Hydrol. Earth Syst. Sci. Discuss., 8, 3233–3269, 2011
www.hydrol-earth-syst-sci-discuss.net/8/3233/2011/
doi:10.5194/hessd-8-3233-2011

Hydrology and Earth System Sciences Discussions

Comparison of catchment grouping methods for flow duration curve estimation at ungauged sites in France

E. Sauquet and C. Catalogne

Cemagref, UR HHLY, 3 bis quai Chauveau – CP 220, 69336 Lyon, France

Received: 18 March 2011 – Accepted: 23 March 2011 – Published: 1 April 2011

Correspondence to: E. Sauquet (eric.sauquet@cemagref.fr)

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

Abstract

The study aims at estimating flow duration curves (FDC) at ungauged sites in France and quantifying the associated uncertainties using a large dataset of 1080 FDCs. The interpolation procedure focuses here on 15 percentiles standardised by the mean annual flow, which is supposed to be known at each site. In particular, this paper discusses the relevance of different catchments grouping procedures on percentiles estimation by regional regression models.

First, five parsimonious FDC parametric models were tested to approximate FDCs at gauged sites. The results show that the model based on Empirical Orthogonal Functions (EOF) expansion outperforms the other ones. In this model each FDC is interpreted as a linear combination of regional amplitude functions with weights – the parameters of the model – varying in space. Here, only one amplitude function was found sufficient to fit well most of the observed curves. Thus the considered model requires only two parameters to be estimated at ungauged locations.

Second, homogeneous regions were derived according to hydrological response on one hand, and geological, climatic and topographic characteristics on the other hand. Hydrological similarity was assessed through two simple indicators: the concavity index (*IC*) that represents the shape of the standardized FDC and the seasonality ratio (*SR*) which is the ratio of summer and winter median flows. These variables were used as homogeneity criteria in three different methods for grouping catchments: (i) according to their membership in one of an a priori French classification into Hydro-Eco-Regions (HERs), (ii) by applying a regression tree clustering and (iii) by using hydrological neighbourhood obtained by canonical correlation analysis.

Finally, regression models between physiographic and/or climatic variables and the two parameters of the EOF model were derived considering all the data and thereafter for each group obtained through the tested grouping techniques. Results on percentiles estimation in cross validation show a significant benefit to form homogeneous regions

before developing regressions, particularly when grouping methods use hydrogeological information.

1 Introduction

A Flow Duration Curve (FDC) is the cumulative frequency distribution of observed flows during a period of interest (month, season, year, or entire period of record). It plots specified flows against their corresponding probability of exceedance that can be also interpreted as the percent of time these specified values are equalled or exceeded. FDC is a commonly used tool in water management applications, since it displays the full range of flows, including low flows and flood events (Vogel and Fennessey, 1995; Smakhtin, 2001). Here long-term flow duration curves were considered and derived from observed daily flows available at each site.

There have been numerous approaches for estimating FDC characteristics at ungauged locations, particularly low-flow percentiles, using regression equations under different climates (see Castellarin et al., 2007 for a recent review). Despite their interest for water management issues FDCs have until now received very little attention in France. The present study is to our knowledge the first attempt to develop regional flow duration models in this country. Previous works have concentred on mapping mean river flow statistics including long-term mean annual and monthly flows (Sauquet, 2006; Sauquet et al., 2008). These results cannot be ignored. A straightforward method for taking benefits from knowing the mean annual flow *qa* is to consider percentiles expressed as proportions of the long-term mean flow of the corresponding catchment as variables of interest. Regionalization can thus focus on the shape of the FDC. The dimensionless FDC and the mean annual flow *qa* are estimated separately and their combination provides the expected percentiles.

This approach, known as “index flow approach”, has been previously adopted by numerous authors (e.g., Holmes et al., 2002; Singh et al., 2001; Castellarin et al., 2004; Ganora et al., 2009) leading to various procedures to estimate normalised percentiles.

The simplest model assumes that the shapes of the FDC at all sites within the study area are identical. In practice, dimensionless FDCs from monitored catchments within the same region are pooled and averaged to create the representative shape. Since the hypothesis of similarity may be too restrictive, the alternative way has been chosen here: a reliable mathematical model with few parameters, which vary in space and are estimated at gauging stations, approximates the dimensionless FDC. The main advantages of the adopted approach are:

- It ensures consistency between river flow statistics (*qa* and percentiles) through the choice of the index value.
- It reduces the number of steps in the regionalisation procedure (only few parameters are involved in the procedures).
- It enables to distinguish the part related to the water balance (i.e. *qa*) from the characteristic response of the catchment to rainfall (i.e. the parameters of the shape of the dimensionless FDC) and thus to better identify the most important sources of spatial variability of FDC properties.

The last step of the procedure involved empirical relationships between the variables of interest and basin descriptors. Indeed this approach is by far the most often employed in regionalisation. In practice empirical formulas, usually established by multivariate regression, may perform poorly when applied at large scale due to high variability of hydrological behaviours, providing estimates with large errors. A way to improve the performance is to delineate homogeneous subregions assuming that pooled river catchments with similar hydrological, physiographical and meteorological characteristics will behave in a similar manner before developing separate regional regressions (Smakhtin, 2001).

The identification of homogeneous regions – both in theory and practice – has received much attention in hydrology, but no general methodology has emerged. Hence different ways to form homogeneous regions can be found in the literature, leading to

HESSD
8, 3233–3269, 2011

fixed geographically regions (either spatially contiguous or not) or hydrological neighbourhoods around each target site. Examples of contiguous regions defined for estimating regional FDCs are provided by Singh et al. (2001) in the Himalayan region of India based on a pre-existing partition into hydrometeorological subregions, and by Laaha and Blöschl (2006a) in Austria where grouping according to seasonality indices was tested. Geographically non-contiguous regions are usually identified using multivariate techniques such as multiple regression, principal component analysis or classification procedures, all of them incorporating catchment characteristics as well as flow statistics (e.g., Isik and Singh, 2008 in Turkey; Nathan and MacMahon, 1990 in Australia; Laaha and Blöschl, 2006a,b and Laaha et al., 2009 in Austria; Vezza et al., 2010 in Italy and Ganora et al., 2009 in northwestern Italy and Switzerland). In the neighbourhood approach, each site is supposed to have its own homogeneous region formed by gauging stations. Two main neighbourhood methods are commonly used. Both used auxiliary variables to define a hydrological space where distances are computed: the region of influence developed by Burn (1990a,b) (e.g., Holmes et al., 2002 in the UK) and the canonical correlation analysis (CCA) promoted by Ouarda et al. (2001).

Since the a priori efficiency of the grouping methods for regionalizing FDC characteristics is unknown, we here assess the relative performance of three of them: (i) contiguous regions obtained manually from expertise; (ii) regions obtained through Classification and Regression Trees algorithm (CART) and (iii) neighbourhood based on canonical correlation analysis (CCA). The choice of these methods was motivated (i) by a pre-existing partition established in France to answer some basic questions related to the European Water Framework directive, (ii) by published works demonstrating the potential of CART models in river flow regime regionalisation in France (Snelder et al., 2009) and (iii) by the wish to test a well-established method formerly developed to address issues in flood estimation.

In this paper we successively investigate two main issues related to the choice of the most adapted parametric model to fit observed dimensionless FDC at gauged sites

HESSD
8, 3233–3269, 2011

and the way to define homogeneous regions regardless of the interpolation procedure used to estimate FDC characteristics. Regarding the last point, this study is in line with previous benchmark studies on the performance of different grouping techniques for estimating low flow percentiles (Laaha and Blöschl, 2006b; Vezza et al., 2010). The paper is organised as follows. The study area and data used are first presented in Sect. 2. Hereafter, Sect. 3 compares the various mathematical models tested to approximate FDCs at gauged sites. Once the best performing parametric model has been identified, the variable on which homogeneity is tested are introduced in Sect. 4. Three approaches for delineating homogeneous regions are applied and compared (Sect. 5). The results of the fitted regional regressions are discussed in Sect. 6 and some conclusions including future research directions are drawn in the final section.

2 Study area and data

Climate and geology are quite diverse in France (area approx. 550 000 km^2): the northern and western parts of France are under maritime temperate climate influences whereas Mediterranean climate with hot and dry summer prevail in the south. In the latter areas, rainfall and evaporation drive the seasonal variations of runoff, in contrast to mountainous areas (high-altitude rivers in both the Pyrenees and the Alps) where snowmelt-fed regimes are observed. From a geological standpoint, France is roughly composed of two major geological formations: Hercynian crystalline impermeable substratum principally located in the north-western part of France (Brittany) and in mountainous areas (Alps, Pyrenees and Massif Central) and more or less permeable sedimentary rocks (limestone and clay) in flat plain areas (e.g., in the northern part of France where large aquifers sustain flows).

The dataset (Fig. 1) consists in 1080 gauging stations among more than 3500 stations that are available in the French database HYDRO (http://www.hydro.eaufrance.fr/). The following selection criteria were imposed to select these gauging stations: (i) no significant human influence on flow, (ii) record covering at least 18 years during

HESSD
8, 3233–3269, 2011

Comparison of catchment grouping methods

E. Sauquet and C. Catalogne

the period 1970–2008 (iii) high quality of measurements. This corresponds to an average density of about 2 gauging stations per 1000 km^2. The distribution of gauging stations across the country is however not uniform, with two notable areas of low station density located in the northern part of France and south Brittany. A total of 40% of the selected catchments have a record length varying between 35 and 45 years in most cases. Continuous observations during the period 1983–2000 are available for 90% of all selected stations, which ensures the temporal consistency of runoff statistics in terms of climatic variability. The drainage areas vary in size between 1.4 and 109 930 km^2. Most of the gauged catchments (44%) have areas from 100 to 500 km^2.

