Hydrol. Earth Syst. Sci. Discuss., 11, 945–985, 2014
www.hydrol-earth-syst-sci-discuss.net/11/945/2014/
doi:10.5194/hessd-11-945-2014

Hydrology and Earth System Sciences
Open Access
Discussions

The influence of methodological procedures on hydrological classification performance

F. J. Peñas[1], J. Barquín[1], T. H. Snelder[2], D. J. Booker[3], and C. Álvarez[1]

[1]Environmental Hydraulics Institute “IH Cantabria”, Universidad de Cantabria, Parque Científico y Tecnológico de Cantabria, C/Isabel Torres 15, 39011 Santander, Spain
[2]Aqualinc Research Ltd, P.O. Box 20-462, Bishopdale, 8543 Christchurch, New Zealand
[3]National Institute of Water and Atmospheric Research, P.O. Box 8602, Christchurch, New Zealand

Received: 29 November 2013 – Accepted: 7 January 2014 – Published: 17 January 2014

Correspondence to: F. J. Peñas (penasfj@unican.es)

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

Abstract

Hydrological classification has emerged as a suitable procedure to disentangle the inherent hydrological complexity of river networks. This practice has contributed to determine key biophysical relations in fluvial ecosystems and the effects of flow modification. Thus, a plethora of classification approaches, which agreed in general concepts and methods but differed largely in specific procedures, have emerged in the last decades. However, few studies have compared the implication of applying contrasting approaches over the same hydrological data. In this work, using cluster analysis and modelling approaches, we classify the entire river network covering the northern third of the Iberian Peninsula. Specifically, we developed classifications of increasing level of detail, ranging from 2 to 20-class levels, either based on raw and normalized daily flow series and using two contrasting approaches to determine class membership: Classify-Then-Predict (ClasF) and Predict-Then-Classify (PredF). Classifications were compared in terms of their statistical strength, the hydrological interpretation, the ability to reduce the bias associated to the underrepresented parts of the hydrological space and the spatial correspondence. The results highlighted that both the data processing and the classification strategy largely influenced the classification outcomes and properties, although differences among procedures were not always statistically significant. The normalization of flow data removed the effect of flow size and generated more complex classifications in which a wider range of hydrologic characteristics were considered. The application of the PredF strategy produced, in most of the cases, classifications with higher discrimination ability, greater ability to address the bias associated with the presence of distinctive gauges and classifications in which classes were more evenly distributed than using the ClasF strategy.

HESSD

11, 945–985, 2014

1 Introduction

Understanding hydrological natural variability has become crucial for river ecology and management because of three main reasons: (i) it is a primary factor influencing river geomorphology (Peñas et al., 2012; Richter et al., 1998; Benda et al., 2004), water (Álvarez-Cabria et al., 2010; Chinnayakanahalli et al., 2011) and biological characteristics (Poff and Zimmerman, 2010), (ii) its variability reflects climate (Morán-Tejeda et al., 2011) and catchment attributes (Monk et al., 2007) and (iii) freshwater resources are essential to maintain many human activities (Naiman and Dudgeon, 2011).

Much progress has been made over the last 20 yr in understanding hydrologic variability and how it promotes self sustaining ecosystems (Poff et al., 2006; Gurnell et al., 2000). However, the inherently complexity of flow regimes hinders both the quantification of direct responses of hydrology to catchment characteristics, and the identification of key hydrology and ecology relationships. The identification and characterization of relevant ecological aspects of the flow regime and the organization of similar rivers into a geographical context (Poff, 1996), trough the definition of hydrological classifications, has emerged as a relevant procedure to structure analyses in hydroecological studies. Specifically, inductive hydrological classification approaches have been used to group river reaches into classes within which key flow regime (Snelder et al., 2009) and ecological attributes (McManamay et al., 2012) are assumed to be similar.

Many of the existing hydrological classifications following the inductive approach rely on the use of statistical procedures to minimize the redundancy of the hydrological information (Olden and Poff, 2003) and also, to reduce the intra-group and increase the inter-groups variability (Snelder and Booker, 2013). These tasks are usually accomplished using Principal Components Analysis (PCA) and Cluster Analysis (CA), respectively (Olden et al., 2012). Nevertheless, many steps within the hydrological classification process may be influenced by a series of subjective decisions depending on the rationale, objectives and available data. For example, many hydrological classifications are based on normalized flow data (McManamay et al., 2012; Kennard et al., 2010; Reidy Liermann et al., 2012) while others used raw flow series (Zhang et al., 2012; Belmar et al., 2011; Alcázar and Palau, 2010). The main reason for normalization is to remove the scale dependence of flow magnitude indices to promote the classification of rivers according to the shape of the regimes. However, normalization can be viewed as a completely subjective choice in the classification process that depends on the objectives of the study (Olden et al., 2012). The shape of the hydrograph provides valuable information about the seasonality, the timing of specific flow events or the patterns of rise and fall of the flow. Undoubtedly these aspects influence river reach ecology (Bunn and Arthington, 2002; Richter et al., 1998) and are key elements for understanding the relationship between climatic and streamflow patterns (Gámiz-Fortis et al., 2011). Nonetheless the size of a river reach and the absolute magnitude of flows also play a key role in ecological processes (Bunn and Arthington, 2002; Vannote et al., 1980).

In addition, beyond the classification of specific sites for which hydrologic data are available (gauged or modelled sites), the scientific and management utility of hydrologic classifications relies on the capacity to extrapolate the class membership to ungauged sites, providing a map of natural flow regimes (Snelder et al., 2009; Reidy Liermann et al., 2012). The Classify-then-Predict (ClasF) strategy has been the most common approach to fulfil this objective (e.g. Kennard et al., 2010; Reidy Liermann et al., 2012). ClasF predicts river reach class membership based on environmental data (climate, topography, geology or land-use) at observed locations. However, this method might pose some flaws when predicting onto an entire region, especially if the distribution of gauges is biased, i.e. specific kind of rivers are under or overrepresented (Snelder and Booker, 2013). If this is the case, the cluster step would fail in accounting for those hydrological features underrepresented in the data set. The presence of sites with exceptional hydrologic character might produce two effects. The first effect is that it may produce higher within-class heterogeneity, while the second is related to the loss of the “rare” hydrologic character when classes are predicted to the whole river network. Because of these reasons some researchers have attempted other approaches such

HESSD
11, 945–985, 2014

as the Predict-then-Classify (PredF) strategy (Ferrier and Guisan, 2006; Snelder and Booker, 2013). Using this approach, hydrological indices obtained from the flow series are predicted onto the entire river network using climate and catchment characteristics, and classification of all river segments is performed as a final stage within the procedure.

The aim of this study was to investigate how the normalization of flow series data previous to the classification procedure and the use of ClasF and PredF influences (i) the classification performance, (ii) the hydrological interpretation of the classifications and their ability to discriminate different hydrological characters, (iii) their ability to reduce the bias associated to the underrepresented parts of the hydrological space and (iv) the degree of spatial correspondence between classifications. To achieve this aim we will develop hydrological classifications of natural conditions over an entire river network in the northern third of the Iberian Peninsula, covering catchments of contrasting climate and spatial configuration. We hypothesised that normalization of river flow data will tend to classify rivers according to their annual regime and not only to the size of the river and also increase the contribution of other hydrological variables not related to flow magnitude. In addition, we hypothesised that the application of the PredF classification procedure will reduce within class heterogeneity, especially when gauges presenting distinctive regimes are included in the classification.

2 Methods

2.1 Study area

The study area comprises the northern third of the Iberian Peninsula (Fig. 1) covering a total area greater than 124 000 km^2. It represents heterogeneous environmental conditions and can be broadly segregate in three main zones. On one hand, the area draining into the Cantabric sea encompass several small basins with drainage areas ranging from 30 to 4907 km^2 covering a total area of 22 000 km^2. Rivers are confined

