102 \ln(\text{High}.p) + 0.866 \ln(\text{NP}) \quad (14b)$$

where  $\text{Mountain}$  = percentage of mountainous areas within the watershed.

The variables included in the linear regression equation are different for each class, especially for the  $\hat{\sigma}$  estimation equation. On the other hand, the area of the basin ( $A$ ) and the mean annual precipitation ( $\text{MAP}$ ) appear in all equations for estimating the mean.

The implementation of a regional model for each hydrological class led to better results, as shown in Fig. 13, where  $R^2$  statistics obtained using the general Eqs. (11a) and (11b) and the equation valid for the specific class, are shown as well. The  $R^2$  value increases for each of the three clusters when considering the specific class equation.

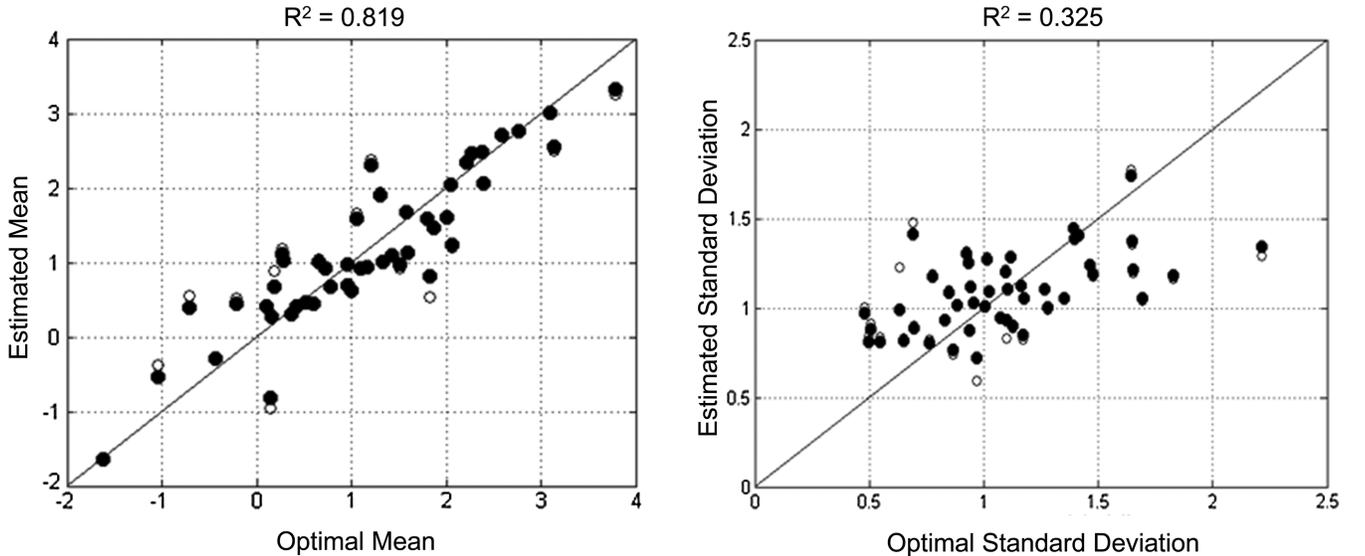
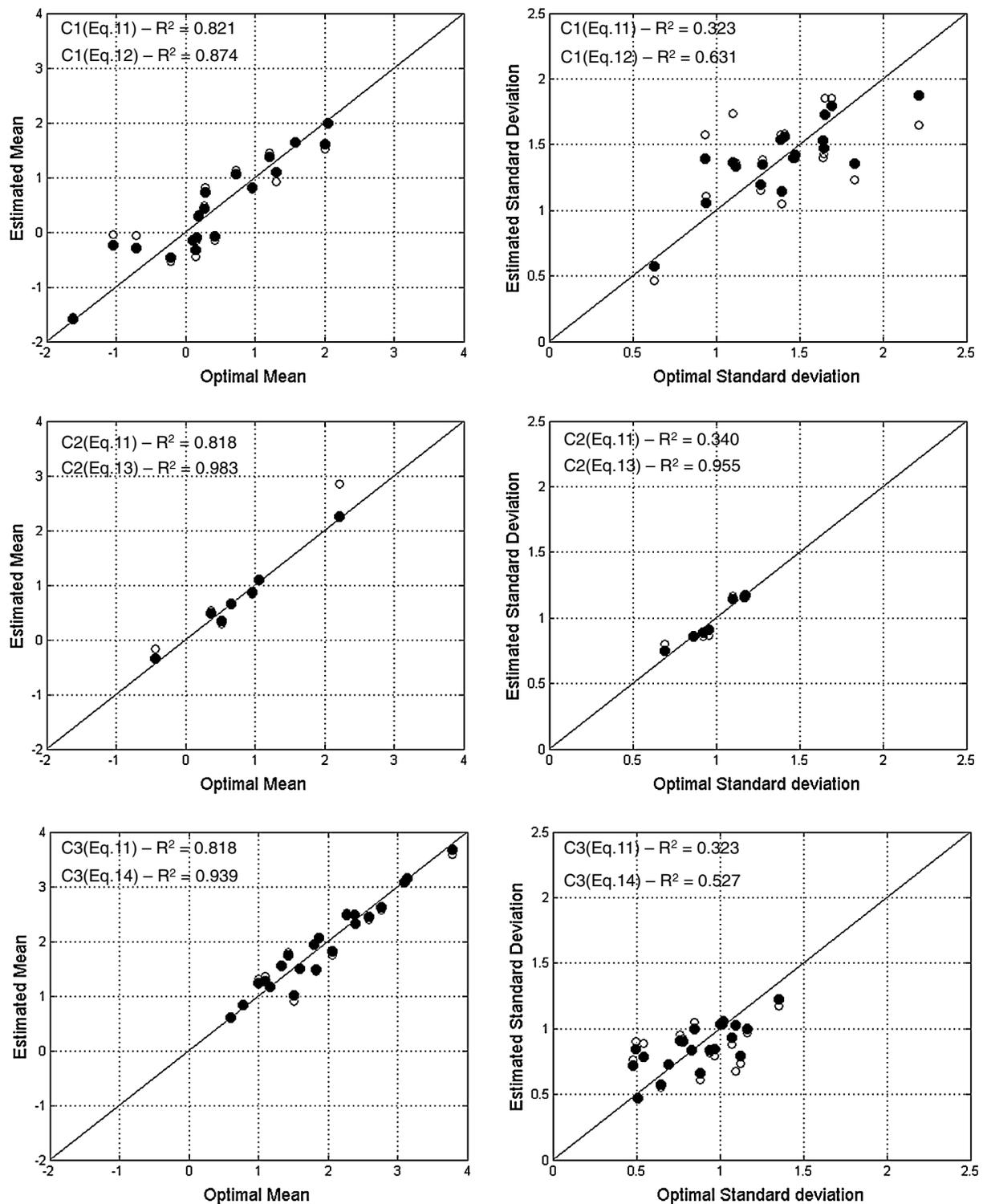