The catchment characteristics selected for use in the delineation of hydrological regions and in the development of regression equations were GIS-derived combining the SAFRAN high-resolution atmospheric reanalysis (Quintana-Seguí et al., 2008; Vidal et al., 2010), a 1-km grid digital elevation model and the associated drainage pattern (Sauquet, 2006). 18 catchment characteristics were selected for their possible influence on the shape of the standardised flow duration curve. The variables considered in this study include the drainage area, the coordinates of the centre of gravity, the mean catchment slope, the three quartiles of the hypsometric curve, the mean annual catchment air temperature, the mean summer catchment potential evapotranspiration using the formulation suggested by Oudin et al. (2005), the mean annual catchment actual evapotranspiration according to Turc formulation (1954), the mean annual catchment precipitation, the variance of the twelve mean monthly catchment precipitations, the mean seasonal precipitations, the fraction of the drainage catchment with impermeable substratum.

In addition, we used the Hydro-EcoRegion classification (HER) developed by Wasson et al. (2002). The HERs delineation was performed by experts incorporating different aspects of the geology, climate, physiography, drainage density, vegetation and topography of France. In particular, HER is the result of the interpretation in terms of erosion resistance, permeability, and hydrochemistry of a original geological map provided by the Bureau de Recherches Géologiques et Minières (B. R. G. M., 1996). The

HER was not specifically developed to discriminate river flow regimes. In the absence of quantitative information on hydrogeology, HERs were considered the most reliable surrogate. This classification divides France into 22 main regions (HER1) that are subdivided into 112 subregions (HER2). The dominant class in terms of fraction of the drainage catchment underlain by each HER was also computed.

3 A parametric model for flow duration curve

As suggested in Sect. 1, the identification of parsimonious models for summarize FDCs is advantageous to reduce the number of steps in the regionalisation procedure (only few parameters are required at ungauged sites to estimate dimensionless FDC).

Numerous formulas have been suggested to approximate FDCs (e.g., Quimpo et al., 1983; Franchini and Suppo, 1996; Yu et al., 2002; Castellarin et al., 2004; Li et al., 2010). Four parametric functions including the exponential model (Eq. 1), the logarithm model (Eq. 2), the power law model (Eq. 3) and the model suggested by Franchini and Suppo (Eq. 4) were tested on the dataset in this study. They approximate FDC at each site i, $i = 1, ..., N$:

$$Q_\mathrm{p}(i) = b(i)\, e^{a(i)\mathrm{p}} \tag{1}$$

$$Q_\mathrm{p}(i) = b(i) + a(i)\, \ln(p) \tag{2}$$

$$Q_\mathrm{p}(i) = b(i)\, p^{a(i)} \tag{3}$$

$$Q_\mathrm{p}(i) = b(i) + a(i)\,(1 - p)^{c(i)} \tag{4}$$

where Q_p is the p-th standardized flow percentiles and $a(i)$, $b(i)$ and $c(i)$ are the parameters at location i.

In addition to these four analytical functions, we tested a different approach based on the discrete decomposition into Empirical Orthogonal Functions expansion

HESSD
8, 3233–3269, 2011
Comparison of catchment grouping methods
E. Sauquet and C. Catalogne

(Holmström, 1963). This mathematical technique, also known as the Karhunen-Loeve transform, aims at extracting common patterns that represent a large fraction of the variability contained in a sample of N time series. EOF analysis has been already used for several purposes in hydrology (e.g., Hisdal and Tveito, 1991; Braud and Obled, 1991; Krasovskaia et al., 1999). In this application, EOF analysis expresses logarithmically transformed FDC as a linear combination of M shape functions β_i:

$$\ln\left(Q_\mathrm{p}(i)\right) = \overline{\ln\left(Q(i)\right)} + \sum_{i=1}^{M} \alpha_i(i)\,\beta_i(p), \quad i=1,\ldots,N, \quad p=1,\ldots,M \tag{5}$$

where N is the number of gauging stations, $\alpha_j(i)$, $i=1, \ldots, N$ are weights which vary with location and β_i are orthogonal functions with zero mean. Transforming the raw data has been adopted to avoid negative unrealistic estimates. The interest in applying this method is to keep the most part of the dataset variance in a limited number of shape functions. It is thus possible to truncate the series expansion to a subset of $L < M$ functions to limit the number of model parameters without significant loss of information.

In this study, all models were calibrated using 15 standardized percentiles Q_p, with respective exceedance probabilities $p = 1$, 2, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, 98 and 99% of the observed FDCs. Analytical models parameters were optimised on observations by applying ordinary least square procedures on logarithmically transformed data to reduce the influence of the largest observations. Prior to optimization, standardized percentiles equalled to zero were replaced by 0.001 to apply the logarithmic transformation.

The EOF decomposition applied on the dataset provides fourteen shape functions characterized by different patterns. The first shape functions, with a contribution of 97.2% to the total variance, represent the most common pattern of French FDCs. The other shape functions stand for a negligible part of the total variance and allow readjustment for very particular FDCs patterns. Considering these results, it was decided to keep only the first shape function. Thus the number of the parameters for the EOF

HESSD
8, 3233–3269, 2011

model is limited to two: the mean of the log-transformed standardized percentiles $\overline{\ln(Q)}$ and the weight associated with the first shape function α_1.

The performance/uncertainty of each model was measured by the deviations from the 15 standardized percentiles Q_p on which the five models are fitted. Unrealistic values (negative) were also replaced by 0.001. Boxplots in Fig. 2 give a graphical overview of the performance of each model. The median and the whiskers of the boxplots measure the bias and the accuracy of the model, respectively. In addition, the fitted curves are displayed on Fig. 3 for four gauged catchments representative of the diversity of FDC patterns within the reference dataset. Results show that:

- None of the models are perfect; in particular, all the models fails to reproduce correctly low-flow percentiles (relative errors may exceed 150% for some catchments). One should note that this criterion is very selective for low values (relative errors may reach large values when estimates are divided by a reference value close to zero).
- The biases appear most pronounced for the power law model (Eq. 3); low-flow percentiles as well as high-flow percentiles tend to be largely overestimated.
- Comparable biases are found for the exponential model (Eq. 1) and the Franchini and Suppo model (Eq. 4): standardized percentiles Q_p are underestimated for $p \leq 0.02$ and for $0.7 \leq p \leq 0.9$ whereas Q_p are overestimated for $p \geq 0.98$ and for $0.1 \leq p \leq 0.4$.
- The relative error range is smaller for the exponential model (Eq. 1) and the Franchini and Suppo model (Eq. 4) for the two standardized percentiles ($p = 0.01$, 0.02). However, there is a systematic negative bias in estimated high-flow standardized percentiles.
- Results for the logarithm model (Eq. 2) follow a very similar pattern to those for the EOF model (Eq. 5): on average, they both overestimate standardized percentiles with $0.4 \leq p \leq 0.8$ while high-flow and low-flow percentiles are underestimated.

HESSD
8, 3233–3269, 2011

The degree of bias differs substantially depending on the fitted model. The power law model (Eq. 3) provides the worst estimates in terms of relative error (bias and spread are the largest among the models). Comparable biases are found for the exponential model (Eq. 1) and the Franchini and Suppo model (Eq. 4). The EOF model (Eq. 5) appears to outperform among the models tested despite poor performance for high-flow percentiles. It performs nearly as well as the logarithm model (Eq. 2) but it also produces globally less biased estimates (median relative errors are the closest to zero and most of the interquartile ranges include zero for all the exceedance probabilities). The advantage of the EOF model is probably a better flexibility (the other models are not enough flexible to reproduce possible inflexion points in the observations) as it results from an empirical modelling of the shapes of the FDC. Considering these results the EOF model is the only one to be kept in the following steps.

4 Variables for testing hydrological homogeneity

The application of grouping methods is conditioned by the prior definition of variables to measure the degree of similarity between catchment behaviour and the level of homogeneity within the region. Several possible characteristics derived from river flow time series were tested and two variables were finally chosen for their correlation to the shape of the FDC and for their interpretation in terms of underlying hydrological processes.

The first variable is directly related to empirical properties observed on FDCs. The analysis of observed FDCs suggests that the 10-th percentile is a breakpoint delineating two parts of the curves: gradient tends to be higher in the upper branch ($10\% < p < 99\%$) than in the lower branch ($1\% < p < 10\%$). On this basis, a concavity index is computed as follows:

$$IC = \frac{Q_{10} - Q_{99}}{Q_1 - Q_{99}} \tag{6}$$

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

HESSD

8, 3233–3269, 2011

Comparison of catchment grouping methods

E. Sauquet and C. Catalogne

This descriptor is a measure of the contrast between low flow and high flow regime. A map of the concavity index in France including the location of the selected stations is presented in Fig. 4. The parameter takes values between 0 and 1. Values close to 1 are observed where large aquifers (e.g., in the northern part of France) and storages in snow pack (e.g., in the mountainous area) moderate the variability of daily flow. Values close to 0 are related to catchments exposed to contrasted climate (e.g., small catchments in the Mediterranean area experiencing hot and dry summers and intense short rainy events in autumn) and also to catchments with no storage capacity (e.g., on impermeable substratum) resulting in severe low-flows and quick runoff responses to rainfall events. It is worth noting that *IC* is well correlated with the parameters of the analytical FDC models and the average base flow index as well (not published here).

The second variable is a seasonality index. Laaha and Blöschl (2006a) demonstrated the value of such a variable for regionalizing the low-flow percentile Q_{95} in Austria. Indeed, grouping based on seasonality indices performed better than alternative groupings since these indices enable to discriminate well low flow processes at the regional scale when seasonal variability of runoff is high. Laaha and Blöschl (2006a) have used the ratio of the 95-th percentile of the winter (December to March) FDC divided by the 95-th percentile of the summer (April to November) FDC. Since our objective encompasses low flows, a Seasonality Ratio (*SR*) based on the medians was used here instead:

$$SR = Q_{50}(\text{summer})/Q_{50}(\text{winter}) \tag{7}$$

$SR \approx 1$ relates to catchments with nearly uniform flows through the year, often when significant groundwater contributions filter out seasonal climatic variability. Catchments influenced by snowmelt-fed processes display $SR < 1$ whereas for typical rainfall-fed catchments with low flow in summer and high flow in winter *SR* is above 1. *SR* is used here as a complement to *IC* to better identifying the causes of low seasonal variability in runoff (snow or groundwater storages). The variation in *SR* is governed by geology and air temperature and consequently in France by topographic influences.