HESSD
11, 945–985, 2014

by the Cantabrian Cordillera, which reaches up to 2600 m a.s.l. and runs parallel to the coast. Thus, they are characterized by high slopes and short main stream lengths. This region has a humid oceanic temperate climate (Rivas-Martínez et al., 2004). Precipitation is abundant throughout the year with mean of 1300 mm yr^{-1}, with maximum rainfalls in December (150 mm $month^{-1}$) and minimum in July (50 mm $month^{-1}$). However, the precipitation magnitude and distribution varies significantly according to local topography. Snow precipitation is frequent in winter above 1000 m a.s.l. More than 50 % of the surface is occupied by deciduous forest, scrubs and grasslands, while 10 % is occupied by agriculture. The population in this area amounts to almost 3 500 000 inhabitants with a population density of 175 hab km^{-2} although it varies between regions. On the other hand, the Mediterranean area is mainly occupied by the Ebro basin along with a set of medium size basins in the eastern zone. The Ebro basin covers a total extension of 85 530 km^2. It is enclosed by the Cantabrian Mountains and the Pyrenees (3400 m a.s.l.) in the north, by the Catalan Coastal Chain (1712 m a.s.l.) in the east and from the north-west to the south-east by the Iberian massif (2300 m a.s.l.) which creates a dense river network in the catchment boundaries and an extended flat surface in the interior. The Ebro Basin receives both temperate and Mediterranean climate influences. The Pyrenean area (northwest) and the northern part of the Iberian massif present oceanic temperate climate that change gradually to a typical Mediterranean climate in the central Ebro depression. Annual precipitation is 656 mm, however it varies from 300 mm in the centre to the 1700 mm in the highest mountains (Bejarano et al., 2010) where snow is also common during the winter months. The precipitation regime in the Mediterranean region has its maxima in autumn and spring and minima in winter and summer. The temperature regime also oscillates through the year with temperatures over 30 °C in summer and below 5 °C during winter. Population density is below 35 hab km^{-2} which could be considered low, however more than 40 % of the surface is occupied by agricultural land and, thus, the catchment is subjected to an intensive water resource control by more than 216 large dams and other water engineering systems. The eastern zone of the study area comprises several medium catchments ranging from 72 to 5000 km^2, occupying a total extension of 16 500 km^2 that drain directly from the Pyrenees or the Catalan costal chain to the sea. This area is dominated by the Mediterranean oceanic climate in the coast and by a temperate climate in the mountains. Precipitation declines from an annual mean of 1200 mm yr^{-1} in the northern river heads to less than 500 mm yr^{-1} in the southern catchments. Coniferous and broadleaf forest, scrubs and grasslands occupies more than 60 % of the surface in the northern catchments which are progressively replaced by agriculture lands in the south. There are a total of 6 600 000 inhabitants in this area, mostly concentrated in the city of Barcelona and its metropolitan area. Therefore, most of the water resources are allocated to urban and industrial uses.

2.2 Hydrologic data

The initial data set consisted in series of mean daily flow recorded at 428 gauging stations operated by different Spanish water agencies and regional governments. Only gauges unaffected by impoundments (defined as large engineering structures) or large upstream abstractions were selected for analyses. In addition, we selected those gauges with available data for the period 1976–2010 and analyzed the quality of the series. First, an analysis of the flow series was carried out to eliminate those years without desirable data quality, which could be due to the presence of (i) periods of consecutive repeated values, (ii) non-natural extreme low flows for short time periods, (iii) periods of zero flow values in non-intermittent rivers, (iv) non-natural flow magnitude rises and falls or (v) large differences between two periods, probably due to changes of flow record method. Years with more than 30 days of missing data were removed from the analysis. In the last step, we discarded the gauges that accounted with less than 8 yr. After applying these restrictions, 156 gauges were selected with an average length of 17 yr of data (Table 1).

In this study we developed two sorts of classifications, one obtained from normalized flow series and the other from non-normalized (raw) series. Normalization is used to eliminate the influence of flow magnitude (Snelder et al., 2009). Thus, flow series were normalized by dividing all daily flow values by the mean annual flow (Poff et al., 2006)

A set of 103 and 101 hydrologic indices, which represent a wide range of ecologically meaningful aspects of the flow regime (Olden and Poff, 2003), were calculated for the raw and normalized flow series, respectively (Appendix A). These indices characterize the central tendency and dispersion of: (i) magnitude of annual and monthly flows conditions, (ii) magnitude of severe high and low flow conditions, (iii) timing of flows, (iv) frequency and duration of high flow pulses and (v) rate of change of flow (Richter et al., 1996; Olden and Poff, 2003).

Given the strong correlation between several indices, the initial set of indices was reduced to a set of non-correlated synthetic indices using the procedure outlined in Olden and Poff (2003) and followed by many others (Chinnayakanahalli et al., 2011; Zhang et al., 2012; Belmar et al., 2011). Principal Component Analysis (PCA) and the broken stick method (Jackson, 1993) were performed to obtain and define the optimal set of synthetic indices. Two PCAs were carried out independently, one for the hydrologic indices calculated from the raw flow series and another for hydrologic indices calculated from the normalized flow series. Each PC was standardized before conducting further analysis to give them equal weights. Snelder and Booker (2013) demonstrated that this additional step increased classification performance.

2.3 Environmental data

A Synthetic River Network (SRN) was delineated using a 25 m digital elevation model (DEM) using the NestStream software (Miller, 2003). The SRN comprises 667 406 segments with lengths ranging from 16 to 800 m and was used as a spatial network to integrate the hydrological and environmental information.

Climate, topography, land cover and geology are hypothesised to be important discriminator of the hydrologic regime regardless of geographic location. Thus, environmental variables were used to explain the hydrological character of the recorded flow series and predict this character onto the whole river network. Predictor variables

HESSD
11, 945–985, 2014

describing several environmental attributes including climate ($n = 3$), topography ($n = 5$), land cover ($n = 6$) and geology ($n = 2$) were extracted from existing databases provided by several national and regional organizations. The variables for each segment represented the mean value of the variables in the upstream catchment. An initial set of 25 environmental variables with potential influence on the hydrological regimes were selected. Pearson's correlation coefficient between each pair of variables was calculated and variables with correlation higher than 0.7 were discarded. A final set of 16 variables were selected (Table 2): (i) climate: precipitation, precipitation range and evapotranspiration were derived from monthly climate variables calculated in a 1 km grid map by means of interpolation procedure based on data recorded in more than 5000 weather stations of the Spanish network. These data were originally developed to be implemented into the Integrated System for Rainfall-Runoff modelling (in Spanish SIMPA model) by the Centre for Hydrographic Studies (CEDEX, Ministry of Public works and Ministry of Agriculture and Environment, Spain). (ii) Topography: catchment area, slope, elevation, confluence density and drainage density were derived from the 25 m DEM. (iii) Land cover: the percentage surface occupied by broadleaf forest, coniferous forest, pasture, agricultural land, denuded areas and urban areas was derived from the Soil Occupancy Information System (in Spanish SIOSE) developed by the National Geographic Institute of the Spanish Government. SIOSE presents a scale of 1 : 25 000 and integrates satellite and aerial images from several sources of information. (iv) Geology: the average rock hardness and the terrain permeability were derived from the litostatigraphic and permeability map at scale 1 : 200 000 developed by the Spanish Geologic and Miner Institute of the Spanish Government. These variables were calculated using procedures described elsewhere (Fernández et al., 2012; Snelder et al., 2008).

2.4 Classification procedures

In this study, we derived classifications with increasing number of levels using the synthetic hydrologic indices derived from the PCA performed on each of the raw and

HESSD
11, 945–985, 2014

normalized flow series using two contrasting strategies (sensu Snelder and Booker, 2013): (i) the classify-then-predict (rawClasF and norClasF) and the (ii) predict-then-classify (rawPredF and norPredF). The prefix raw and nor indicates whether classification was based on the hydrological indices extracted from the raw or normalized flow series respectively.

2.4.1 Classify-Then-Predict classification (ClasF)

Partitioning Around Medoids (PAM; Kauffman and Rousseeuw, 1990) algorithm on the synthetic indices was used to cluster gauges (Fig. 2). This technique allows the user to specify the number of clusters. We produced classifications with numbers of classes ranging from 2 to 20. We then used Random Forest (RF; Breiman, 2001) to developed predictive models that relate class memberships and catchment properties (Fig. 2). We fitted one specific RF for each classification level (2 to 20-class level) and then, these models were used to establish the most probable class of each segment of the SRN for each classification, i.e. 19 sets of predictions.

2.4.2 Predict-Then-Classify classification (PredF)

For the PredF strategy, empirical models were fitted to each of the standardized synthetic indices as a function of predictor catchment variables using RFs (Fig. 2). Then predictions of the synthetic indices are made for the whole SRN, thereby generating predicted distributions for each synthetic index. Finally, classifications were produced by clustering all the modelled sites using the PAM algorithm varying again between 2 and 20 class levels.

Given the high number of gauges removed due to the presence of impoundments or abstraction upstream, the selected gauges represented “reasonably natural hydrological conditions” only, and probably do not represent the whole spectrum of natural hydrologic conditions in the study area. In addition, the SRN developed for this study presented many rivers of first and second order. Many of them were intermittent or perennial rivers, which is a character underrepresented in the hydrological data base. The prediction of the hydrological synthetic indices or class membership beyond the hydrological space represented in the selected gauges could lead to misleading results. Therefore, instead of using the whole SRN (667 406 segments) in the prediction stage of each approach, those segments of the SRN that presented values of predictor variables out of the range (maximum/minimum) defined by these predictors in the selected the gauges were discarded. Thereby, 178 297 segments were kept.

As stated before, both strategies are based in the use of RF (Breiman, 2001). RF fits many classification and regression trees (CART; Breiman et al., 1984), each of them grown with a randomized subset of sites and predictor variables from the initial data. Each CART is then used to predict the sites initially excluded from the data set, named the out-of-bag (OOB) samples. These predictions are used to calculate the predictive accuracy of the model and the importance of each predictor variable (Snelder et al., 2011).

2.5 Comparison of classification performance

Both the performance of classifications with a given number of classes constructed by different strategies and the performance of classifications with different number of classes derived with the same strategy were compared. The performance of the classifications was measured using the classification strength (CS; Van Sickle, 1997) and ANOVA.