Fig. 12. Scatterplots between optimal (i.e., obtained by fitting lognormal distribution to observations) and estimated parameters



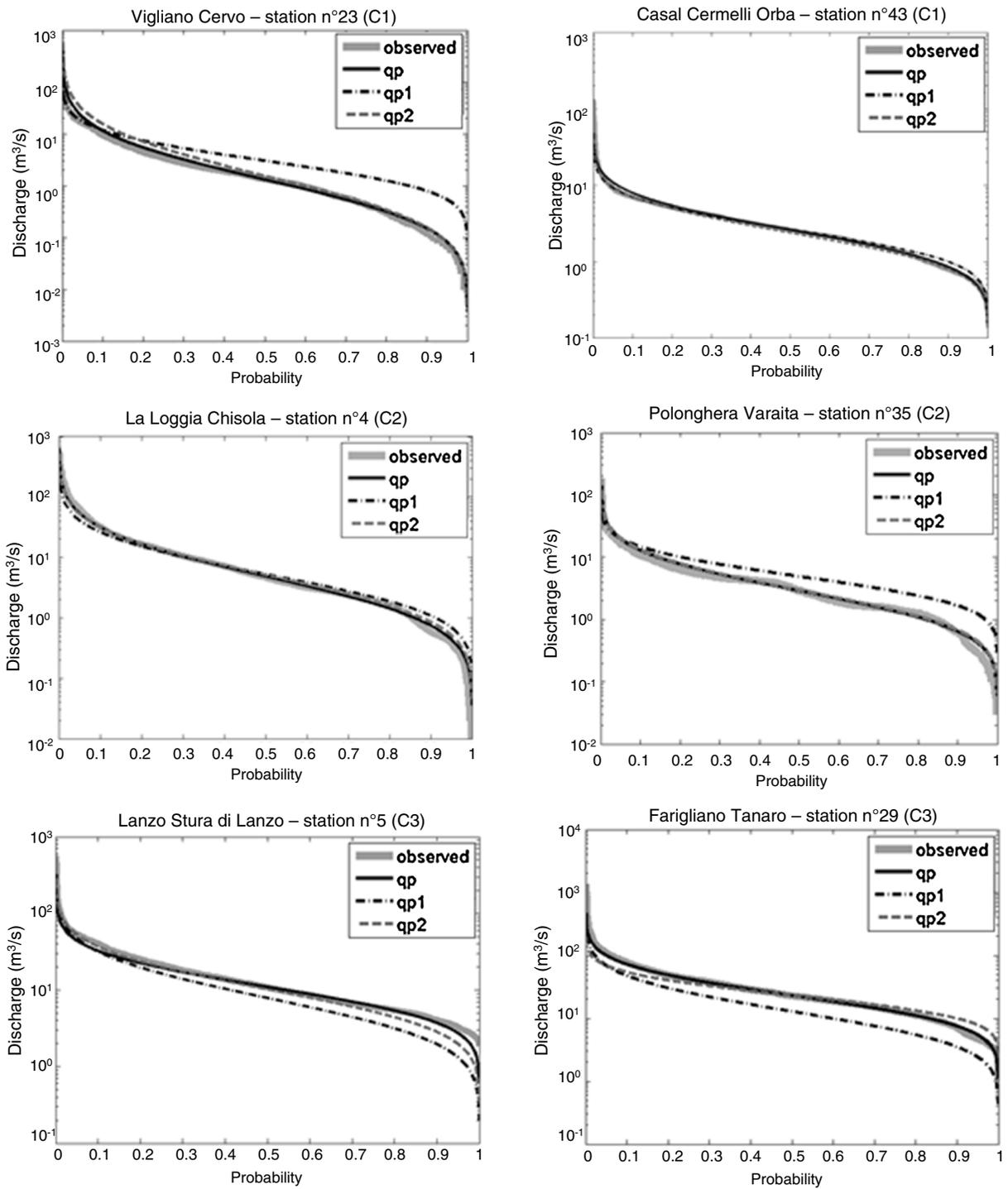
**Fig. 13.** Scatterplots of distribution parameters subdivided into three hydrological classes performed by classification

Some examples of results in term of FDC are shown in Fig. 14, where observed FDC is compared with

- $qp$ , the FDC estimated with log-normal distribution and the optimal value of  $\mu$  and  $\sigma$ ;
- $qp1$ , the FDC estimated with log-normal distribution in which  $\mu$  and  $\sigma$  are estimated with Eqs. (11a) and (11b); and
- $qp2$ , the FDC estimated with lognormal distribution in which  $\mu$  and  $\sigma$  are estimated with Eqs. (12)–(14), depending on the hydrological class of membership.

In all cases presented in Fig. 14, the estimate of FDC with the regionalization that considers hydrological classes shows better results.

In order to test model reliability, a cross-validation was performed. The 46 FDCs were constructed using the leave-one-out estimates of  $\mu$  and  $\sigma$ . Then they were compared with the empirical ones using the index of reliability Eqs. (7)–(10). Figs. 15 and 16 show MAE and MAPE values, respectively, for all analysed basins after cross-validation.