These two variables *IC* and *SR* are the flow characteristics used to delineate homogeneous groups. Methods and results are presented in the subsequent section.

5 Grouping methods

5.1 Methods

5.1.1 Visual Grouping (VG)

Non-overlapping regions of approximately homogeneous low-flow indices *SR* and *IC* have first been identified visually. The starting point was the partition of France into 112 Hydro-EcoRegions (HER2s) at the finest level. These HER2s have been pooled based on hydrological expert knowledge.

The boundaries of HER2s have been first superimposed to the map displayed in Fig. 4. The most similar neighbouring HER2s have been progressively pooled by respecting contiguity, minimizing the dispersion within each cluster and maximizing the dissimilarity between the clusters based on visual inspection. The pooling process is far from obvious. In particular, due to the uneven density of the reference network, some of the HER2s contain too few stations to relate undoubtedly them to other neighbouring HER2s. Hence we used additional information such as rough description of hydrogeology to merge the ungauged HER2s with one of the adjacent clusters. Lastly, inspection of *SR* values led to a partition of the preliminary groups into sub-groups of HER2s, homogenous in terms of seasonality. Figure 5 presents the division of France into 18 different regions so obtained. Mixed regions may persist due to the heterogeneity at HER2 scale or due to the merging of HER2s containing a small number of gauged sites to large clusters. The identified regions include from 21 to 138 gauged sites and the average size is 57 (5% of the dataset).

5.1.2 Regression Tree (RT)

The aim of the analyses via tree-building algorithms is to predict dependent variables from a set of factor effects. Classification and Regression Trees approaches perform successive binary partitions of a given dataset according to decision variables. One advantage of this method is its ability to handle qualitative data (e.g., membership to a specific class). In general, RT leads to a set of if-then logical conditions as basis for classification. The algorithm identifies the best possible predictors, starting from the most discriminating factors and proceeding to the less important controls, to divide the clusters (*nodes*) into two successive parts. The optimal choices are determined recursively by increasing the homogeneity within the two resulting clusters. In this application the R software package *rpart* (Therneau and Atkinson, 2010) was used. The decision variables were selected automatically by the algorithm among the 19 catchment descriptors (i.e. including the dominant HER2) to ensure an optimal homogeneity of *IC* chosen as the dependent variable, in the successive clusters. The only constraint was to include at least 30 gauging stations in each region. At last 22 hydrological regions were identified with a mean number of 54 gauging stations per region (Fig. 6).

5.1.3 Canonical Correlation Analysis (CCA)

Canonical Correlation Analysis (Hotelling, 1936) is a multivariate statistical method suited to study interrelations between two sets of variables. CCA has been previously suggested by Ouarda et al. (2001) as a neighbourhood definition method. CCA provides two sets of canonical variables V_i, $i = 1, ..., p$ and W_i, $i = 1, ..., p$ obtained as follows:

- V_i, $i = 1, ..., p$ are linear combinations of k standardized hydrological variables X_i, $i = 1, ..., k$.

- W_i, $i = 1, ..., p$ are linear combinations of r standardized physiographic and climatic characteristics of the catchment Y_i, $i = 1, ..., r$.

HESSD
8, 3233–3269, 2011

– (V_i, W_i) have maximum correlation.

– (V_i, V_j), (V_i, W_j) and (W_i, W_j) $(i \neq j)$ are uncorrelated.

Canonical variables V_i, $i = 1, ..., p$ and W_i, $i = 1, ..., p$ can be interpreted as coordinates in hydrological and catchment-related physical spaces, respectively. Knowing Y_i, $i = 1, ..., r$ at ungauged location it is then possible to compute W_i, $i = 1, ..., p$ and through the calculation of correlation coefficients between canonical variables (V_i, W_j) their possible proximity – according to Mahanalobis distance – to the gauged stations in the hydrological space, which delineates neighbourhood around each site.

CCA has been formerly applied to regional flood frequency estimation (e.g., Ouarda et al., 2001; Chokmani and Ouarda, 2004; Shu and Ouarda, 2007). The present study is probably one of the first published works on CCA application to predict FDCs at ungauged locations. Here CCA was carried out between the two indicators *IC* and *SR* and all the catchment descriptors (excepted dominant HER2, since traditional CCA cannot manage qualitative variables). Geological description is thus reduced to the percentage of impervious areas. All combinations of 2 to 18 variables among the 18 remaining basin descriptors (listed in Sect. 2) were tested and at last we retained a combination of six characteristics which provides to the highest correlations between the first two pairs of canonical variables (V_i, W_i) $(p = 2)$. These catchment characteristics relate to location (the coordinates of the centre of gravity), climate (the mean annual catchment actual evapotranspiration and the variance of the twelve mean monthly catchment precipitations), geology (the fraction of the drainage catchment with impermeable substratum) and altitude (the third quartile of the hypsometric curve).

In addition to the variables involved in CCA, one should define the boundaries of the neighbourhood to exclude gauging stations too far from the target site. Ouarda et al. (2001) suggested a distance threshold depending on a given confidence level and on target site. Preliminary tests showed the difficulty to define a satisfactory confidence level for our dataset, in particular for very atypical sites for which too few similar sites are selected to derive, thereafter, reliable regional regressions. Consequently we chose

HESSD
8, 3233–3269, 2011

here to fix the number of stations contributing to neighbourhood to 50, i.e. the 50 closest gauging stations to the target site, to allow objective comparisons with the results of the two other grouping methods.

5.2 Results

Figures 5 and 6 present maps obtained by VG and RT, respectively. One colour is assigned to each reach of the main river network (i.e. all locations draining more than 50 km^2). Displaying results from CCA on a map is not feasible since each site has its own neighbourhood. The comparison between the two maps suggests that:

- The two procedures based on the same auxiliary variables lead to different divisions. The spatial pattern provided by RT is patchier than the one obtained by VG: small tributaries may belong to different classes than the main stem they flow into. The relative influence of the location is naturally moderate on class allocation since mountainous basins in the Alps and Pyrenees are pooled together. This result is in direct line with conclusions of previous studies dedicated to flood quantile estimation (Merz and Blöschl, 2005; Ouarda et al., 2001) that concluded that geographical proximity does not involve hydrological similarity.

- Common geographical groupings can be found e.g., in the north part of France (in brown in Fig. 5 and in cyan in Fig. 6) and in the west part of France (in orange in Fig. 5 and in dark blue in Fig. 6), supporting visually the fact that the two partitions are not totally inconsistent.

To supplement this analysis, we examined the empirical distributions of both *SR* and *IC* per regions (identified by a letter on the x-axis). Box plots are presented in Fig. 7. There is no obvious difference between the spread of *SR* and *IC*. The absence of significant improvement in terms of homogeneity within each group (regarding the interquartile provided by the empirical distribution of each variable) and in discrimination between groups (regarding the differences between the medians of each groups for

HESSD
8, 3233–3269, 2011

Comparison of catchment grouping methods

E. Sauquet and C. Catalogne

each variable) is due to the valuable information contained in the Hydro-EcoRegions. Both methods lead to two very distinct regions with high values for *IC*. As a proof the membership to clusters of HER is chosen as the first splitting variables (Fig. 8).

Regarding CCA we decided to compare results with published works in terms of correlation structure. Figure 9 indicates weak correlations between the canonical variables: $r_1 = 0.71$ between W_1 and V_1 and $r_2 = 0.57$ between W_2 and V_2. As comparison, for flood quantile estimation, Ouarda et al. (2001) obtained r_1 between 0.959 and 0.960 and r_2 between 0.279 and 0.422 in an application in the Province of Ontario (Canada), Haché et al. (2002) obtained $r_1 = 0.986$ et $r_2 = 0.842$ in the Saint-Maurice river region (Canada) and Ouarda et al. (2008) obtained $r_1 = 0.966$ and $r_2 = 0.247$ in Mexico. These studies used at least one T-year flood quantile QT expressed in $m^3\,s^{-1}$ as one of the hydrological variables and the drainage area A as one of the physiographical variables. Since catchment area is certainly the factor with the greatest influence on flood magnitude above climate, geology and land-use as one of the physiographical variables, CCA suggests automatically a first pair of canonical variables (V_1, W_1) highly correlated with QT and A, respectively. Roughly speaking, the presence of a strong link between one hydrological variable and one physiographical variable ensures at least one highly correlated pair of canonical variables. This is not the case here: the hydrological variables are two ratios free from scale effect and so A was excluded from the final variables involved in the definition of canonical variables and the highest coefficient of determination observed between *SR* and the first quartile of the hypsometric curve is just 0.34.

All groupings were tested as bases for developing regional regression. The next section presents briefly the method and their relative performance.

HESSD
8, 3233–3269, 2011

6 Regional regression

6.1 Method

The homogeneous regions are now identified. Multivariate regression model relations between the EOF model parameters and catchment descriptors can be developed. Both linear and power form models dependences were investigated:

$$\alpha_1 = \lambda_0 + \sum_{i \in [1;18]} \lambda_i Y_i \tag{8}$$

$$\overline{\ln (Q(i))} = \lambda'_0 + \sum_{i \in [1;18]} \lambda'_i Y_i \tag{9}$$

$$\alpha_1 = \lambda_0 \prod_{i \in [1;18]} Y_i^{\lambda_i} \tag{10}$$

$$\overline{\ln (Q(i))} = \lambda'_0 \prod_{i \in [1;18]} Y_i^{\lambda'_i} \tag{11}$$

Models were adjusted on observations to each homogeneous group by the ordinary least squares method (using log transformed data to fit power-form models).