CS estimate the degree of dissimilarity of the hydrological character between gauges explained by the classifications (Snelder and Booker, 2013). This analysis was performed on the hydrological indices with the highest loading on each of the PCs. Briefly, CS results from the difference between the mean dissimilarity of the gauges in the same class (D_{within}) and the mean dissimilarity of gauges in the other classes (D_{between}). Higher values of CS indicate a greater uniformity within classes and greater differences between classes (Van Sickle, 1997). We calculated CS for each classification (rawClasF, rawPredF, norClasF and norPredF each with 2–20 classes). We applied the restriction that classes comprised a minimum of five gauges to reduce the influence in the analysis of classes represented by very few gauges.

In addition, we performed an ANOVA on all the hydrologic indices (103 and 101 for raw and normalized series, respectively) with the class as the explanatory variable to analyze the potential of classifications to discriminate each of the hydrological index. The coefficient of determination (r^2) was calculated for each level (2–20 classes) of the 4 classifications. The restriction of the five gauges per class was also applied.

Following the procedure outlined in Snelder and Booker (2013) and Snelder et al. (2012), both the CS and ANOVA analysis were performed on gauges not used in the fitted models by means of a five-fold cross validation procedure (Hastie et al., 2001). This allowed us focusing on the "predictive performance" of the classifications. Each cross validation procedure was repeated 5 times in order to "smooth out" the variability inherent to each subset. Therefore, results of 25 estimates of predictive CS and r^2 statistics for each hierarchical level of classifications were obtained. Based on the "one standard error rule", two classifications were assumed significantly different if standard errors of the statistics did not intersect.

2.6 Hydrological interpretation of classifications

The original indices with the five highest values in each retained axis of the PCAs were used to interpret the hydrological meaning of the new synthetic indices. In addition, we used the ANOVA results to interpret each classification by looking at the different coefficients of determination for specific indices. We assumed that the higher the coefficient of determination the higher the importance of that index to discriminate among classes.

2.7 Analysis of distinctive gauges

We also analyzed how each classification strategy resolved the problem of the bias associated with the presence of gauges that showed the most distinctive regimes (i.e.

HESSD
11, 945–985, 2014

distant hydrological character). We quantified the effect that the most distinctive gauges produce on the specific classes where they were included. Independent analyses were made for classification based on raw and normalized flow series. First we calculated, based on the standardized synthetic indices scores, the dissimilarity between each pair of gauges and then, the corresponding mean dissimilarity for each gauge. We then selected the 4 most dissimilar gauges and recorded the classes they belonged to when the entire river network was classified. For each distinctive gauge two analyses were performed. Firstly, we calculate the distance between the distinctive gauge and the medoid of the classes in which it was were included. Large distances indicated high class heterogeneity. Secondly, we analyzed the proportion of the classification domain assigned to the classes where the distinctive gauges were included. Low frequency of these classes indicated the inability of the procedure to represent properly certain characteristics of the hydrological space in the entire SRN.

2.8 Correspondence between classifications

The spatial agreement between each pair of classifications was evaluated by means of the Adjusted Rand Index (ARI; Hubert and Arabie, 1985). ARI analyze the relationship of each pair of gauges and how they differ between two cluster solutions. It ranges between 0 (indicating that agreement between two clustering solutions is not better than chance) and 1 (indicating perfect agreement). Given the large number of segments in the SRN, we randomly selected a subset of 500 segments and computed ARI for all pairs of the four classifications.

Bespoke functions written in R were use to analyse flow series, calculate hydrological indices, develop and compare classifciations (Snelder and Booker, 2013).

HESSD
11, 945–985, 2014

3 Results

3.1 PCA and predictive mapping

The broken stick method selected the first five PCs of the PCA performed on the raw series. They explained 91 % of the variance, accounting the PC1 alone for the 68 % (Table 3). The OBB misclassification rate of the RF models in the rawClassF ranged from 0.13 for the 2 classes level to 0.77 for the 20-classes level (Fig. 3). The misclassification rate increased 0.17 (from 0.35 to 0.52) when the classification level increment from 7 to 8 classes. The most important predictor variables of the RF were catchment area, precipitation, agriculture, pasture and elevation. For the rawPredF classification, the mean OBB r^2 for the RF models of the 5 synthetic indices was 0.4 decreasing from 0.65 for PC1 to 0.18 for the PC5. Predictors varied according to the modelled PC, but most of them included topography (catchment area, slope), climate (precipitation) and land cover (agriculture, coniferous and broadleaf forest) variables.

Parallel, the first six PCs of the PCA performed on the normalized flow series were retained. They explained 83.3 % of the variance, with the PC1 and PC2 explaining 38.6 and 20.4 %, respectively (Table 3). For the norClasF strategy the OOB misclassification rate for the RF models range from 0.22 to 0.66 for the 3 and the 18-classes levels, respectively (Fig. 3). Abrupt changes in this rate were recorded between 6 to 7 (decrease) and 7 to 8 (increase) class levels. The most important variables differed between classifications comprising different class levels but in general precipitation, elevation, gradient and broadleaf forest were present in most models. For the norPredF strategy the mean OBB r^2s was 0.31 for the 6 PCs decreasing from 0.63 for PC2 to 0.08 for the PC6. The most important variables were not consistent between RF models although precipitation, elevation, pasture and broadleaf forest were present in most of them.

HESSD
11, 945–985, 2014

3.2 Comparison of classification performance

CS statistics for the classifications based on the raw flow series (rawClasF and rawPredF) showed similar patterns. CS increased from 2 to 5-class level, more pronounced in rawPredF, and decreased slightly beyond this level but, in general, the analysis did not reveal significant differences (i.e. overlapped among standard error bars) between levels of classification (Fig. 4a). RawPredF showed generally higher CS values than rawClasF, although in most cases differences were not significant.

The discrimination power of classifications for each of hydrological index got higher with increasing number of classes (Fig. 5 and Supplement, Table S1). However, in most cases there were not significant differences from 6 or 7 to 20-class levels. Moreover, rawPredF outperformed rawClasF, especially for those indices representing flow magnitude and duration (Fig. 5 and Supplement, Table S1).

NorPredF presented a progressive increment of CS from 2 to 10-class level where it reached the maximum value, suffering then only slight variations (Fig. 4b). NorClasF presented a more unstable CS pattern than norPredF with constant rise and fall of the CS with the increase of class level. Except for specific class levels (2 and 4-class levels), norPredF reached higher CS than norClasF presenting significant differences in the classifications with 6, 7 and 14-class levels.

The discrimination ability of norClasF and norPredF on individual indices showed similar patterns to those found for classifications based on raw series. An increase in r^2 with increasing number of classes and the presence of an inflexion located between 6 and 10-class levels (Fig. 6 and Supplement, Table S2) were observed. In addition, although norPredF performed better than norClasF, differences were not significant in several cases. In general, the classifications based on the raw flow series provided slightly higher CS (Fig. 4) and r^2 values (Figs. 5 and 6) than those based on normalized series.

HESSD
11, 945–985, 2014

3.3 Hydrological interpretation of classifications

According to the hydrological indices with the highest values on each axis in the PCA performed on the raw flow series, PC1 represented the magnitude of the mean annual flow and the magnitude and duration of high flows, while PC2 represented the frequency of high flow events and the magnitude of low flows. PC3 was also related to the frequency of high flow events while PC4 and PC5 represented the variability of rate of change, the asymmetry of flow series and the interannual variability of different hydrological characteristics, respectively (Table 3). The physical interpretation of the PCs became more difficult as variance explained decreased. In addition, ANOVA analysis revealed higher r^2 values of indices related to flow magnitude and duration (I1, M10, MxM4, MnM7, 7LF, 7HF, sdM10, sd7LF, sd7HF) and frequency (FRE3) than those representing other aspects of the flow regime (JMax, Rev, sdFRE3, sdJMax, Rev; Fig. 5 and Supplement, Table S1).

The PCA performed on the normalized flow series showed that PC1 represented the variability of the annual mean flow and the magnitude and duration of extreme low flows and PC2 represented the variability of the magnitude and duration of high flow events. PC3 was related to the mean and variability of the magnitude of monthly flows in the beginning of the humid season (October) while PC4 represented the variability of the magnitude of annual flows and the magnitude of the minimum flows. PC5 was related to mean winter (January) and spring (May) flows while PC6 represented the magnitude and variability of summer (August) flows (Table 3). However, it should be pointed out that interpreting axes becomes rather difficult when variability explained decreases. The highest r^2 values (maximum value around 0.5) were obtained for the indices representing mean monthly flows while the maxima for those indices representing mean and duration of extreme flows was 0.3 (Fig. 6 and Supplement, Table S2). In addition, both norClasF and norPredF showed high discrimination ability on indices representing the frequency of high flow events (FRE), despite these indices were not identified as important in the PCAs.