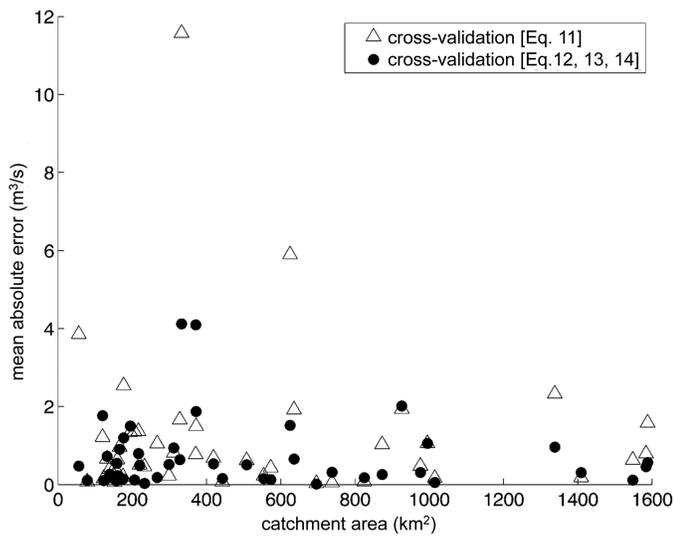
**Fig. 14.** Examples of observed and simulated FDCs with the different proposed models

MAE values obtained with leave-one-out FDCs generally show absolute error lower than  $2 \text{ m}^3/\text{s}$ , while for MAPE, the percentage error is generally lower than 15%. As far as average errors are concerned, MAE, MAPE, and AD obtained with Eqs. (11a) and (11b) are, respectively,  $1.19 \text{ m}^3/\text{s}$ , 10.9%, and 22.9% while the ones obtained with Eqs. (12)–(14) decrease to  $0.71 \text{ m}^3/\text{s}$ , 6.8%, and 8.2% respectively.

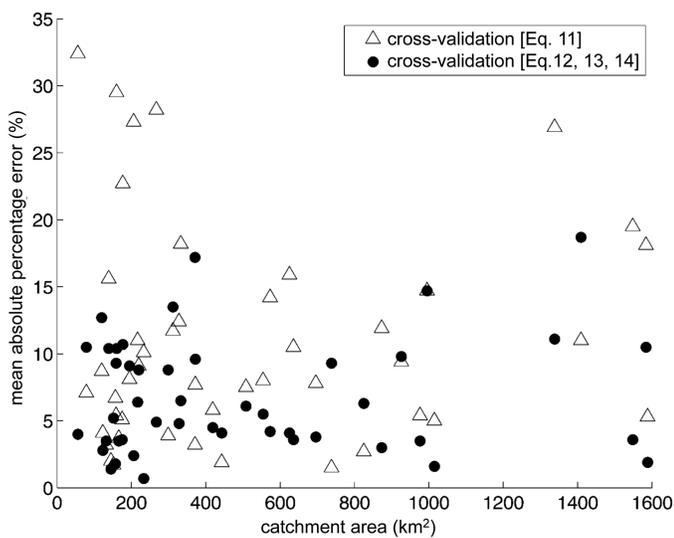
In general, the catchment classification and regionalization of flow duration curves give a good performance. A slight increase of error characterizes basins with low or intermediate elevations. This is probably due to the presence of human disturbances

(irrigation systems, mini hydroelectric power stations, and urban areas).

The values of the efficiency index defined in Eq. (10) were computed for each station and were then used to calculate three further descriptors of the overall quality of flow-duration curves, denoted as P1, P2, and P3, respectively, defined as the percentages of cases for which  $E_s > 0.75$ ;  $0.50 \leq E_s \leq 0.75$ ; and  $E_s \leq 0.50$ . P1 indicates the percentage of cases with good fit, P2 indicates those with fair fit, and P3 indicates those with poor fit. Considering the model performance without catchment classification, the percentage of samples with poor performance (P3) is exactly 50%, so it is not