In order to define the most appropriate model for each region, all combinations including one to four variables among the 18 quantitative variables were tested and the 10 best regression models in terms of adjusted coefficient of determination were retained. These models were then refined/filtered through an interactive scheme: (i) outliers using Cook's distance were removed first, (ii) the statistical properties of residuals (including normality and homoscedasticity) were checked by visual inspection (only for the first two grouping methods) and (iii) the robustness of each empirical formula was finally assessed by leave-one-out cross-validation. The final models were selected regarding to the best value of the coefficient of determination obtain by leave-one-out cross-validation.

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

HESSD

8, 3233–3269, 2011

6.2 Results

To measure the values of prior region delineation a global regression using the whole available gauging stations dataset and the procedure described in Sect. 4.5 was derived.

The predictive performance of each method is assessed using the root mean square error (RMSE) and the coefficient of determination of the regression R^2 between observed and predicted values for the EOF model parameters, $\overline{\ln(Q)}$ and α_1. In addition to these statistics, scatter plots were drawn and inspected visually to compare the spread of the predictions. These results are reported in the next four figures (from Figs. 10 to 13). The two upper panels plot estimated values against observed ones ($\overline{\ln(Q)}$ on the left and α_1 on the right). Each point is related to one gauging station. A one-to-one line (in red) is added to each graph. Absolute relative errors were also computed for each of the 15 selected standardized percentiles Qp and their empirical statistical distributions were summarised by box plots displayed on the lower panel.

The cross validation results for the global regression are presented in Fig. 10. As expected the scores are unsatisfactory: dispersion is high around the one-to-one line ($R^2 < 0.20$ for both EOF model parameters) and the low-flow percentiles were poorly predicted. By comparison, the three next figures (Figs. 11 to 13) illustrate the performance of the three tested methods and suggest that:

- The regional regression based on the three grouping approaches is superior to global regression like in Laaha and Blöschl (2006b) and in Vezza et al. (2010); results for all models follow a similar pattern in terms of relative error on standardized percentiles: the highest errors are obtained for the lowest values.

- RT is the best regionalisation method, and VG performs nearly as well with comparable R^2; however one should note that the estimations by VG approach are probably heteroscedastic (the spread of errors increases along with α_1).

HESSD
8, 3233–3269, 2011

Comparison of catchment grouping methods

E. Sauquet and C. Catalogne

– CCA outperforms only slightly global regression. This finding is astonishing since CCA is known as a very efficient regional estimation method.

To understand the unexpected performance for CCA, we performed additional computations and compared the neighbourhoods defined by CCA to the expected ones, ideally defined in the hydrological space. For the 50 closest gauging stations, the percentages of concordance between the neighbourhoods predicted by CCA and the theoretical ones are weak and may even equal to zero. Therefore the selected neighbours were almost never the expected ones. Besides one should note that estimates from regional regressions obtained with the expected neighbourhoods showed that performance is far better.

Hence the correlation between canonical variables is not strong enough to guarantee the correspondence between the physiographical and hydrological spaces and thus to ensure the efficiency of CCA. It probably points out the lack of efficient catchment characteristics to strengthen the link between the two spaces – certainly characteristics linked to hydrogeology since the application of the two other methods differs only by the introduction of such a variable (i.e. dominant HER2).

7 Conclusions

In this study, a regionalisation method is suggested to estimate flow duration characteristics. The developed approach supposes that the mean annual flow qa is known before estimating FDCs at ungauged sites. Efforts have been therefore concentrated on the estimation of the shape of the normalised FDC using a large data set of FDC derived from 1080 gauging stations.

First, a parametric and parcimonious model based on EOF decomposition has been developed to fit the observed shapes of the FDC. A comparison to other models referenced in the literature demonstrates that the EOF model leads to the best estimates at gauging stations. A reason could be that, conversely to the empirical approach, analytical formulas are not flexible enough to accommodate the full range of observed shapes.

Thus it would be unrealistic to support the idea of one parametric model adapted to all the hydrological conditions.

In a second step, different grouping techniques for identifying homogeneous regions and developing separate regression models have been compared. Two of the grouping procedures, VG and RT, with comparable performance, demonstrate the significant gain to develop regional regressions. One should note that the RT classification procedure has the advantage to be automatic and objective whereas heterogeneity may persist in the VG groups that could explain its ranking (2nd). Nevertheless a large portion of the variance remains unexplained. Further effort could be devoted to the interpolation of the residuals. One could apply techniques such as adapted kriging (Sauquet, 2006), Top-Kriging (Skøien et al., 2006) or physiographical space based interpolation (Castiglioni et al., 2009) for this purpose.

Despite a greatest flexibility in neighbourhood selection, i.e. a neighbourhood is defined individually for each target site, the third and last grouping method, CCA, performed poorly. These bad and unexpected scores for CCA may result from the difficulty to obtain a sufficient correlation link between hydrological and physiographical spaces in the absence of relevant characteristics to describe the hydrogeological properties within the catchments. Indeed, for the other two grouping techniques hydrogeology is summarized by one qualitative variable, i.e. the class of the dominant HER2, which provides sufficient information to increase homogeneity within regions and to ensure more efficient regional regressions. As a result the application of CCA in predefined regions with homogeneous hydrogeological properties should be investigated to compare equitably CCA to other methods on the same bases.

Acknowledgements. The authors wish to thank the French National Agency for Water and Aquatic Environments (ONEMA) for partially funding the present work.

HESSD

8, 3233–3269, 2011

Comparison of catchment grouping methods

E. Sauquet and C. Catalogne

References

Braud, I. and Obled, C.: On the use of Empirical Orthogonal Function (EOF) analysis in the simulation of random fields, Stoch. Hydrol. Hydraul., 5, 125–134, 1991.

B. R. G. M.: Carte géologique au 1/1.000.000^{eme}, 1996.

Burn, D. H.: An appraisal of the "region of influence" approach to flood frequency analysis, Hydrolog. Sci. J., 35, 149–165, 1990a.

Burn, D. H.: Evaluation of regional flood frequency analysis with a region of influence approach, Water Resour. Res., 26, 2257–2265, 1990b.

Castellarin, A., Galeati, G., Brandimarte, L., Montanari, A., and Brath, A.: Regional flow-duration curves: reliability for ungauged basins, Adv. Water Resour., 27, 953–965, 2004.

Castellarin, A., Camorani, G., and Brath, A.: Predicting annual and long-term flow-duration curves in ungauged basins, Adv. Water Resour., 30, 937–953, 2007.

Castiglioni, S., Castellarin, A., and Montanari, A.: Prediction of low-flow indices in ungauged basins through physiographical space-based interpolation, J. Hydrol., 378(3–4), 272–280, 2009.

Chokmani, K. and Ouarda, T. B. M. J.: Physiographical space-based kriging for regional flood frequency estimation at ungauged sites, Water Resour. Res., 40, W12514, doi:10.1029/2003WR002983, 2004.

Franchini, M. and Suppo, M.: Regional analysis of flow duration curves for a limestone region, Water Resour. Manage., 10, 199–218, 1996.

Ganora, D., Claps, P., Laio, F., and Viglione, A.: An approach to estimate nonparametric flow duration curves in ungauged basins, Water Resour. Res., 45, W10418, doi:10.1029/2008WR007472, 2009.

Haché, M., Ouarda, T. B. M. J., Bruneau, P., and Bobee, B.: Regional estimation by canonical correlation analysis: Comparison of types of hydrological variables, Can. J. Civ. Eng., 29(6), 899–910, 2002.

Hisdal, H. and Tveito, O. E.: Generation of runoff series at ungauged locations using Empirical Orthogonal Functions in combination with kriging, Stoch. Hydrol. Hydraul., 6, 255–269, 1991.

Holmes, M. G. R., Young, A. R., Gustard, A., and Grew, R.: A region of influence approach to predicting flow duration curves within ungauged catchments, Hydrol. Earth Syst. Sci., 6, 721–731, doi:10.5194/hess-6-721-2002, 2002.

Holmström, I.: On a method for parametric representation of the state of the atmosphere, Tellus, 15, 127–149, 1963.
Hotelling, H.: Relations between two sets of variates, Biometrika, 28, 321–377, 1936.
Isik, S. and Singh, V. P.: Hydrologic regionalization of watersheds in Turkey, J. Hydrol. Eng., 13(9), 824–834, 2008.
Krasovskaia, I., Gottschalk, L., and Kundzewicz, Z. W.: Dimensionality of Scandinavian river flow regimes, Hydrolog. Sci. J., 44(5), 705–723, 1999.
Laaha, G. and Blöschl, G.: Seasonality indices for regionalizing low flows, Hydrol. Process., 20, 3851–3878, 2006a.
Laaha, G. and Blöschl, G.: A comparison of low flow regionalisation methods – catchment grouping, J. Hydrol., 323(1–4), 193–214, 2006b.
Li, M., Shao, Q., Zhang, L., and Chiew, F. H. S.: A new regionalization approach and its application to predict flow duration curve in ungauged basins, J. Hydrol., 389(1–2), 137–145, 2010.
Merz, R. and Blöschl, G: Flood frequency regionalisation – spatial proximity vs catchment attributes, J. Hydrol., 302(1–4), 283–306, 2005.
Nathan, R. J. and McMahon, T. A.: Identification of homogeneous regions for the purpose of regionalization, J. Hydrol., 121, 217–238, 1990.
Ouarda, T. B. M. J., Girard, C., Cavadias, G. S., and Bobée, B.: Regional flood frequency estimation with canonical correlation analysis, J. Hydrol., 254(1–4), 157–173, 2001.
Ouarda, T. B. M. J., Bâ, K. M., Diaz-Delgado, C., Cârsteanu, A., Chokmani, K., Gingras, H., Quentin, E., Trujillo, E., and Bobée, B.: Intercomparison of regional flood frequency estimation methods at ungauged sites for a Mexican case study, J. Hydrol., 348(1–2), 40–58, 2008.
Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andreassian, V., Anctil, F., and Loumagne, C.: Which potential evapotranspiration input for a lumped rainfall-runoff model? Part 2 Towards a simple and efficient potential evapotranspiration model for rainfallrunoff modelling, J. Hydrol., 303, 290–306, 2005.
Quintana-Seguí, P., Le Moigne, P., Durand, Y., Martin, E., Habets, F., Baillon, M., Canellas, C., Franchisteguy, L., and Morel, S.: Analysis of near surface atmospheric variables: Validation of the SAFRAN analysis over France, J. Appl. Meteorol. Clim., 47, 92–107, 2008.
Quimpo, R. G., Alejandrino, A. A., and McNally, T. A.: Regionalised flow duration curves for Philippines, J. Water Resour. Pl.-ASCE, 109(4), 320-330, 1983.