HESSD
11, 945–985, 2014

3.4 Analysis of distinctive gauges

Three of the four selected distinctive gauges within the classifications based on raw flow series were situated in the Ebro catchment and one in the Cantabric region. The distance between each distinctive gauge and its respective class medoid in the rawPredF classifications was lower than the distance in the rawClasF classification 63 % of the times although only four times the relative differences were greater than 10 % (Table 4). In addition, for the rawClasF it was observed that the proportion of the classification domain assigned to the classes in which the distinctive gauges were included was very low compared to the most evenly distributed classification, i.e. if all the classes had the same proportion, and beyond the 6-class level this proportion was below 1 % for the four distinctive gauges (Fig. 7a). Regarding the rawPredF the proportions of the classes containing the distinctive gauges were higher than for the rawClasF but in general these proportions were below the most even distributed classification (Fig. 7b).

The classifications based on the normalized flow series presented two distinctive gauges situated in the Ebro catchment and the other two in two Catalan catchments. NorPredF showed smaller distances between the distinctive gauges and their respective class medoids than norClasF 89 % of the times and the relative differences were many times over 40 % (Table 4). The proportion of the classes containing the distinctive gauges in the norClasF was, in general, below the frequency showed by the most even distributed classification (Fig. 7c) while the norPredF classifications presented the most similar proportions to the most evenly distributed classification (Fig. 7d).

3.5 Correspondence between classifications

The ARIs for each pair of classifications were in the range 0.15–0.4 for the 6-class level and in the range 0.15–0.3 for the 11 and 16-class levels and the mean of all classification levels (Table 5). The highest ARI was obtained between rawPredF and norPredF (≥ 0.4) and rawPredF and rawClasF (≥ 0.2). Contrary rawClasF and norClasF showed the lowest correspondence (≤ 0.15).

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

HESSD

11, 945–985, 2014

The influence of methodological procedures

F. J. Peñas et al.

Title Page
Abstract Introduction
Conclusions References
Tables Figures
◂◂ ▸▸
◂ ▸
Back Close
Full Screen / Esc

Printer-friendly Version
Interactive Discussion

4 Discussion

As expected the normalization of flow data generated hydrological classifications in which a greater number of hydrological aspects not related with flow magnitude and the size of the river were considered than if data were not normalised, making these classifications more difficult to interpret and predict. In addition, classifications based on the PredF procedure outperformed those obtained with ClasF procedure and in general, dealt better than ClasF with the bias associated to the underrepresented parts of the hydrological space in the original data set.

4.1 Comparison of classification performance

Similar classification performance measured through CS and ANOVA was observed in relation to the results obtained by Snelder and Booker (2013) in New Zealand rivers, which highlights the possibility of applying similar approaches to classify rivers and obtain equivalent results independently of their geographical location as is the case for hydrological regionalization where contiguous regions are delineated.

Our analysis demonstrated that in general, the PredF strategy performed better than ClasF and significant differences in the ability to discriminate hydrological characters were found for several class levels, especially when the classification approaches were applied over the raw flow series. The higher performance of PredF classifications is supported by the conceptual basis of this approach. ClasF imposes sharp barriers to the observed hydrological space and not over the whole hydrologic domain of the fluvial network. Then, the prediction step enforces congruence of all the river reaches with those previously known classes, whereas the real extent to which such discrete groupings exist is uncertain (Kennard et al., 2010). In contrast, the aim of PredF is to account for the whole hydrological variability in the SRN before conducting the classification. This process generates a more complete distribution of the hydrologic variables which is in accordance with the hydrologic reality of the SRN, avoiding the bias associated to gauge location. Moreover, PredF does not assume any interactions between

HESSD
11, 945–985, 2014

the various dependent variables for each RF, which is true as the PCA created orthogonal and independent variables. In addition, it must be pointed out that the PredF approach has not been commonly used in other hydrological classification studies and therefore, further analyses should be done to completely understand the strengths and weakness associated with this strategy.

In general, classifications based on raw flow series had higher discrimination ability for individual indices than those based on normalized flow series (Figs. 5 and 6). As discussed below, classifications based on raw series discriminated rivers based almost exclusively on flow magnitude, which greatly depends on river size. In contrast, classifications based on normalized flow series considered a greater range of hydrological aspects. While the variability of river size shows in general a clear pattern across river networks and thus it is a straightforward approach to segregate river reaches, the consideration of a higher spectrum of hydrologic aspects hampered the creation of classes and thus classifications achieved lower discrimination ability.

4.2 Hydrological interpretation of classifications

Most of the published hydrological classifications are based on normalized flow series (Snelder et al., 2009; Solans and Poff, 2013; Reidy Liermann et al., 2012) or normalized hydrological indices (Kennard et al., 2010; McManamay et al., 2012) while few authors have used the untransformed (raw) data (Belmar et al., 2011; Poff, 1996; Zhang et al., 2012). The use of normalized flow series or indices down weight the influence of flow magnitude on classifications and the application of this criterion significantly affects the final classification outcome. However, to our knowledge this is the first study that has assessed the concern of choosing one of the two approaches. The PCA performed on the raw series showed that the first PC, which was related to mean annual flows and magnitude and duration of high flows explained more than two thirds of the hydrological variability. Thereby, the classifications developed from these data segregated rivers according to their size, as expected. In addition, indices accounting with the frequency of high flow events obtained the highest loadings in the PC2 and PC3 and therefore, this flow regime attribute was relatively well represented in the classifications (Table 6). Moreover, the ANOVA analysis also showed that all the indices related to flow magnitude, even those not included as the most important ones in the PCA presented important differences between classes. This is not surprising given the high correlation between all the flow magnitude indices. However, although these classifications segregated river reaches according to flow magnitude, they were unable to incorporate the severity of droughts. Severity of droughts should be essential aspects to be considered in the classifications given the Mediterranean character of the study zone. For instance, Belmar et al. (2011) working in a Mediterranean catchment and Chinnayakanahalli et al. (2011) covering the Western United States found that, besides the flow magnitude, other hydrologic characteristics related to drought events were contained in the synthetic hydrologic indices. We expected that the characteristic intermittency of many Mediterranean streams had been represented in the synthetic indices, although the lack of this attribute in our classifications may be attributed to the scarcity of gauges situated in intermittent streams. Moreover, the fact that the high differences in flow magnitude between large and small rives have accounted with the largest percentage of variability, have probably masked the effects of low flow attributes.

On the other hand, the interpretation of the classifications based on normalized flow series differed completely to those derived from raw flow series (Table 6). The main differences can be summarized in two essential aspects. First, despite both PCAs explained a similar portion of the total variance, the percentages explained by the different PCs were more evenly distributed in the normalized series. Therefore these classifications were not so obviously conditioned to just one hydrologic character as classifications based on raw series. Second, it was observed that the indices with the highest loading in each PC and hence, their interpretation, varied considerably depending on the data processing (Table 6). In the case of normalized flow series, PC1 represented the magnitude and duration of low flow conditions which means that this classification accounted, to some extent, for the Mediterranean character of the rivers. In addition, PC3 to PC6 were related to the magnitude of flows in different months and periods through the year, therefore classification accounted with the shape of the hydrograph as it has been observed in other works (Bejarano et al., 2010; Solans and Poff, 2013; Snelder et al., 2009). Contrary to expected, other indices not related to flow magnitude, such as the frequency of high flow events were not included as important indices in any PC. Nonetheless, the ANOVA analysis highlighted the ability of these classifications to discriminate the indices representing frequency and therefore it was assured that such an important hydrological aspect was incorporated into the classification.

Finally, it must be pointed out that any of the classifications, whether they were based on raw or normalized data, failed to represent some other important hydrologic aspects such as timing of extreme flow events, predictability, duration of high flow events and rate of change (Table 6). Other analysis based on daily flow series also found low representativeness of these flow regime aspects (Olden and Poff, 2003) which ultimately resulted in a small contribution to the hydrologic classifications (Snelder and Booker, 2013; Snelder et al., 2009).

4.3 Analysis of distinctive gauges

The analyses demonstrated that the PredF approach presented greater capability than ClasF to deal with the underrepresented parts of the hydrological space in the data set. In contrast to Snelder and Booker (2013), we found an underrepresentation of unmodified gauges on large rivers, given the intense flow regime modification that these river types suffer in the study area. Nonetheless, the presence of gauges in small rivers of first and second order was also scarce. If data were not normalized, rawClasF approach generated classes that were comprised by the distinctive gauge plus a very limited number of gauges, in most of the cases less than four. Therefore classes were relatively homogeneous presenting dissimilarity values close to those found in the classifications based on the rawPredF strategy. However rawClasF produced classes with frequencies lower than 1 % which probably were well below the actual frequencies of those river types. On the other hand, the normalization of the flow series smoothed the differences between gauges due to the reduction of the river size effect, which implied that distinctive gauges in the norClasF classifications were not isolated into independent classes. This greatly reduced the problem associated with the low frequency of these classes but in contrast, produced classes with high heterogeneity because distinctive gauges were grouped with other gauges with which they were not that similar. Contrary, the prediction of these rare hydrologic characteristics to a greater number of rivers reaches previous to the classification step through the PredF approach promoted that the proportion of reaches accounting with these rare characteristics increases. Therefore, in the subsequent step of classification, river reaches were grouped together generating classes with a greater degree of homogeneity and classes were more evenly distributed.