**Fig. 15.** Mean absolute error (MAE) values for all analyzed basins after cross-validation



**Fig. 16.** Mean absolute percentage error (MAPE) values for all analyzed basin after cross-validation

**Table 5.** Performance Index  $E_s$  after Cross-Validation

Model	P1 (%)	P2 (%)	P3 (%)
Eq. (9)	33	17	50
Eq. (10) (Class C1)	44	28	28
Eqs. (11a) and (11b) (Class C2)	57	29	14
Eqs. (12a) and (12b) (Class C3)	43	29	29
Total [sum of Eqs. (10)–(12a) and (12b)]	46	28	26

satisfactory. When considering the combination of Eqs. (12)–(14), fit is poor in only 26% of the catchments, as shown in Table 5.

## Conclusions

This study presented a classification framework applied to catchments of the Upper Po river basin. Three homogeneous classes

have been detected: Class C1 with the lowest runoff value and characterized by flash-flood events; Class C3 characterized by the highest runoff values; and Class C2 with intermediate behaviour. The results show that there is not a strong subdivision of catchment classifications from a geographic point of view, meaning that not only climate influences runoff regime. Geographic closeness is not a proper criterion for classifying basins in this area.

While the methodology used is already known in literature, this study presented an original combination of indexes that proved to be appropriate for the classification of the complex mountainous river basins of the Alpine region. Usually, climate indexes are considered as mostly influencing catchments' response whereas in this study, the introduction of morphometric and land-use indexes proved to be fundamental. A further element of novelty in this approach is the validation of the found classification scheme, conducted as a sort of test of the predictive power of the found results. The classification was validated by using eight catchments, not used to train SOM; these eight basins were initially supposed to be ungauged so that the membership in hydrological classes C1, C2, or C3 was determined by calculating only physiographic and climate characteristics. The validation brought an overlap of 62.5%. The results show the possibility to identify groups of similar catchments in a relatively small area.

Advantages of applying catchment classification were demonstrated through an application for the regionalization of flow-duration curves. Two kind of regional models were computed, using a stepwise multiple linear regression approach: the first model was implemented considering all 46 basins; the second one considers a set of three models, each for one of the hydrological classes found by cluster analysis. The performance and robustness of these models were checked by means of leave-one-out cross-validation. The obtained results show great performance improvement when the regionalization model is found by taking account of the three different hydrological classes, with a mean absolute percentage error that decreases from 11%, for the single region case to 7% in the three homogeneous regions case. The differences in hydrological regimes are not always a consequence of geographical distance, so the subdivision in similar clusters seems to be a key point in defining a solid and good working model of regionalization.

## Acknowledgments

The authors acknowledge ARPA Piemonte (Regional Environment Protection Agency) for meteorological and flow data, and digital cartography. Elena Toth is thankfully acknowledged for her helpful suggestions and comments on this work.

## References

- Acreman, M. C., and Sinclair, C. D. (1986). "Classification of drainage basins according to their physical characteristics: An application for flood frequency analysis in Scotland." *J. Hydrol.*, 84(3), 365–380.
- Booker, D. J., and Snelder, T. H. (2012). "Comparing methods for estimating flow duration curves at ungauged sites." *J. Hydrol.*, 434–435, 78–94.
- Boscarello, L., Ravazzani, G., and Mancini, M. (2013). "Catchment multisite discharge measurements for hydrological model calibration." *Procedia Environ. Sci.*, 19, 158–167.
- Brath, A., Castellarin, A., Franchini, M., and Galeati, G. (2001). "Estimating the index flood using indirect methods." *Hydrol. Sci. J.*, 46(3), 399–418.