HESSD
8, 3233–3269, 2011

Comparison of catchment grouping methods

E. Sauquet and C. Catalogne

Sauquet, E.: Mapping mean annual river flows: geostatistical developments for incorporating river network dependencies, J. Hydrol., 331(1–2), 300–314, 2006.
Sauquet, E., Gottschalk, L., and Krasovskaia, I.: Estimating mean monthly runoff at ungauged locations: an application to France, Hydrol. Res., 39(5–6), 403–423, 2008.
Shu, C. and Ouarda, T. B. M. J.: Flood frequency analysis at ungauged sites using artificial neural networks in canonical correlation analysis physiographic space, Water Resour. Res., 43, W07438, doi:10.1029/2006WR005142, 2007.
Singh, R. D., Mishra, S. K., and Chowdhary, H.: Regional flowduration models for large number of ungauged Himalayan catchments for planning microhydro projects, J. Hydrol. Eng., 6(4), 310-316, 2001.
Skøien, J. O., Merz, R., and Blöschl, G.: Top-kriging – geostatistics on stream networks, Hydrol. Earth Syst. Sci., 10, 277–287, doi:10.5194/hess-10-277-2006, 2006.
Smakhtin, V. U.: Low flow hydrology: a review, J. Hydrol., 240, 147–186, 2001.
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, 2009.
Therneau, T. M. and Atkinson, B.: rpart: Recursive Partitioning, R package, http://cran.r-project.org/web/packages/rpart/ (last access: March 2011), 2010.
Turc, L.: Water balance of soil: relation between precipitation, evapotranspiration and runoff, Ann. Agronom., 5, 491–595, 1954.
Vezza, P., Comoglio, C., Rosso, M., and Viglione, A.: Low Flows Regionalization in North-Western Italy, Water Resour. Manage., 24, 4049–4074, 2010.
Vidal, J. P., Martin, E., Baillon, M., Franchistéguy, L., and Soubeyroux, J. M.: A 50-year high-resolution atmospheric reanalysis over France with the Safran system, Int. J. Climatol., 30(10), 1627–1644, doi:10.1002/joc.2003, 2010.
Vogel, R. M. and Fennessey, N. M.: Flow duration curves II: a review of applications in water resources planning, Water Resour. Bull., 31(6), 1029–1039, 1995.
Wasson, J. G., Chandesris, A., Pella, H., and Blanc, L.: Typology and reference conditions for surface water bodies in France: the hydro-ecoregion approach, TemaNord, 566, 37–41, 2002.
Yu, P. S., Yang, T. C., and Wang, Y. C.: Uncertainty analysis of regional flow duration curves, J. Water Res. Pl.-ASCE, 128(6), 424–430, 2002.

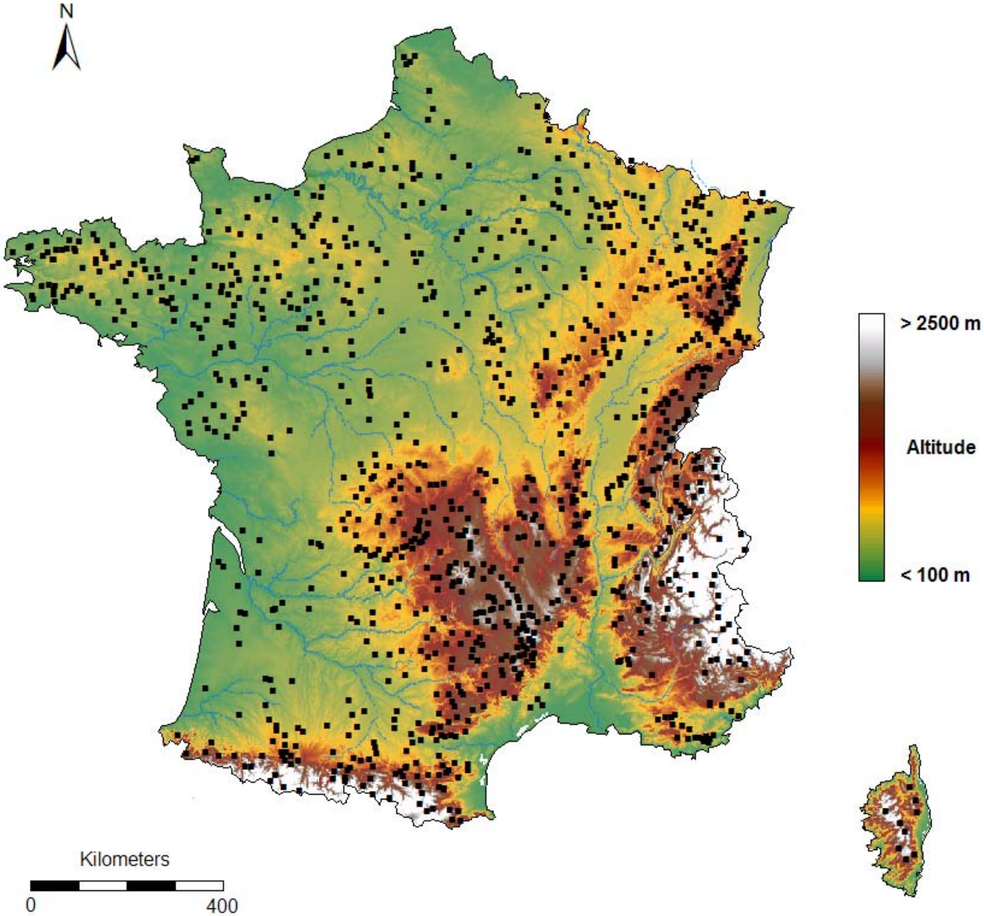

Fig. 1. Study area and gauging stations identified by their respective centre of gravity (black square).

HESSD
8, 3233–3269, 2011

Comparison of catchment grouping methods

E. Sauquet and C. Catalogne

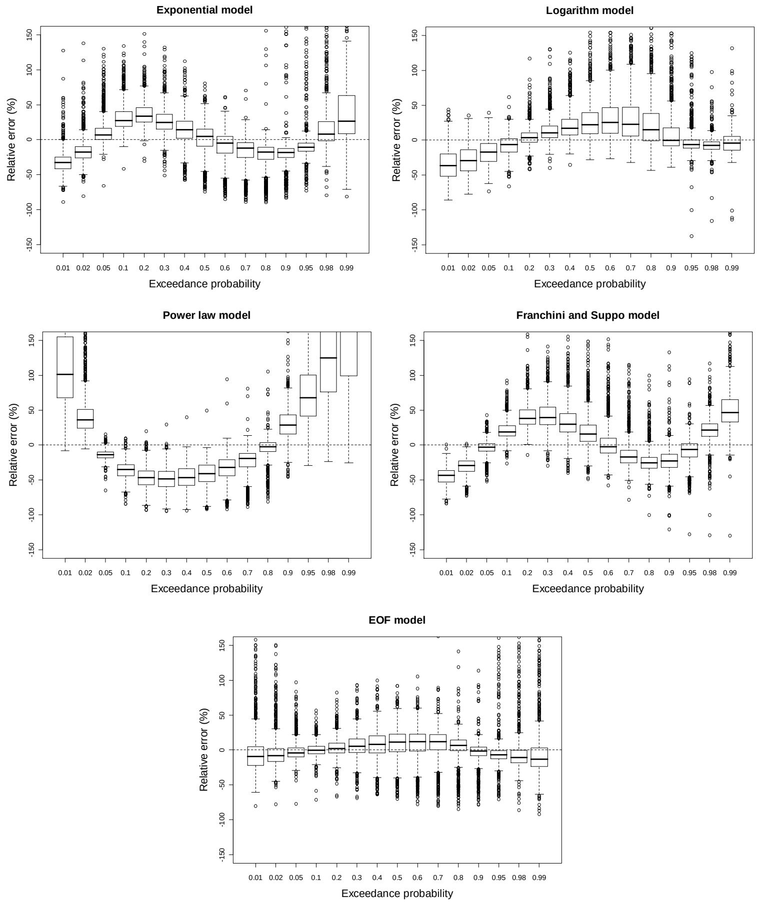

Fig. 2. Empirical distribution of the relative error for each percentile and each model. The boxplots are defined by the first quartile, the median and the third quartile. The whiskers extend to 1.5 the interquartile range; open circles indicate outliers.