4.4 Correspondence between classifications

The ARI analysis has shown that classifications performed over the same data (raw or normalized) with contrasting approaches (ClasF or PerdF) presented a similar correspondence. In general, many ARI values were around 0.2 which implies a certain degree of similarity but still important differences in the spatial distribution of classes. Therefore, although the comparison of the classifications performance did not revealed significant differences for several classification levels, it did not imply that the classification were equivalent regarding the spatial arrangement and highlighted the importance of the classification procedure in the final outcome. In contrast to the expected outcome, ARI analyses also showed that classifications produced using the PredF approach, independently of the data being processing, presented a higher correspondence between them than any other pair. This result highlights that the prediction of the hydrological characteristics to the entire SRN before classifying is probably generating classifications more adjusted to the actual spatial hydrology variability even if classifications presented different interpretation.

HESSD
11, 945–985, 2014

5 Conclusion

In conclusion, this study shows that the methodological procedures used throughout the classification process greatly influences classification outcomes and performance. Although the comparison between ClasF and PredF did not reveal significant differences for several classification levels, the classifications based on PredF produced, in general, higher classification performance, higher ability to discriminate individual indices between classes and greater ability to deal with the bias associated to the presence of gauges with distinctive regimes in the data set. Moreover, the application of the PredF strategy produced more evenly distributed classifications than the ClasF strategy and produced classifications more adjusted to the actual spatial arrangement of hydrologic variability. Therefore, we recommend the application of the PredF strategy although further analyses should be done to completely understand its strengths and weakness. Finally, the pre-processing of flow data influenced the meaning and interpretation of the hydrological classes. The normalization of flow data removed the effect of flow magnitude and generated classifications in which a wider spectrum of hydrologic characteristics was considered. However, the use of raw or normalized data is subject to the final objective and particular application of the classification.

Supplementary material related to this article is available online at http://www.hydrol-earth-syst-sci-discuss.net/11/945/2014/hessd-11-945-2014-supplement.pdf.

HESSD
11, 945–985, 2014

Acknowledgements. This study was partly funded by the Spanish Ministry of Economy and Competitiveness as part of the project MARCE (Ref: CTM-2009-07447; http://marce.ihcantabria.es/) and the Ministry of Agriculture, Food and Environment as part of the project RECORAM (Ref: 132/2010). José Barquín is supported by a Ramon y Cajal grant (Ref: RYC-2011-08313) of theof Ministry of Economy and Competitiveness. We would also like to thank Confederación Hidrográfica del Cantábrico, Confederación Hidrográfica del Ebro, Agencia Vasca del Agua, Agencia Catalana del Agua and Gobierno de Navarra for providing flow series data.

References

Alcázar, J. and Palau, A.: Establishing environmental flow regimes in a Mediterranean watershed based on a regional classification, J. Hydrol., 388, 41–51, doi:10.1016/j.jhydrol.2010.04.026, 2010.

Álvarez-Cabria, M., Barquín, J., and Juanes, J. A.: Spatial and seasonal variability of macroinvertebrate metrics, do macroinvertebrate assemblages track river health?, Ecol. Indic., 10, 370–379, doi:10.1016/j.ecolind.2009.06.018, 2010.

Bejarano, M. D., Marchamalo, M., García de Jalón, D., and González del Tánago, M.: Flow regime patterns and their controlling factors in the Ebro basin (Spain), J. Hydrol., 385, 323–335, doi:10.1016/j.jhydrol.2010.03.001, 2010.

Belmar, O., Velasco, J., and Martínez-Capel, F.: Hydrological classification of natural flow regimes to support environmental flow assessments in intensively regulated Mediterranean rivers, Segura River Basin (Spain), Environ. Manage., 47, 992–1004, doi:10.1007/s00267-011-9661-0, 2011.

Benda, L., Poff, N. L., Miller, D., Dunne, T., Reeves, G., Pess, G., and Pollock, M.: The network dynamics hypothesis: how channel networks structure riverine habitats, Bioscience, 54, 413–427, doi:10.1641/0006-3568(2004)054[0413:TNDHHC]2.0.CO;2, 2004.

Breiman, L.: Bagging predictors, Mach. Learn., 24, 123–140, doi:10.1023/A:1018054314350, 1996.

Breiman, L.: Random forest, Mach. Learn., 45, 5–32, doi:10.1023/A:1010933404324, 2001.

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J.: Classification and Regression Trees, Wadsworth, Inc., Monterey, Calif., USA, 1984.

HESSD
11, 945–985, 2014

Bunn, S. E. and Arthington, A. H.: Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity, Environ. Manage., 30, 492–507, doi:10.1007/s00267-002-2737-0, 2002.
Chinnayakanahalli, K. J., Hawkins, C. P., Tarboton, D. G., and Hill, R. A.: Natural flow regime, temperature and the composition and richness of invertebrate assemblages in streams of the western United States, Freshwater Biol., 56, 1248–1265, doi:10.1111/j.1365-2427.2010.02560.x, 2011.
Colwell, R. K.: Predictability, Constancy, and Contingency of Periodic Phenomena, Ecology, 55, 1148–1153, 1974.
Fernández, D., Barquín, J., Álvarez-Cabria, M., and Peñas, F. J.: Quantifying the performance of automated GIS-based geomorphological approaches for riparian zone delineation using digital elevation models, Hydrol. Earth Syst. Sci., 16, 3851–3862, doi:10.5194/hess-16-3851-2012, 2012.
Gámiz-Fortis, S. R., Hidalgo-Muñoz, J. M., Argüeso, D., Esteban-Parra, M. J., and Castro-Díez, Y.: Spatio-temporal variability in Ebro river basin (NE Spain): global SST as potential source of predictability on decadal time scales, J. Hydrol., 409, 759–775, doi:10.1016/j.jhydrol.2011.09.014, 2011.
Gurnell, A. M., Hupp, C. R., and Gregory, S. V.: Linking hydrology and ecology, Hydrol. Process., 14, 2813–2815, doi:10.1002/1099-1085(200011/12)14:16/17<2813::AID-HYP120>3.0.CO;2-Q, 2000.
Hastie, T., Tibshirani, R., and Friedman, J. H.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer-Verlag, New York, 2001.
Hubert, L. and Arabie, P.: Comparing Partitions, J. Classif., 2, 193–218, 1985.
Jackson, D. A.: Stopping rules in principal components-analysis – a comparison of heuristic and statistical approaches, Ecology, 74, 2204–2214, doi:10.2307/1939574, 1993.
Kauffman, L. and Rousseeuw, P. J.: Finding Groups in Data, An Introduction to Cluster Analysis, Wiley and Sons, New-York, 1990.
Kennard, M. J., Pusey, B. J., Olden, J. D., MacKay, S. J., Stein, J. L., and Marsh, N.: Classification of natural flow regimes in Australia to support environmental flow management, Freshwater Biol., 55, 171–193, doi:10.1111/j.1365-2427.2009.02307.x, 2010.
McManamay, R. A., Orth, D. J., Dolloff, C. A., and Frimpong, E. A.: A regional classification of unregulated stream flows: spatial resolution and hierarchical frameworks, River Res. Appl., 28, 1019–1033, doi:10.1002/rra.1493, 2012.

HESSD
11, 945–985, 2014

Miller, D.: Programs for DEM Analysis, Earth System Institute, Mount Shasta, CA, 38 pp., 2003.
Monk, A. W., Wood, P. J., and Hannah, D. M.: Examining the influence of flow regime variability on instream ecology, in: Hydroecology and Ecohydrology: Past, Present and Future, edited by: Wood, P. J., Hannah, D. M., and Sadler, J. P., John Wiley & Sons, Ltd., Chichester, 2007.
Morán-Tejeda, E., López-Moreno, J. I., Ceballos-Barbancho, A., and Vicente-Serrano, S. M.: River regimes and recent hydrological changes in the Duero basin (Spain), J. Hydrol., 404, 241–258, doi:10.1016/j.jhydrol.2011.04.034, 2011.
Naiman, R. J. and Dudgeon, D.: Global alteration of freshwaters: influences on human and environmental well-being, Ecol. Res., 26, 865–873, doi:10.1007/s11284-010-0693-3, 2011.
Olden, J. D. and Poff, N. L.: Redundancy and the choice of hydrologic indices for characterizing streamflow regimes, River Res. Appl., 19, 101–121, doi:10.1002/rra.700, 2003.
Olden, J. D., Kennard, M. J., and Pusey, B. J.: A framework for hydrologic classification with a review of methodologies and applications in ecohydrology, Ecohydrology, 5, 503–518, doi:10.1002/eco.251, 2012.
Peñas, F. J., Barquín, J., Snelder, T., Booker, D., Fernandez, D., and Álvarez-Cabria, M.: Do rivers reaches differ in habitat-flow relationships according to hydrologic classification and river size?, 9th International Symposium on Ecohydraulics Proceedings, 17–21 September 2012, Vienna, Austria, 14956, 2012.
Poff, N. L.: A hydrogeography of unregulated streams in the United States and an examination of scale-dependence in some hydrological descriptors, Freshwater Biol., 36, 71–91, doi:10.1046/j.1365-2427.1996.00073.x, 1996.
Poff, N. L. and Zimmerman, J. K. H.: Ecological responses to altered flow regimes: a literature review to inform the science and management of environmental flows, Freshwater Biol., 55, 194–205, doi:10.1111/j.1365-2427.2009.02272.x, 2010.
Poff, N. L., Olden, J. D., Pepin, D. M., and Bledsoe, B. P.: Placing global stream flow variability in geographic and geomorphic contexts, River Res. Appl., 22, 149–166, doi:10.1002/rra.902, 2006.
Reidy Liermann, C. A., Olden, J. D., Beechie, T. J., Kennard, M. J., Skidmore, P. B., Konrad, C. P., and Imaki, H.: Hydrogeomorphic classification of Washington state rivers to support emerging environmental flow management strategies, River Res. Appl., 28, 1340–1358, doi:10.1002/rra.1541, 2012.