- Castellarin, A., Camorani, G., and Brath, A. (2007). "Predicting annual and long-term flow-duration curves in ungauged basins." *Adv. Water Resour.*, 30(4), 937–953.
- Castellarin, A., Claps, P., Troch, P. A., Wagener, T., and Woods, R. (2011). "Catchment classification and PUB." *Hydrol. Earth Syst. Sci.*, in press.
- Castellarin, A., Galeati, G., Brandimarte, L., Montanari, A., and Brath, A. (2004). "Regional flow-duration curves: Reliability for ungauged basins." *Adv. Water Resour.*, 27(10), 953–965.
- Chapman, T. (1999). "A comparison of algorithms for stream flow recession and baseflow separation." *Hydrol. Processes*, 13(5), 701–714.
- Cheng, L., Yaeger, M., Viglione, A., Coopersmith, E., Ye, S., and Sivapalan, M. (2012). "Exploring the physical controls of regional patterns of flow duration curves—Part 1: Insights from statistical analyses." *Hydrol. Earth Syst. Sci.*, 16(11), 4435–4446.
- Claps, P., and Fiorentino, M. (1997). "Probabilistic flow duration curves for use in environmental planning and management." *Integrated approach to environmental data management systems*, Harmancioglu, Springer, Netherlands, 255–266.
- Clausen, B., and Biggs, B. J. F. (2000). "Flow variables for ecological studies in temperate streams: Grouping based on covariance." *J. Hydrol.*, 237(3–4), 184–197.
- Coopersmith, E., Yaeger, M. A., Ye, S., Cheng, L., and Sivapalan, M. (2012). "Exploring the physical controls of regional patterns of flow duration curves. Part 3: A catchment classification system based on regime curve indicators." *Hydrol. Earth Syst. Sci.*, 16(11), 4467–4482.
- Di Prinzio, M., Castellarin, A., and Toth, E. (2011). "Data-driven catchment classification: Application to the pub problem." *Hydrol. Earth Syst. Sci.*, 15(6), 1921–1935.
- Efron, B. (1982). "The jackknife, the bootstrap and other resampling plans." *Society for industrial and applied mathematics*, SIAM, PA.
- Fennessey, N., and Vogel, R. (1990). "Regional flow-duration curves for ungauged sites in Massachusetts." *J. Water Resour. Plann. Manage.*, 10.1061/(ASCE)0733-9496(1990)116:4(530), 530–549.
- Galelli, S., and Castelletti, A. (2013). "Tree based iterative input variable selection for hydrological modeling." *Water Resour. Res.*, 49(7), 4295–4310.
- Ganora, D., Claps, P., Laio, F., and Viglione, A. (2009). "An approach to estimate non-parametric flow duration curves in ungauged basins." *Water Resour. Res.*, 45(10), W10418.
- Hall, M., and Minns, A. (1999). "The classification of hydrologically homogeneous regions." *Hydrol. Sci. J.*, 44(5), 693–704.
- Hannah, D. M., Kansakar, S. R., Gerrard, A. J., and Rees, G. (2005). "Flow regimes of Himalayan rivers of Nepal: Nature and spatial patterns." *J. Hydrol.*, 308(1–4), 18–32.
- Hellebrand, H., Müller, C., Matgen, P., Fenicia, F., and Savenije, H. (2011). "A process proof test for model concepts: Modeling the meso-scale." *Phys. Chem. Earth*, 36(1–4), 42–53.
- Herbst, M., and Casper, C. (2008). "Towards model evaluation and identification using self-organizing maps." *Hydrol. Earth Syst. Sci.*, 12(2), 657–667.
- Köplin, N., Schädler, B., Viviroli, D., and Weigartner, R. (2012). "Relating climate change signals and physiographic catchment properties to clustered hydrological response types." *Hydrol. Earth Syst. Sci.*, 16(7), 2267–2283.
- Ley, R., Casper, M. C., Hellebrand, H., and Merz, R. (2011). "Catchment classification by runoff behavior with self-organizing maps (SOM)." *Hydrol. Earth Syst. Sci.*, 15(9), 2947–2962.
- Li, M., Shao, Q., Zhang, L., and Chiew, F. H. S. (2010). "A new regionalization approach and its application to predict flow duration curve in ungauged basins." *J. Hydrol.*, 389(1–2), 137–145.