HESSD
8, 3233–3269, 2011

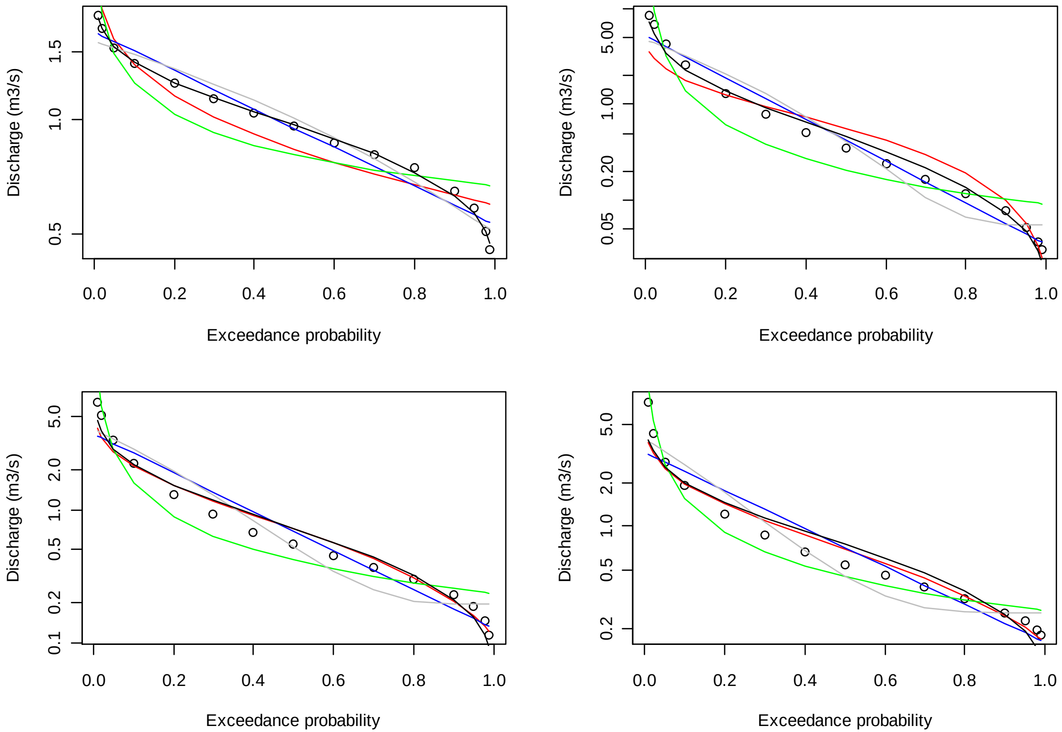

Fig. 3. Comparison of observed (open circle) and modeled flow duration curves – logarithm (red), exponential (blue), power law (green), Franchini and Suppo (grey), EOF (black).

HESSD
8, 3233–3269, 2011

Comparison of catchment grouping methods

E. Sauquet and C. Catalogne

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

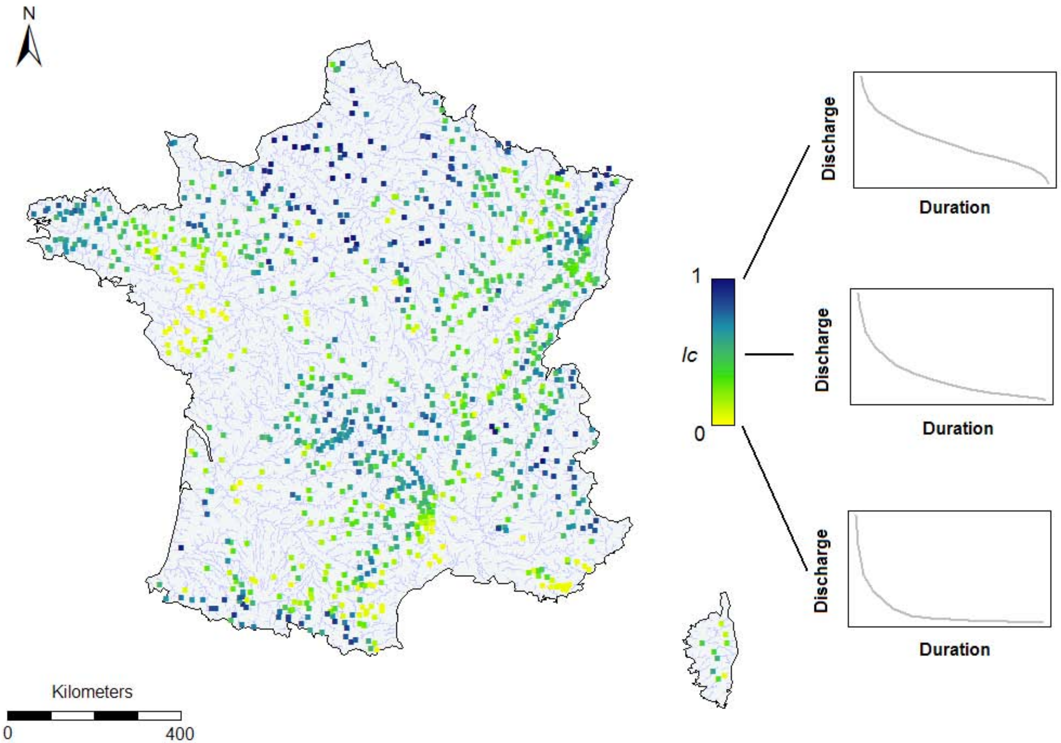

Fig. 4. Spatial distribution of the concavity index *IC* observed at gauged catchments identified by the location of their centre of gravity.

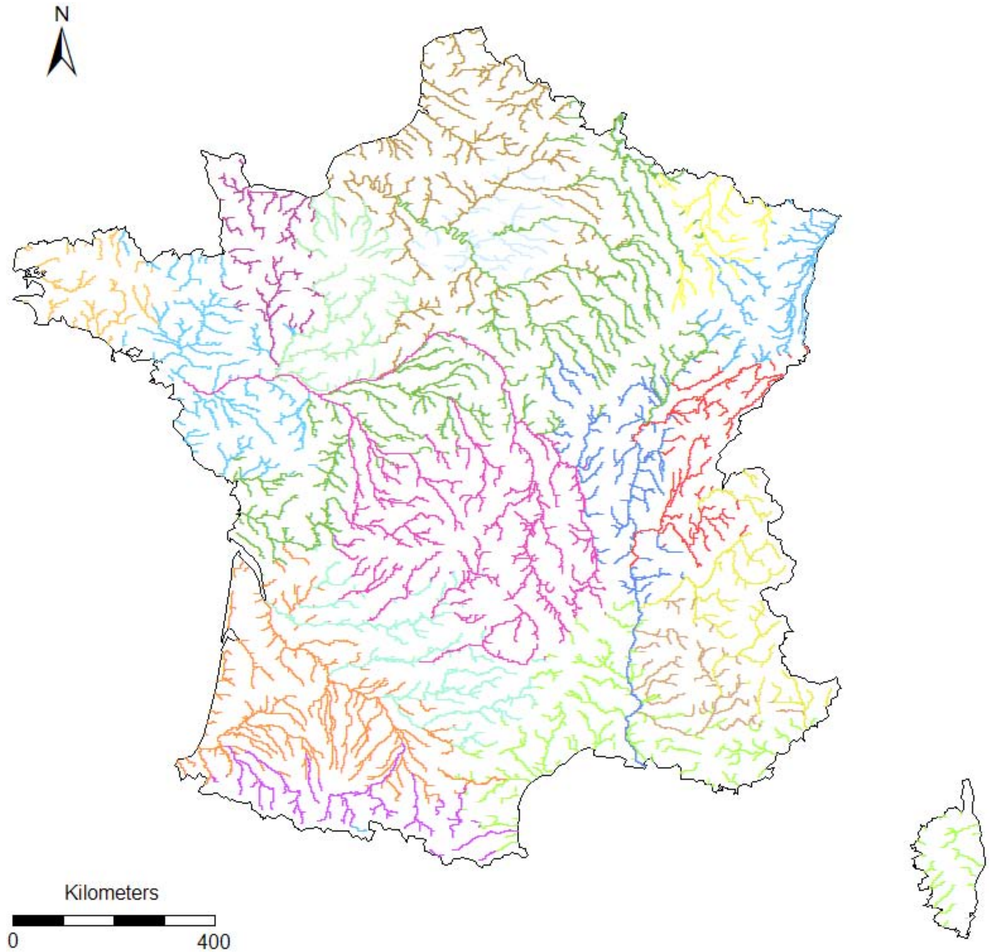

Fig. 5. Results of classification based on visual grouping (VG).

HESSD

8, 3233–3269, 2011

Comparison of catchment grouping methods

E. Sauquet and C. Catalogne

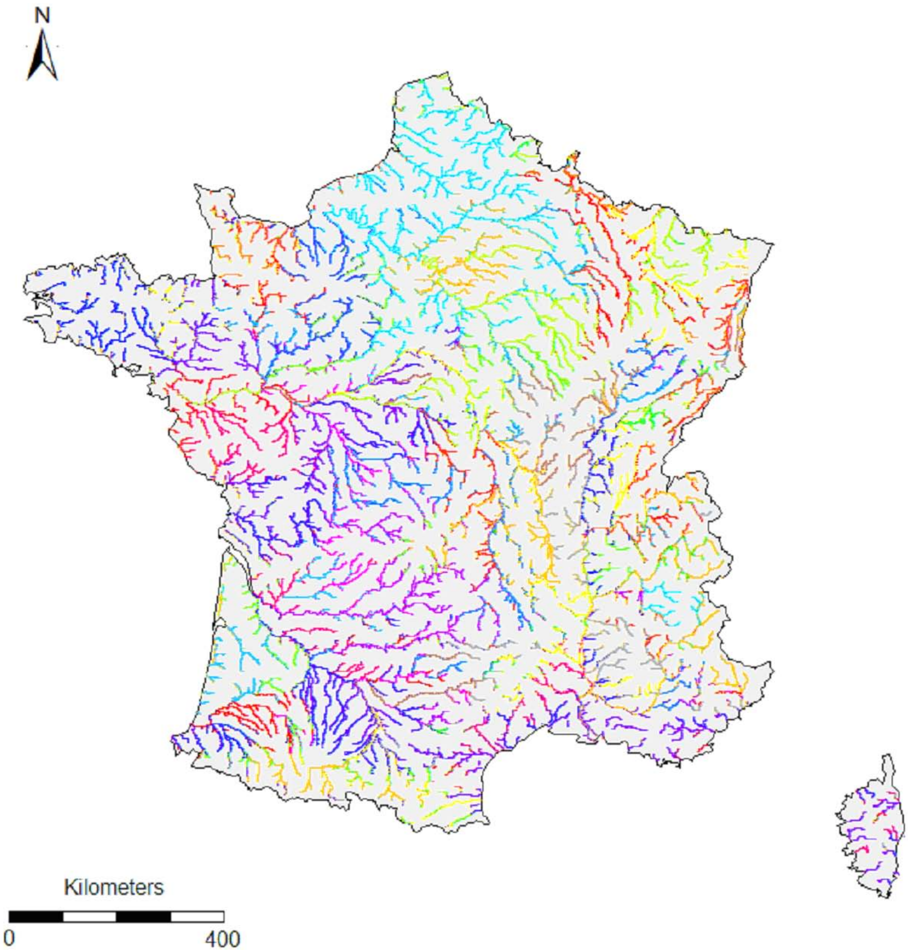

Fig. 6. Results of classification based on regression trees (RT).