Richter, B. D., Baumgartner, J. V., Powell, J., and Braun, D. P.: A method for assessing hydrologic alteration within ecosystems, Conserv. Biol., 10, 1163–1174, doi:10.1046/j.1523-1739.1996.10041163.x, 1996.

Richter, B. D., Baumgartner, J. V., Braun, P. D., and Powell, J.: A spatial assessment of hydrologic alteration within a river network, Regul. River., 14, 329–340, doi:10.1002/(SICI)1099-1646(199807/08)14:4<329::AID-RRR505>3.0.CO;2-E, 1998.

Rivas-Martínez, S., Penas, A., and Díaz, T. E.: Bioclimatic Map of Europe, Bioclimates, Cartographic Service, University of León, León, Spain, 2004.

Snelder, T. H. and Booker, D.: Natural flow regime classifications are sensitive to definition procedures, River Res. Appl., 7, 822–838, doi:10.1029/2009WR008839, 2013.

Snelder, T. H., Pella, H., Wasson, J.-G., and Lamouroux, N.: Definition procedures have little effect on performance of environmental classifications of streams and rivers, Environ. Manage., 42, 771–788, doi:10.1007/s00267-008-9188-1, 2008.

Snelder, T. H., Lamouroux, N., Leathwick, J. R., Pella, H., Sauquet, E., and Shankar, U.: Predictive mapping of the natural flow regimes of France, J. Hydrol., 373, 57–67, doi:10.1016/j.jhydrol.2009.04.011, 2009.

Solans, M. A. and Poff, N. L.: Classification of natural flow regimes in the Ebro Basin (Spain) by using a wide range of hydrologic parameters, River Res. Appl., 9, 1147–1163, doi:10.1002/rra.2598, 2013.

Vannote, R. L., Minshall, G. W., Cummins, K. W., Sedell, J. R., and Cushing, C. E.: The river continuum concept, Can. J. Fish. Aquat. Sci., 37, 130–137, 1980.

Van Sickle, J.: Using mean similarity dendogram to evaluate classifications, J. Agr. Biol. Envir. St., 2, 370–388, 1997.

Zhang, Y., Arthington, A. H., Bunn, S. E., Mackay, S., Xia, J., and Kennard, M.: Classification of flow regimes for environmental flow assessment in regulated rivers: the Huai River Basin, China, River Res. Appl., 28, 989–1005, doi:10.1002/rra.1483, 2012.

HESSD
11, 945–985, 2014

Table 1. Number of retained years for flow time-series used in the analysis.

N. of years	N. of gauges	Frequency	Freq. acum. frequency
> 19	52	33.3	33.3
19	3	1.9	35.3
18	7	4.5	39.7
17	6	3.8	43.6
16	16	10.3	53.8
15	7	4.5	58.3
14	8	5.1	63.5
13	8	5.1	68.6
12	11	7.1	75.6
11	9	5.8	81.4
10	9	5.8	87.2
9	9	5.8	92.9
8	11	7.1	100.0

HESSD
11, 945–985, 2014

The influence of methodological procedures

F. J. Peñas et al.

Table 2. Environmental variables used to predict classes or the synthetic hydrologic indices onto the ungauged segments of the river network (TG: Topography; CL: Climatic LC: Land Cover; GL: Geology).

Variable	Type	Units	Description	Source
Precipitation	CL	mm	Annual catchment precipitation	SIMPA
Precipitation range	CL	mm	Range between maximum and minimum seasonal precipitation	SIMPA
Evapotranspiration	CL	Mm	Annual catchment evapotranspiration	SIMPA
Catchment area	TG	km^2	Total catchment area	DEM
Slope	TG	%	Average catchment gradient	DEM
Elevation	TG	m	Average catchment elevation	DEM
Confluence density	TG	–	Number of rivers confluences by catchment area	DEM
Drainage density	TG	–	Number of segments divided by the catchment area	DEM
Broadleaf forest	LC	%	Surface occupied by broadleaf forest	SIOSE
Coniferous forest	LC	%	Surface occupied by coniferous	SIOSE
Pasture	LC	%	Surface occupied by pasture	SIOSE
Agriculture	LC	%	Surface occupied by agricultural land	SIOSE
Denuded	LC	%	Surface occupied by denuded areas	SIOSE
Urban	LC	%	Surface occupied by urban areas	SIOSE
Permeability	GL	–	Terrain permeability	IGM
Hardness	GL	–	Rock hardness	IGM

HESSD
11, 945–985, 2014

Table 3. The 5 hydrologic indices with the highest loadings in each PC and variation explained by the retained PCs using the raw (above) and the normalized flow series (below). A minus sign indicates negative relation with the PC.

Flow series	Axe	Hydrologic variables with the highest values in the PCs	Variation Explained (%)
Raw	PC1	-I1, -X25, -90HF, - 30HF, -M11	68
Raw	PC2	-FRE7, -FRE3, -lcv, BFI, sdBFI	10.6
Raw	PC3	-FRE1, -nPH, -FRE3, dPH, sdZFD	5.9
Raw	PC4	sdnPos, sdnNeg, ikur, Ica	3.6
Raw	PC5	-sdnPH, sdJMax, -sdRev, -sdFRE3, -sdJmin	3.5
Normalized	PC1	-I2, X75, 90LF, 30LF, 7LF	38.6
Normalized	PC2	sd30HF, sd7HF, sd3HF, sd90HF, sdM5	20.4
Normalized	PC3	-M10, -sdM10, -MXM10, -FRE1, sdM9	11.6
Normalized	PC4	ikur, X25, MnM9, MnM2, MnM11	7.1
Normalized	PC5	-M1, M5, sdZFD, -sdM1, -MxM1,	6.1
Normalized	PC6	SdM8, MXM8, sdnPH, -MxM11, -sdM11	4.5

Table 4. Euclidean distance between the distinctive gauges (MG) and the medoid of the classes in which they were included for the 4, 6, 8, 10, 12, 16 and 20-class levels classification. Empty cells indicated that the gauge is the unique gauge in the class. Bold letters indicate the procedure that showed the lowest distance.

Raw series	MG 1		MG 2		MG 3		MG 4	
	rawClasF	rawPredF	rawClasF	rawPredF	rawClasF	rawPredF	rawClasF	rawPredF
4	15.92	**10.56**	13.57	**9.66**	7.01	**6.14**	**4.67**	5.17
6	18.19	**10.39**	9.55	**9.04**	**3.56**	3.88	**5.17**	5.50
8		10.49	9.55	**6.19**	**3.56**	3.79	5.17	**3.78**
10		10.23	9.55	**8.77**	**3.56**	3.62	4.69	**4.56**
12		2.61	9.55	**8.94**	3.56	**3.40**	4.40	**4.32**
16		9.63		6.19	3.56	**3.41**	3.10	**3.09**
20		8.17		6.41	3.56	**3.40**	**2.50**	3.09

Raw series	MG 1		MG 2		MG 3		MG 4	
	norClasF	norPredF	norClasF	norPredF	norClasF	norPredF	norClasF	norPredF
4	11.26	**6.12**	9.89	**5.66**	9.29	**4.85**	9.94	**5.13**
6	10.96	**5.36**	9.67	**5.25**	9.40	**5.08**	9.85	**5.05**
8	10.96	**6.11**	10.30	**5.31**	7.22	**5.72**	5.57	**4.45**
10	9.45	**6.04**	9.67	**5.31**	7.22	**5.17**	5.57	**4.42**
12	9.45	**4.93**	9.67	**4.93**	7.22	**4.85**	5.57	**4.41**
16		5.36	**5.45**	5.61	6.41	**3.55**	6.36	**3.51**
20		4.91	5.45	**5.11**	6.41	**5.19**	6.36	**3.51**

HESSD
11, 945–985, 2014

Table 5. Adjusted Rand Index (ARI) for the 6, 11 and 16-class level and the mean of all class levels classifications following the four approaches.

Level	Classification	Classification rawClasF	rawPredF	norClasF
6	rawPredF	0.20		
	norClasF	0.13	0.18	
	norPredF	0.19	0.41	0.19
11	rawPredF	0.24		
	norClasF	0.15	0.22	
	norPredF	0.18	0.31	0.22
16	rawPredF	0.23		
	norClasF	0.15	0.19	
	norPredF	0.17	0.30	0.19
Mean of all levels	rawPredF	0.22		
	norClasF	0.15	0.19	
	norPredF	0.17	0.31	0.20

HESSD
11, 945–985, 2014

Table 6. Relative representativeness of each flow regime aspect according to the data processing previous to classification procedure.