- Mendicino, G., and Senatore, A. (2013). "Evaluation of parametric and statistical approaches for the regionalization of flow duration curves in intermittent regimes." *J. Hydrol.*, 480, 19–32.
- Nash, J., and Sutcliffe, J. V. (1970). "River flow forecasting through the conceptual models. Part 1: A discussion of principles." *J. Hydrol.*, 10(3), 282–290.
- Olden, J. D., and Poff, N. L. (2003). "Redundancy and the choice of hydrologic indices for characterizing streamflow regimes." *River Res. Appl.*, 19(2), 101–121.
- Oudin, L., Andréassian, V., Perrin, C., Michel, C., and Le Moine, N. (2008). "Spatial proximity, physical similarity, regression and ungauged catchments: A comparison of regionalization approaches based on 913 French catchments." *Water Resour. Res.*, 44(3), W03413.
- Patil, S., and Stieglitz, M. (2014). "Modelling daily streamflow at ungauged catchments: What information is necessary?" *Hydrol. Processes*, 28(3), 1159–1169.
- Priestley, C. H. B., and Taylor, R. (1972). "On the assessment of surface heat flux and evaporation using large-scale parameters." *Mon. Weather Rev.*, 100(2), 81–92.
- Razavi, T., and Coulibaly, P. (2013). "Classification of Ontario watersheds based on physical attributes and streamflow series." *J. Hydrol.*, 493, 81–94.
- Sanborn, S. C., and Bledsoe, B. P. (2006). "Predicting streamflow regime metrics for ungauged streams in Colorado, Washington, and Oregon." *J. Hydrol.*, 325(1–4), 241–261.
- Sawicz, K., Wagener, T., Sivapalan, M., Troch, P. A., and Carrillo, G. (2011). "Catchment classification: Empirical analysis of hydrologic similarity based on catchment function in the eastern U.S." *Hydrol. Earth Syst. Sci.*, 15(9), 2895–2911.
- Ssegane, H., Tollner, E. W., Mohamoud, Y. M., Rasmussen, T. C., and Dowd, J. F. (2012). "Advances in variable selection methods II: Effect of variable selection method on classification of hydrologically similar watersheds in three Mid-Atlantic ecoregions." *J. Hydrol.*, 438, 26–38.
- Sylvain, A., and Celisse, A. (2010). "A survey of cross-validation procedures for model selection." *Stat. Surveys*, 4, 40–79.
- Tan, P. N., Steinbach, M., and Kumar, V. (2006). *Introduction to data mining*, Vol. 1, Pearson Addison Wesley, Boston.
- Toth, E. (2013). "Catchment classification based on characterization of streamflow and precipitation time series." *Hydrol. Earth Syst. Sci.*, 17(3), 1149–1159.
- Vesanto, J., and Alhoniemi, E. (2000). "Clustering of the self-organizing map." *IEEE Trans. Neural Networks*, 11(3), 586–600.
- Veza, P., Comoglio, C., Rosso, M., and Viglione, A. (2010). "Low flows regionalization in north-western Italy." *Water Resour. Manage.*, 24(14), 4049–4074.
- Viglione, A., Claps, P., and Laio, F. (2006). "Utilizzo di criteri di prossimità nell'analisi regionale del deflusso annuo." *Proc., Italian Conf. on Hydraulic and Hydraulic Constructions*, Rome.
- Viola, F., Noto, L. V., Cannarozzo, M., and La Loggia, G. (2011). "Regional flow duration curves for ungauged sites in sicily." *Hydrol. Earth Syst. Sci.*, 15(1), 323–331.
- Wallner, M., Haberlandt, U., and Dietrich, J. (2013). "A one-step similarity approach for the regionalization of hydrological model parameters based on self-organizing maps." *J. Hydrol.*, 494, 59–71.
- Winter, T. C. (2001). "The concept of hydrologic landscapes." *J. Am. Water Resour. Assoc.*, 37(2), 335–349.
- Yadav, M., Wagener, T., and Gupta, H. (2007). "Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins." *Adv. Water Resour.*, 30(8), 1756–1774.
- Ye, S., Yaeger, M., Coopersmith, E., Cheng, L., and Sivapalan, M. (2012). "Exploring the physical controls of regional patterns of flow duration curves. Part 2: Role of seasonality, the regime curve, and associated process controls." *Hydrol. Earth Syst. Sci.*, 16(11), 4447–4465.