

The Quantile Solidarity approach for the parsimonious regionalization of flow duration curves

Carine Poncelet, Vazken Andréassian, Ludovic Oudin, Charles Perrin

▶ To cite this version:

Carine Poncelet, Vazken Andréassian, Ludovic Oudin, Charles Perrin. The Quantile Solidarity approach for the parsimonious regionalization of flow duration curves. Hydrological Sciences Journal, 2017, 62 (9), pp.1364-1380. 10.1080/02626667.2017.1335399. hal-03985673

HAL Id: hal-03985673 https://hal.science/hal-03985673

Submitted on 24 Feb 2023 $\,$

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The Quantile Solidarity approach for the parsimonious regionalization of flow duration curves

Carine Poncelet^a* and Vazken Andréassian^a and Ludovic Oudin^b and Charles Perrin^a

^aUR Hydrosystèmes et Bioprocédés, Irstea, Antony, France; ^bUMR Metis, University of Paris VI, Paris, France

*corresponding author: <u>carine.poncelet@irstea.fr</u>

The Quantile Solidarity approach for the parsimonious regionalization of flow duration curves

This paper presents a novel method to estimate flow duration curves (FDCs) for ungauged catchments with permanent flow. It is based on a dataset of 521 catchments located throughout France. The method consists in a three-step procedure called the Quantile Solidarity (QS) approach. First, a regression-based model is built to estimate FDCs in ungauged catchments, linking each flow quantile independently to physical descriptors. The second step consists in imposing the continuity of the regression-based model parameters along the quantiles (hence quantile solidarity) to obtain a large reduction of the number of parameters used to estimate the FDC. The last step consists in spatially interpolating the model residuals to further improve the performance of the FDC estimation. The QS approach yields a robust and parsimonious FDC estimation without any loss in simulation efficiency and ensures strictly monotonic FDCs.

Keywords: Flow duration curves; parsimony; PUB; regionalization; ungauged catchments

Subject classification codes: include these here if the journal requires them

Introduction

Flow duration curves

The flow duration curve (FDC) is a widely used hydrological characteristic summarizing the statistical distribution of streamflow at a catchment outlet [45]. It graphically shows the percentage of time that a given streamflow value is likely to be exceeded. FDCs have various uses, including sedimentation studies [52], hydrological models calibration [53] and comparison [26], thresholds definition for hydropower management [22], flood estimation [50], catchment classification [31] or estimation of environmental flows [14]. In this study, FDCs are indexed by their non-exceedance probability. For a gauged catchment, an observed/empirical FDC can be obtained directly from the streamflow record. The construction of an empirical FDC is of course not possible in an ungauged catchment: in this case, the FDC must be estimated through an adequate regionalization procedure. An exhaustive review of the existing methods was produced at the end of the Prediction on Ungauged Basins (PUB) decade (see [7] for further detail). The existing methods to estimate FDCs on ungauged catchments can be grouped into three categories:

- (1) Regression-based methods estimate each quantile separately from catchment features (climatic and/or physical characteristics). These methods do not require a hypothesis regarding the streamflow distribution or the FDC shape, but they require the identification of a large number of parameters. Moreover, the application of the relationships on ungauged catchments often induces undesirable numerical problems since flow values may not always increase with the quantiles [7]. The studies reported by [15] and [24] provide good examples of regression-based FDC estimations. In addition, the study reported by [2] also addressed the issue of overparameterization, as they exploited the strong structural relationship among streamflow quantiles.
- (2) Streamflow index-based methods involve normalization of the FDC using mainly the mean annual streamflow [3]. These methods generally require first the regionalization of the mean annual streamflow used to normalize the FDC and second the determination of the shape of the FDC by following either parametric or non-parametric approaches. The parametric approaches assume a regional shape (statistical law) for the FDC. The parameters of the statistical law are calibrated at the regional scale or predetermined using physical or climatic features and then applied to ungauged basins of the surrounding area (see e.g.

[25, 42]). Conversely, the non parametric approaches make no assumption on an underlying statistical law for the regional shape of the FDC. They require the implementation of a catchment classification and the identification of rules that allocate any ungauged catchment to a previously identified group. The ungauged catchment then receives the mean FDC of the group. The articles by [6], [37] and [19] are good examples of non-parametric FDC estimation on ungauged catchments.

(3) Geostatistical-based methods explore the spatial correlations between hydrological variables of gauged catchments to transfer them to ungauged catchments. The FDC is first computed on neighboring gauged catchments and then transferred to the ungauged catchments using the spatial distribution of observed FDCs. The studies of [47] and [9] are good examples of geostatisticalbased FDC estimations.

[7] conducted a comparative study of the different methods. They showed that regression-based and geostatistical-based methods perform better, especially when the gauging network is dense. These conclusions are in agreement with the results of other related studies by [33], [44] and [4].