HESSD

8, 3233–3269, 2011

Comparison of catchment grouping methods

E. Sauquet and C. Catalogne

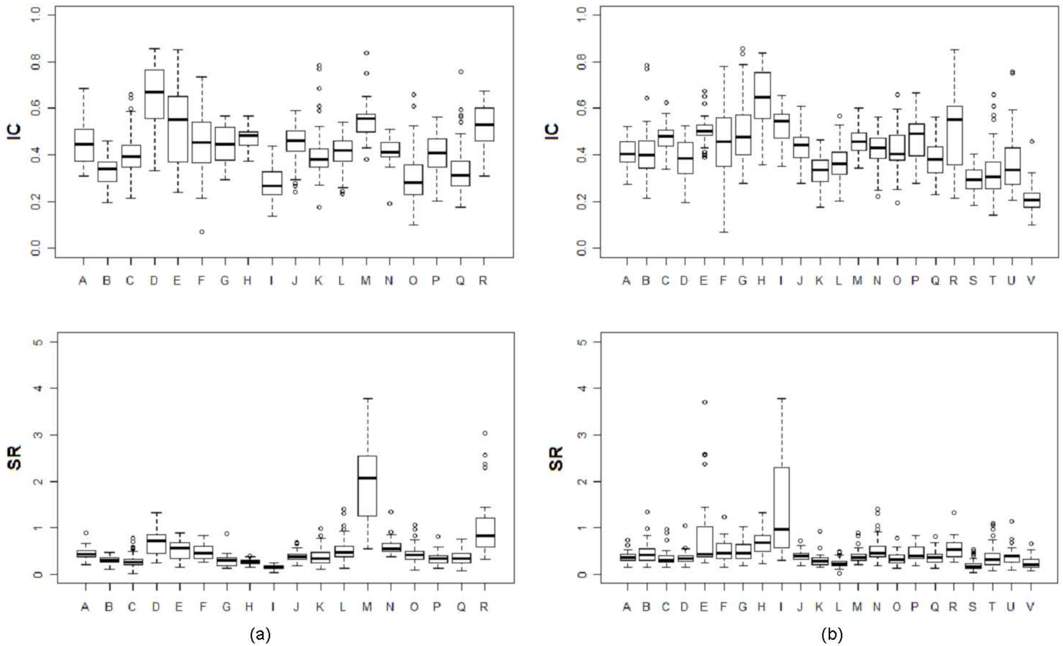

Fig. 7. Empirical distributions of the two hydrological indicators for each cluster according to VG **(a)** and RT **(b)**.

HESSD

8, 3233–3269, 2011

Comparison of catchment grouping methods

E. Sauquet and C. Catalogne

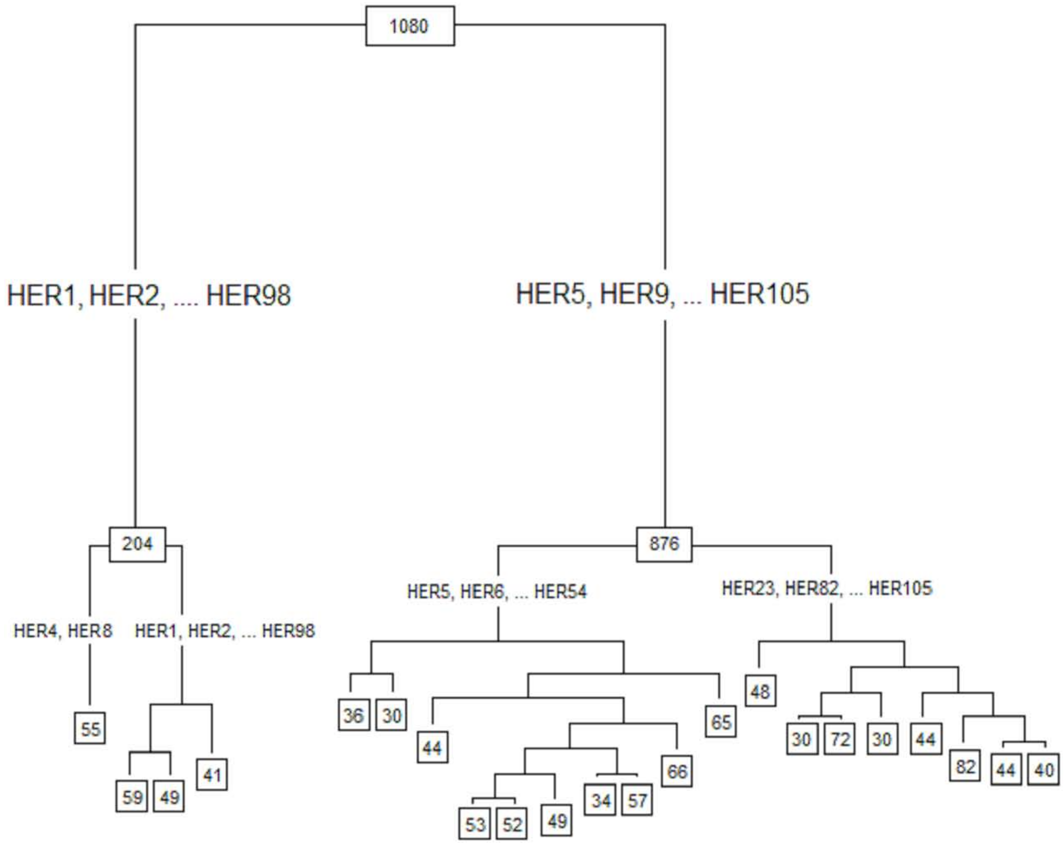

Fig. 8. Regression tree model (the numbers at each node of the tree and the name of the first splitting variables are reported in the boxes and in the middle of the branches, respectively).

HESSD

8, 3233–3269, 2011

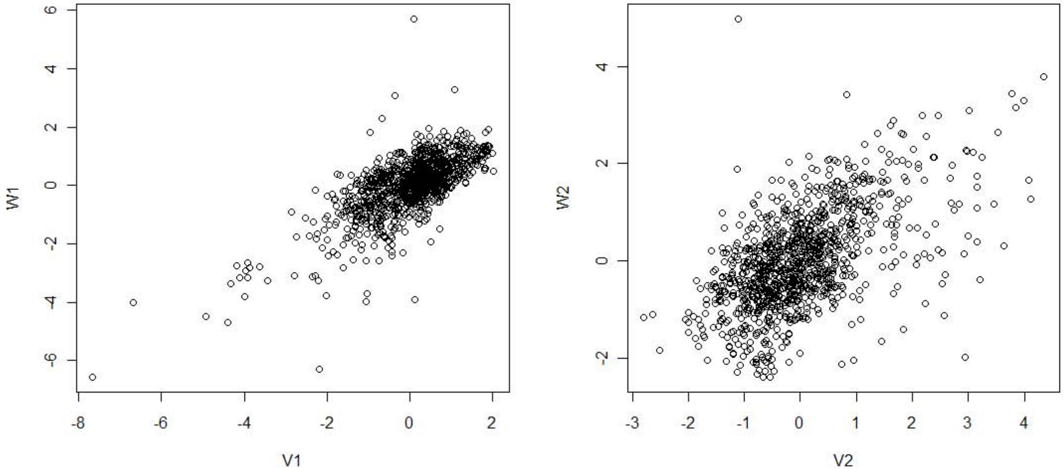

Fig. 9. Correlation between canonical variables.

HESSD

8, 3233–3269, 2011

Comparison of catchment grouping methods

E. Sauquet and C. Catalogne

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

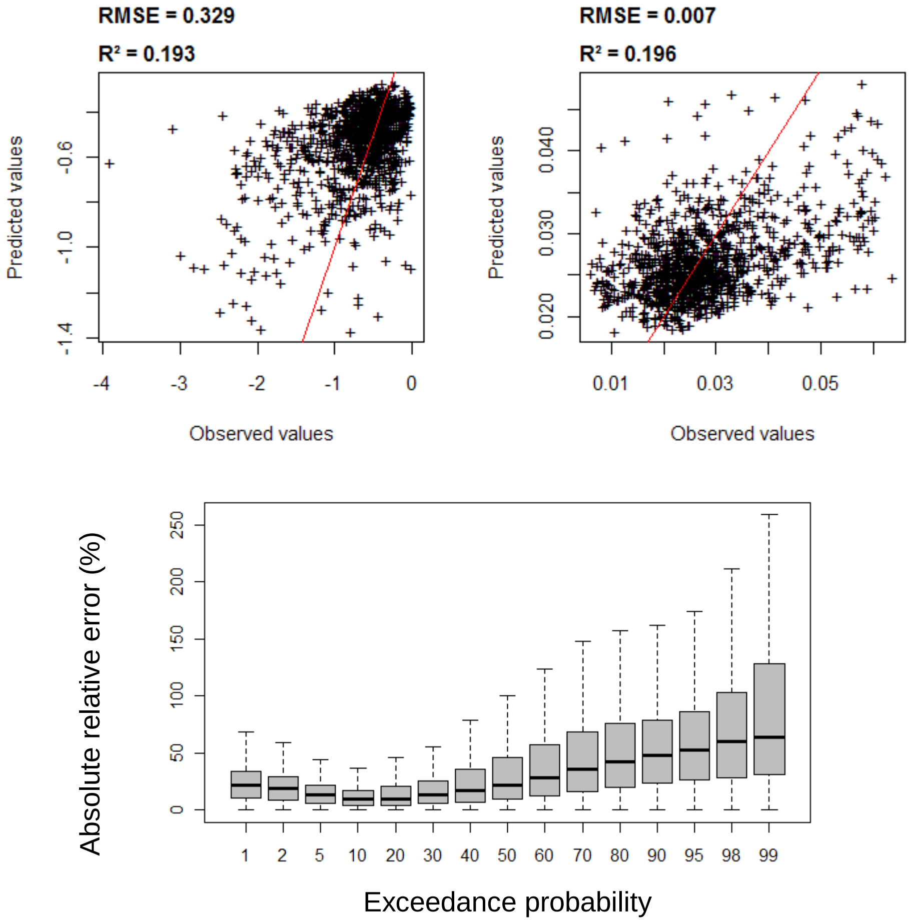

Fig. 10. Results for the global regression model.