Flow Aspect		Raw	Normalized
Magnitude of annual flows	Mean	c	
	Variability	a	c
Magnitude of monthly flows (shape of the hydrograph)	Mean	–	c
	Variability	–	b
Magnitude and duration of low flows	Mean	–	c
	Variability	–	–
Magnitude and duration of high flows	Mean	c	–
	Variability	–	c
Timing of extreme flow events	Mean	–	–
	Variability	a	–
Frequency and duration of high pulses	Mean	b	b
	Variability	–	–
Rate of change	Mean	–	–
	Variability	a	–

– none; [a] limited; [b] moderate; [c] high.

HESSD
11, 945–985, 2014
The influence of methodological procedures
F. J. Peñas et al.

Table A1. Hydrological indices used in the classification. Overall mean and standard deviation (referred in the manuscript by the prefix sd) of annual values for each index except for I1, I2, Ica, Icv, ikur, X5, X25, X75 and X95. I1 was not calculated for Normalized flow series.

Group	Name	Description
(1) Magnitude of annual and monthly flows	I1	Linear moment that represents the mean of the calculated flow duration curve
	I2	Linear moment that represents the variance of the calculated flow duration curve
	Ica	Linear moment that represents the skewness of the calculated flow duration curve
	Icv	Linear moment that represents the coefficient of variation of the calculated flow duration curve
	ikur	Linear moment that represents the kurtosis of the calculated flow duration curve
	M1-M12	Mean monthly flow. Standard deviation for each index was calculated.
	MxM1-MxM12	Maximum monthly flow
	MnM1- MnM12	Minimum monthly flow
(2) Magnitude and duration of annual extremes	1LF	Magnitude of minimum annual flow of 1 day duration. sd was also calculated
	7LF	Magnitude of minimum annual flow of 7 day duration.
	30LF	Magnitude of minimum annual flow of 30 day duration.
	90LF	Magnitude of minimum annual flow of 90 day duration.
	X75	Mean magnitude of flow exceeded 75 % of the time
	X95	Mean magnitude of flow exceeded 95 % of the time
	1HF	Magnitude of maxima annual flow of 1 day duration
	7HF	Magnitude of maxima annual flow of 7 day duration
	30HF	Magnitude of maxima annual flow of 30 day duration
	90HF	Magnitude of maxima annual flow of 90 day duration
	X25	Magnitude of the flows exceeded 25 % of the time. High flow pulses
	X5	Magnitude of the flows exceeded 5 % of the time.
	ZFD	Number of zero flow days
	BFI	Seven-day minimum flow divided by mean annual daily flows
(3) Timing of extreme flow events	JMin	Julian day of minimum flow
	JMax	Julian day of annual maximum flow
	Pred	Predictability (sensu Colwell, 1974)
(4) Frequency and duration of high pulses	FRE1	Number of high flow events per year using an upper threshold of 1 time median flow over all years
	FRE3	Number of high flow events per year using an upper threshold of 3 time median flow over all years
	FRE7	Number of high flow events per year using an upper threshold of 7 time median flow over all years
	nPHigh	Number of high pulses within each year
	dPHigh	Duration of high pulses within each year
(5) Rate of change	Pos	Mean of all positive differences between days
	nPos	Number of days with increasing flow
	Neg	Mean of all negative differences between days
	nNeg	Number of days with decreasing flow
	Rev	Number of hydrologic reversals

HESSD
11, 945–985, 2014

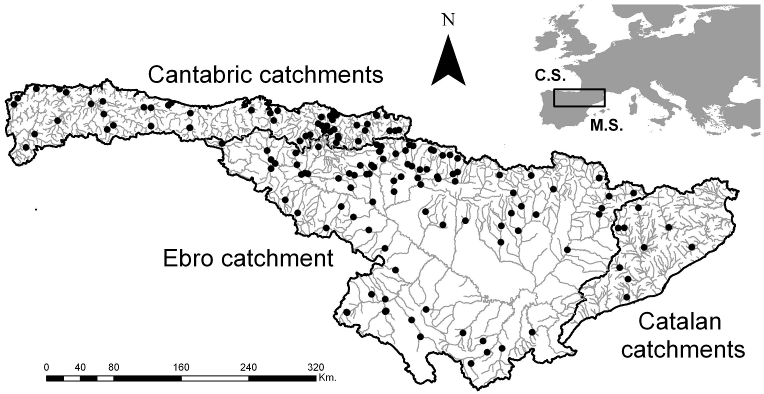

Fig. 1. Map of unregulated gauges (•; $n = 156$) in the study area. Black lines divide the Cantabric, the Ebro and the Catalan catchments. (CS: Cantabric sea; MS: Mediterranean sea).

HESSD
11, 945–985, 2014

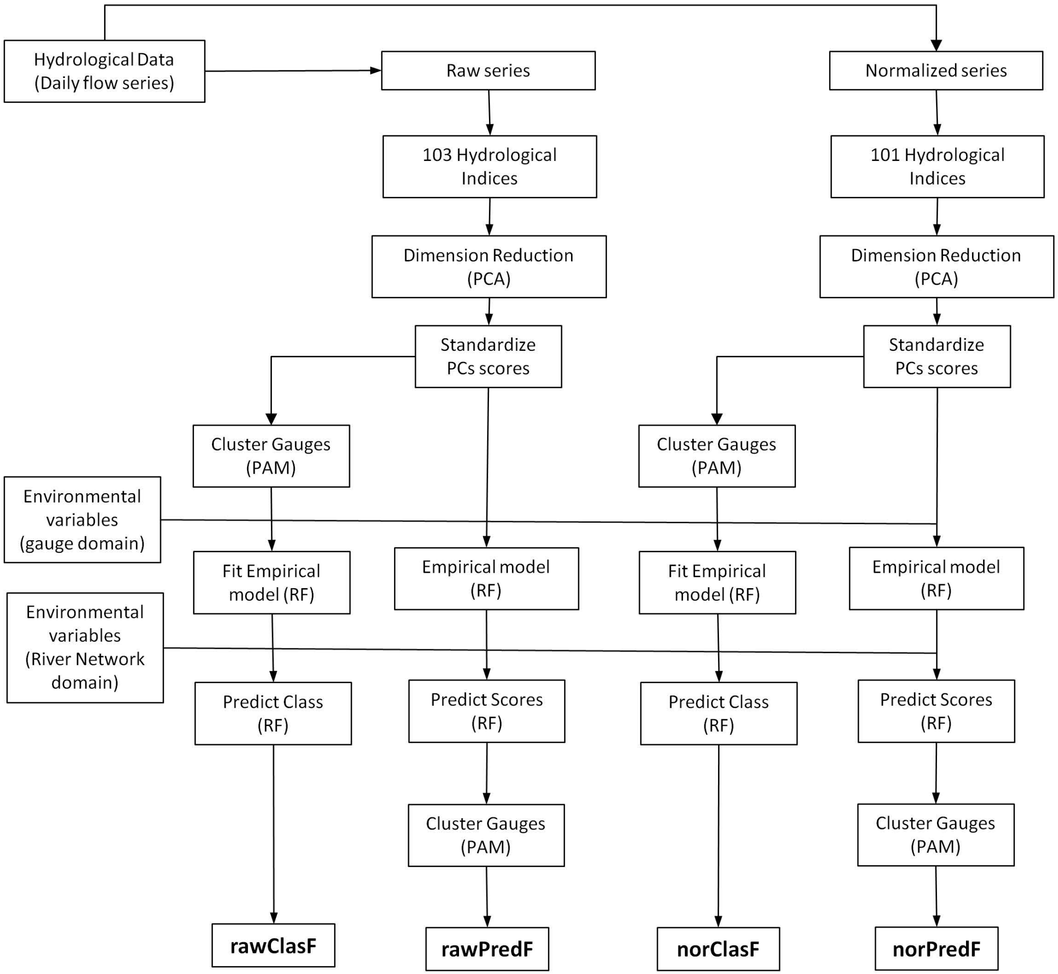

Fig. 2. Schematic diagram summarising the strategies applied to define the 4 classifications.

HESSD
11, 945–985, 2014

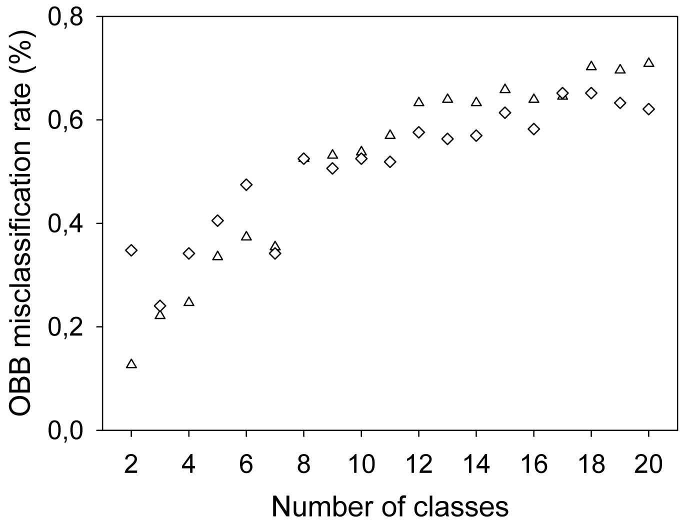

Fig. 3. Out-of-Bag misclassification rate of the random forest models developed for the 2 to 20-class level classification using classify then predict strategy based on the synthetic indices derived from the raw (∆; rawClasF) and the normalized flow series (⋄: norClasF).