Scope of the paper

This study proposes a regression-based regionalization of the FDC on a dataset of 521 French catchments. It is innovative in that we forced the parameter continuity along the quantiles to adress the two major limitations of regression-based methods: (i) lack of robustness (due to the high number of degrees of freedom) and (ii) potential errors due to non-incThis study proposes a regression-based regionalization of the FDC on a dataset of 521 French catchments. It is innovative in that we forced the parameter continuity along the quantiles to adress the two major limitations of regression-based methods: (i) lack of robustness (due to the high number of degrees of freedom) and (ii) potential errors due to non-increasing quantile values in ungauged catchments. The Quantile Solidarity (QS) approach takes advantage of the continuity of the regression parameters along the quantiles to drastically reduce the number of model parameters used to estimate FDC on ungauged sites. Last, we investigate the spatial coherence of the estimated FDCs' residuals to further improve the estimations in ungauged conditions. The final result is a parsimonious non-parametric estimation of the FDC. reasing quantile values in ungauged catchments.

The Quantile Solidarity approach

Flow duration curve estimations through a regression-based model

For each catchment, an empirical FDC is first computed from the observed streamflow using the entire available record. We considered 99 quantiles between 0.01 (low flow) and 0.99 (high flow), as a result the proposed methodology might not be suited to estimate extreme flows (i.e. beyond the 0.01 and 0.99 probabilities). Each quantile was modelled independently, by means of a simple multiplicative regression-based model:

$${}^{\wedge}Q_{i}^{j} = C_{i} (F1_{j})^{(a_{i})} (F2_{j})^{(b_{i})} (F3_{j})^{(c_{i})}$$
(1)

In equation 1, indexes *i* and *j* represent respectively the quantile and the catchment. Q_i^{j} (mm/d) is the estimated flow quantile, and $F1_j$, $F2_j$ and $F3_j$ are the catchment features. C_i is the constant term of the model while a_i , b_i and c_i are its parameters. These parameters are the same for all catchments but are a function of the quantile. Note that this multiplicative model cannot produce streamflow quantiles equal to zero (but the catchments studied do not include basins with intermittent flows).

Hence using the methodology to estimate FDCs of ungauged intermittent catchments can lead to quantile overestimation. To compute the parameter values, we first linearized equation 1:

$$\ln(^{A}Q_{i}^{j}) = K_{i} + a_{i}\ln(F1_{i}) + b_{i}\ln(F2_{i}) + c_{i}\ln(F3_{i})$$
(2)

Equation $\underline{2}$ is solved by following the ordinary least-squares scheme for the 261 catchments of the calibration set (see section 3.1 for further detail). The constant term (K) represents a unit corrector. Since equation 2 has four parameters and since it is repeated for each quantile, i.e., 99 times, this leads to a model with 396 calibrated parameters. Hereafter, this complete regression model is referred as CM396.

Reduction of the number of parameters by imposing quantile solidarity

Regression-based methods to estimate FDCs at ungauged sites often present serious limitations due to overparametrization: lack of robustness and non-monotonic estimated FDCs. A powerful way to address the issues of overparametrization and model transferability/robustness is to promote model parsimony. The QS approach presented here substantially reduces the number of calibrated parameters involvedusing the same notations as above. This parsimonious model has only ten degrees of freedom, and will be called PM10 hereafter. in regression-based approaches. The rationale of the QS approach is that the parameter values associated with each catchment feature will vary smoothly and monotonically with the flow quantiles.

The case of regression parameters

The QS approach hypothesizes that the models' parameters (a, b and c) are linear functions of the quantile *i*, as shown in equation 3. The choice of a linear function to link the parameters to the quantile will be further discussed in section 4.2.1:

$$\hat{\mathbf{a}}_{\mathbf{i}} = \boldsymbol{\alpha}_{\mathbf{a}} + (\boldsymbol{\beta}_{\mathbf{a}}) \, \mathbf{i}/100 \tag{3}$$

where α_a and β_a are called the metaparameters and allow one to compute a_i given quantile i. The metaparameters are identified by ordinary least squares regression for the 99 quantiles. Using this formulation for a, b and c drastically reduces the number of calibrated parameters from 396 to 105 degrees of freedom: two metaparameters for each feature and 99 associated with the constant term (presented below). Since a, b and c add a physically interpretable constraint to the relationship between the features and the quantiles, the combination of a, b, c and the features is referred to hereafter as the physical part of the model. Section 4.1.1 further investigates the physical role of each feature.