HESSD
8, 3233–3269, 2011

Comparison of catchment grouping methods

E. Sauquet and C. Catalogne

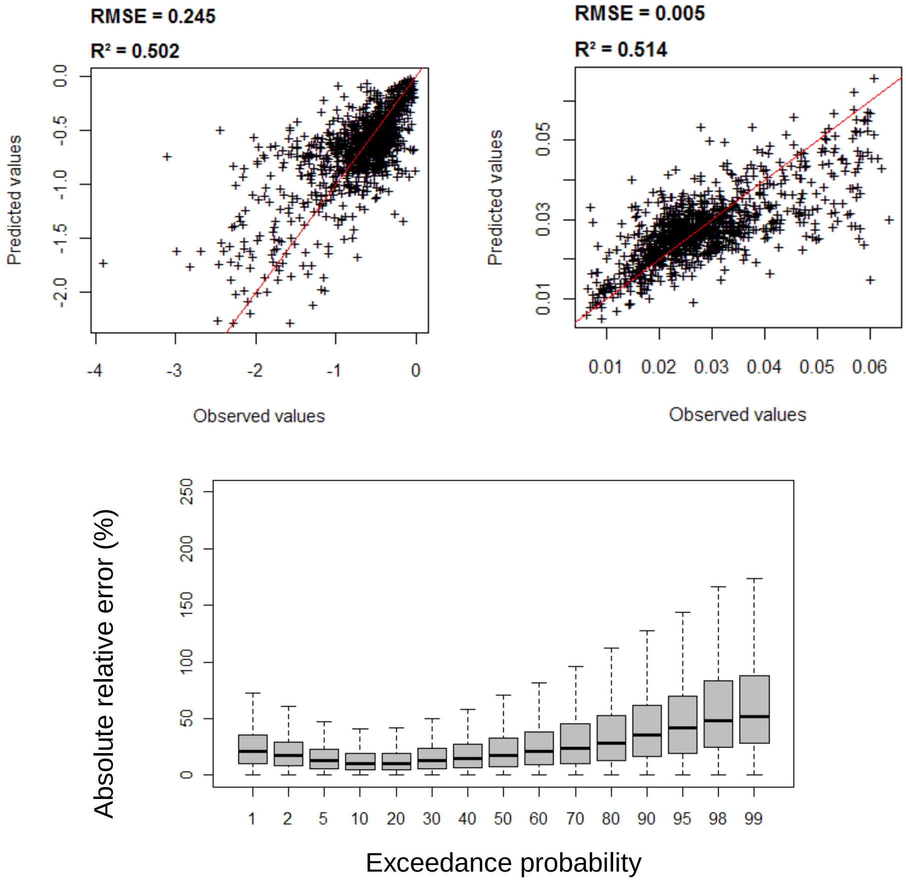

Fig. 11. Results for the regional regression model applied to visual grouping.

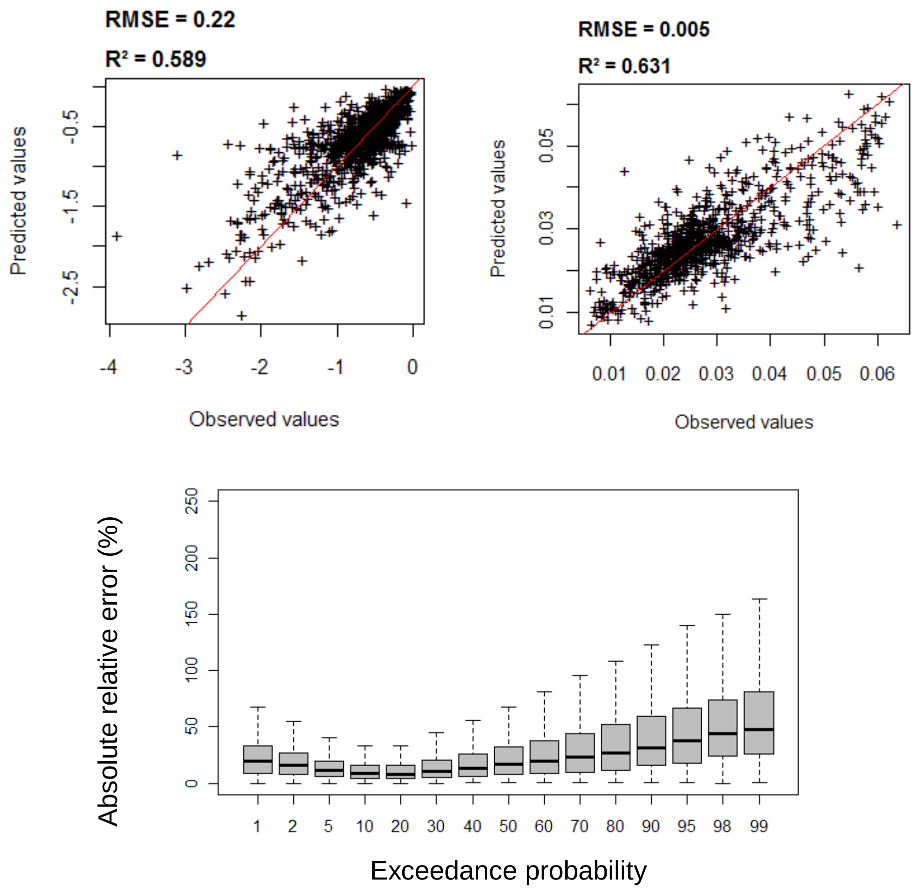

Fig. 12. Results for the regional regression model applied to groups derived from RT.

HESSD
8, 3233–3269, 2011

Comparison of catchment grouping methods

E. Sauquet and C. Catalogne

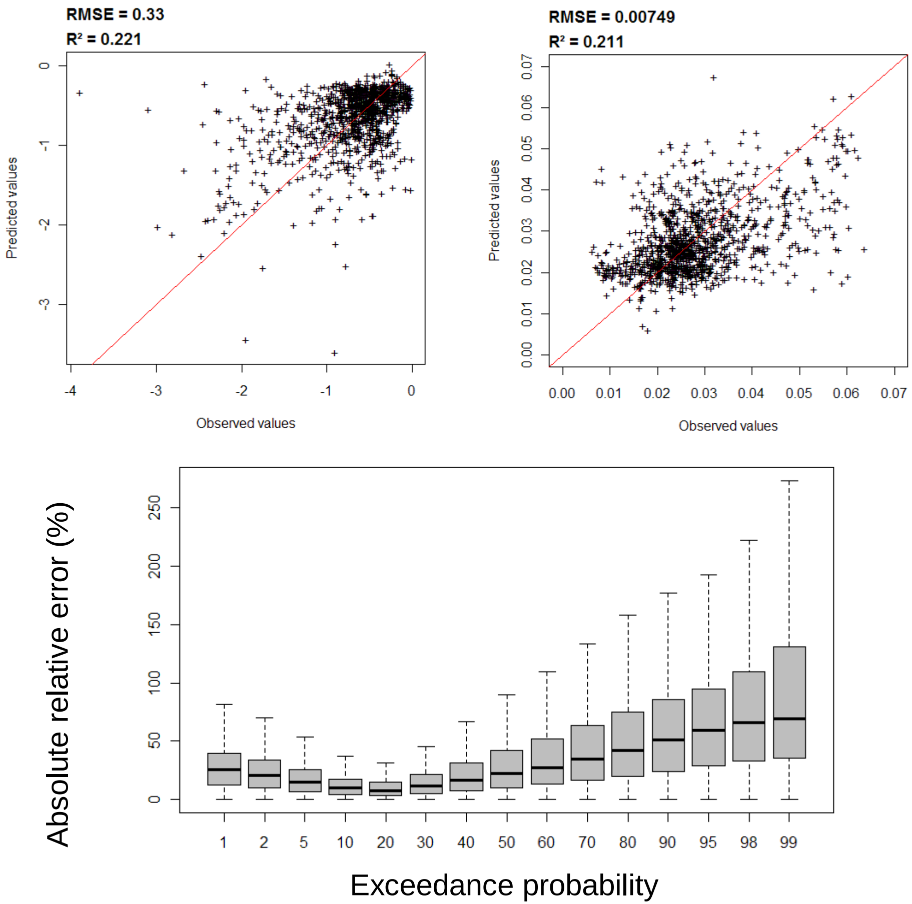

Fig. 13. Results for the regional regression model applied to neighbourhoods derived from CCA.

HESSD

8, 3233–3269, 2011

Comparison of catchment grouping methods

E. Sauquet and C. Catalogne