HESSD
11, 945–985, 2014

Title Page
Abstract Introduction
Conclusions References
Tables Figures
◀ ▶
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

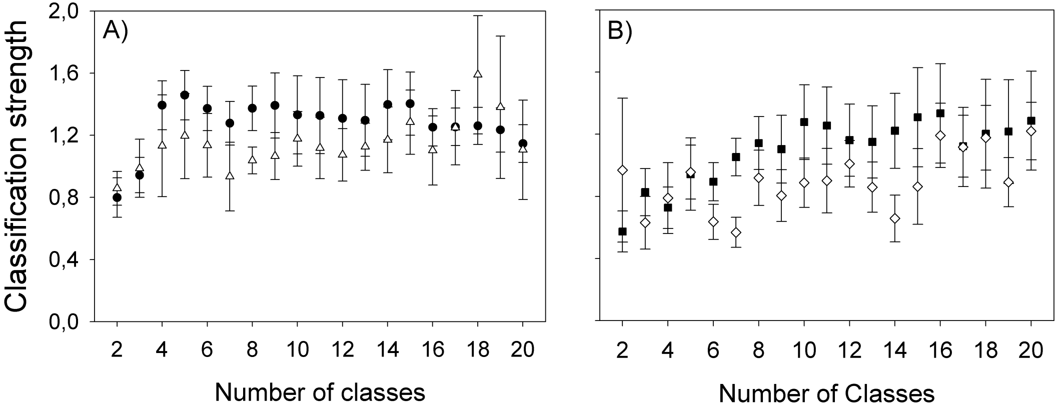

Fig. 4. Performance of the classifications based on the Classification Strength statistic **(A)** classifications based on raw flow series (•: rawPredF; △: rawClasF); **(B)** classifications based on normalized series (■: norPredF; ⋄: norClasF).

HESSD
11, 945–985, 2014

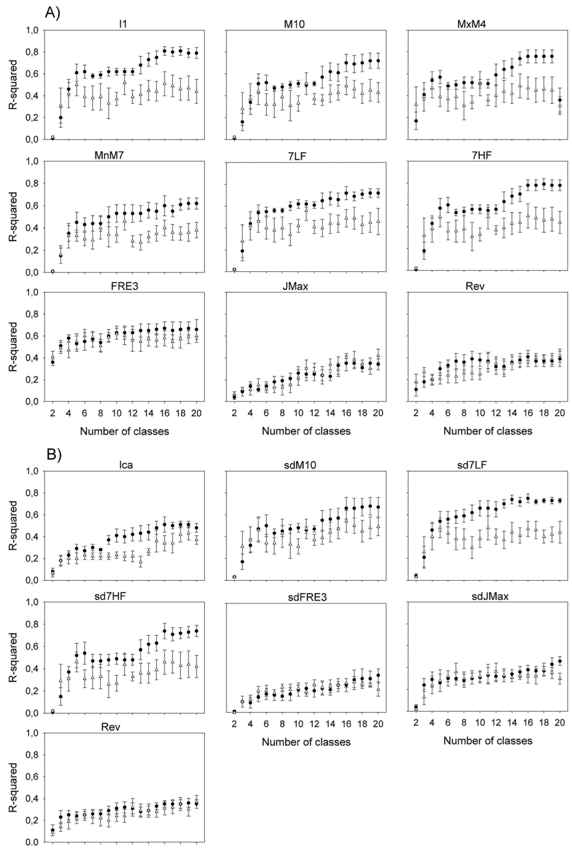

Fig. 5. Performance of the classifications derived from the raw flow series based on individual indices analysis. **(A)** Indices representing mean values. **(B)** Indices representing standard deviation. (•: rawPredF; ∆: rawClasF). We selected one index representing each aspect of the natural flow regime to illustrate the results (the values obtained for the 103 indices are included in Supplement, Table S1).

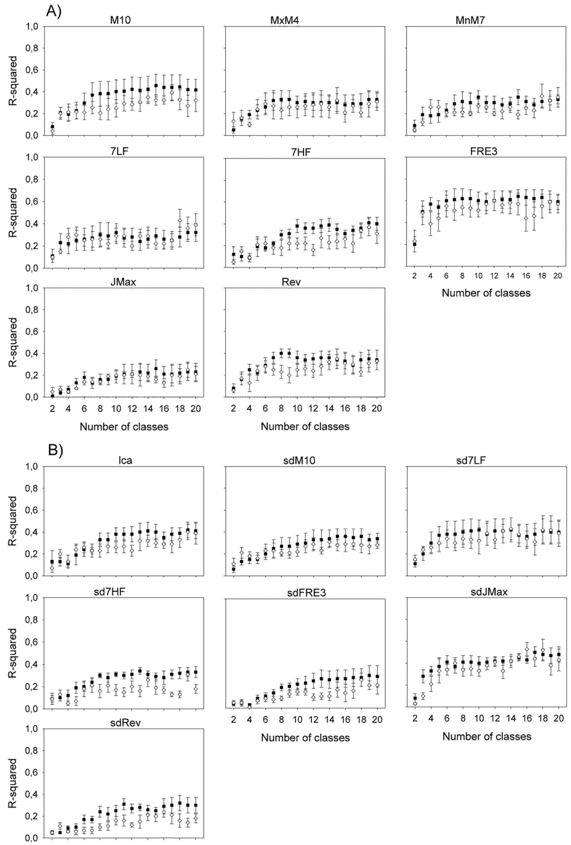

Fig. 6. Performance of the classifications derived from the normalized flow series based on individual indices analysis. **(A)** Indices representing mean values. **(B)** Indices representing standard deviation. (■ norPredF; ⋄ norClasF). We selected one index representing each aspect of the natural flow regime to illustrate results (the values obtained for the 101 indices are included in Supplement, Table S2).

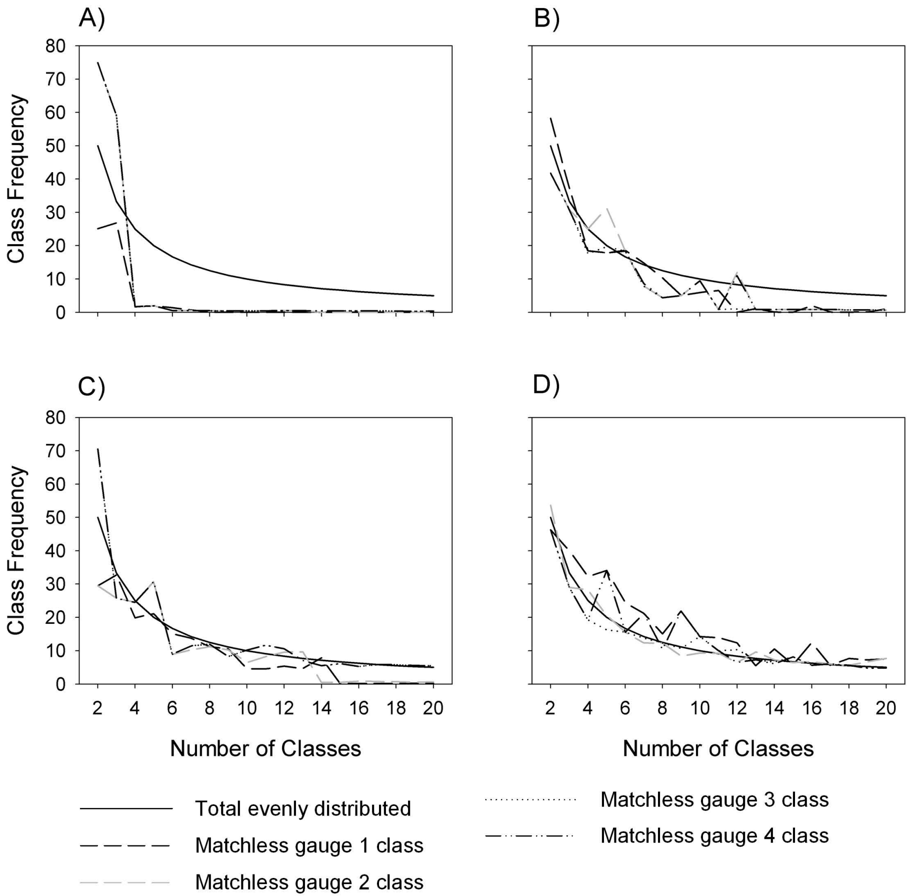

Fig. 7. Frequency (%) of the segments of the classification domain assigned to the classes where the distinctive gauges were included. (**A**: rawClasF; **B**: rawPredF; **C**: norClasF; **D**: norPredF).

HESSD
11, 945–985, 2014
The influence of methodological procedures
F. J. Peñas et al.