The case of the constant term

To further reduce the number of degrees of freedom, one must handle the constant term. After the metaparameters have been computed, the constant term is recalibrated on the calibration catchments as the difference between observed flow quantiles and the quantiles estimated using equation 3. The recalibration allows to take into account the differences between the initial model parameters and their estimates using equation 3. Lastly, K is modelled using an error function, the Gaussian quantile function:

$$K_{i} = \alpha_{K} + \beta_{K} G(i/100; \mu_{K}, \sigma_{K})$$
(4)

The Gaussian quantile function was chosen both for theoretical reasons and fitting suitability. Indeed, when recalibrated, K is an error term since it takes in account both the unexplained variance and the errors made while regularizing the other parameters (a, b and c). The constant term is expressed as a function of the quantile using four parameters (instead of 99). By combining equations 2, 3 and 4 we obtain a far more parsimonious model:

$$\ln(^{Q_{j}}) = \alpha_{K} + \beta_{K} G(i/100; \mu_{K}, \sigma_{K}) + (\alpha_{a} + \beta_{a} i/100) \ln(F1_{j}) + (\alpha_{b} + \beta_{b} i/100) \ln(F2_{j}) + (\alpha_{b} + \beta_{b} i/1$$

using the same notations as above. This parsimonious model has only ten degrees of freedom, and will be called PM10 hereafter.

Residuals regionalization

The model's residuals are defined:

$$\epsilon_{i,j} = (Q_r^{j})/(^{A}Q_r^{j})$$
(6)

where $Q_i j^i$ (mm/d) is the observed flow quantile and $^{A}Q_i j^i$ (mm/d) is the estimated flow quantile. Ninety-nine residuals were computed for each catchment of the calibration subset. Since we are not able to explain all of the FDC variability with the regression model, physical features that are not included in the FDC regression equations make the residuals correlated in space (see section 4.2.2). Therefore, the residuals are not randomly distributed among the calibration catchments and, consequently, a spatial interpolation of residual values can still improve the FDC estimation model. To interpolate this residual information, we chose a simple method based on the squared inverse geographical distance:

$$\epsilon_{i,k} = (1/D) \sum_{j} [(1/\operatorname{dist}(k,j))^2 \epsilon_{i,j}]$$
(7)

where D is the sum of the weights (squared inverse geographical distance). The index k refers to the target catchment (considered ungauged) and index j refers to the neighbour catchments (considered gauged) and dist(k,j) the geographical distance between catchments k and j as defined in equation 8. Equation 7 gives the computed residual for an ungauged catchment k as a linear combination of observed residuals of

the calibration catchments. They are numerous ways to computed geographical distance between catchments. Based on the work of [29] we retained a weighted Euclidean distance:

dist(k, j) = 0.2
$$((XO_k - XO_j)^2 + (YO_k - YO_j)^2)^{1/2} + 0.8 ((XC_k - XC_j)^2 + (YC_k - YC_j)^2)^{1/2}$$

(8)

with (XO_k, YO_k) the outlet coordinates of catchment k and (XC_k, YC_k) the centroid coordinates of catchment k. The advantage of this distance is that it implicitly takes into account catchment area and preferably select neighbouring catchments with a size similar to that of the target catchment.

The parsimonious model with regionalized residuals will be called PM10+REG in the following sections and is defined by equation 9:

$$^{A}Q_{i}^{j} = K_{i} (F1_{j})^{\hat{a}_{i}} (F2_{j})^{^{A}b_{i}} (F3_{j})^{\hat{c}_{i}} (^{^{A}}\epsilon_{i, k})$$
(9)

using the same notations as above.

Dataset and evaluation procedure

Catchment set

We based this study on a set of 521 French catchments, selected based on: (i) availability of daily streamflow records for the 1982-2002 period, with less than three years of missing data, (ii) no catchments with intermittent streamflow (minimum recorded flow is 0.05 mm/d), (iii) unimpacted catchments, i.e., not significantly influenced by anthropic activities (regulation, water pumping, etc). For each catchment, we retrieved the daily series ofwe retained a weighted Euclidean distance observed streamflow from the national HYDRO archive [<u>30</u>]. Prior to this study, all streamflow

records were checked visually and compared using double mass curve analysis by [<u>10</u>] to ensure streamflow data accuracy and consistency. Table 1 summarizes a few hydroclimatic characteristics observed over the catchment set.

Table 1 shows that a wide range of hydroclimatic conditions is represented over the 521 catchments. This large-sample approach is particularly useful because general results can be obtained (see e.g. [1, 20]). The drainage support area listed in Table 1 is expressed as the average catchment area above an observed stream source and will be further described in section 3.2.

Fig. 1 shows the location of the catchments used in the study. The catchment set was split into two subsets for the calibration-validation procedure: 261 catchments were used to calibrate the QS approach and 260 catchments for its validation. We allocated the catchments so that the widest range of hydroclimatic conditions would be covered in both subsets.

Selection of the catchment features

Preliminary tests (not shown here for the sake of brevity but presented as supplementary material) identified the three most relevant catchment features for the estimation of FDCs at ungauged sites using the CM396 model. To identify these dataset features, we used a pool of 17 features (climate, soil, morphology, land cover and geology) and tested all possible combinations of three and four features. To avoid weak model structure, we only kept the combinations in which the Spearman correlation between the catchment features were less than 0.5, for a total of 887 combinations. The scope of the paper being the parsimonious regionalization of the FDC, we will not present here all tested features and will only discuss the choice of the three retained features. The abbreviation for all tested features can be found in Appendix 1. The combination that yielded the best Kling and Gupta efficiency [21] value between empirical and estimated

FDCs over the calibration set and whose feature parameters expressed the most stability across combinations was selected to apply the Quantile Solidarity approach. The use of KGE is not conventional but we believe it is a good option for relative comparison between possible features combination. Fig. 2 shows for each feature the mean performance of all three-feature combinations they are involved in. Since we selected only combinations where feature correlation are lower than 0.5, each feature can be involved in a different number of combinations.

Fig. 2 shows that the humidity index (HI) yields the highest mean performance (KGE = 0.55). However, the rank of the first combination including HI is 24. In other words, the use of HI in a combination does not automatically lead to a high performance but seems to be a safeguard against bad FDC estimations. Fig. 2 highlights some equifinality issues between the combinations. This is due to the large number of degrees of freedom of the regression model and resulting overparametrization. Since the parameter continuity along the quantiles is not affected by the choice of features (see supplementary material), we only worked with one combination of features to illustrate the proposed methodology. We believe that our results (parsimonious regression model and monotonic FDCs) will not be impacted by the choice of other features.

The methodology used to select the relevant features relies on two assumptions: (i) three features are sufficient to estimate FDCs at ungauged sites and (ii) all quantiles can be estimated by the same features. These assumptions are an integral part of the QS approach and will be discussed in section 4. Following this methodology, we retained the DSA of the observed stream network (km²), the mean daily precipitation (P, mm/d) and the mean daily potential evapotranspiration (E0, mm/d) to base the FDC estimation on. Fig. 3. illustrates the computation of DSA. The DSA is defined by [48] and [28] as the area needed to observe a stream within a catchment. It is computed on the basis of the observed river network and digital elevation model (DEM). The observed river network comes from the CARTHAGE database. The DEM serves as basis to create the drainage direction map used to optimize an unique area upstream of each source in the observed river network. This "calibrated" area minimizes the distance between the observed stream and the DEM-extracted stream and is the DSA. For this study, we used the SRTM DEM with a pixel resolution of 100 m as used by [28]. Since the DEM is not used to define the river network, the impact of the DEM resolution on the DSA value is limited. Large DSA indicates a permeable geology, favoring the development of large aquifers.

FDC estimation deals with a double variability: the variability between quantiles and the variability between catchments. It therefore seems logical to assess the overall model performance on these two aspects. We evaluated the catchment variability using the relative error as a percentage:Mean daily precipitation and potential evapotranspiration were computed using time series spanning the period between 1982 and 2002 obtained from the Météo France's SAFRAN reanalysis [51]. Mean daily precipitation is computed directly from the daily P included in SAFRAN. Mean daily potential evapotranspiration is computed using the Oudin formula [39] based on the daily temperature included in SAFRAN and the extraterrestrial radiation derived as a function of the latitude and the Julian day.

Baseline FDC estimation at ungauged sites (reference method)

Even though only 261 stations are used for calibration in our study, the gauged network is considered dense because all validation catchment have calibration catchments nearby (the density of the gauge network is 1 gauging station for 1050 km²). Because spatial

proximity-based methods perform very well when the gauge network is dense, we chose as reference method a slightly modified form of the classical index-station method presented by [45]. The method proposed by [45] initially aimed at adjusting short-term records for long-term representativity. It consisted in plotting the flow duration curves computed from short but concurrent records of neighboring catchments. The assumption was that the relation obtained from the short records is constant in the long-term so that the information obtained on catchments with long records can be used to estimate the FDC of catchments with short records. For our baseline method, we retained five calibration catchments because it is a good compromise between using local information and obtaining robust regionalization [38]. A composite streamflow time series for the validation catchment was built using the median observed streamflow (in mm/d) of the five closest gauged catchments at each time step and we then computed the FDC from it. The comparison of the performance of this method with a comparable study [44] showed that this non-parametric and simple method it is an efficient benchmark, which is why we did not consider other grouping strategy or reference methodsFDC estimation deals with a double variability: the variability between quantiles and the variability between catchments. It therefore seems logical to assess the overall model performance on these two aspects. We evaluated the catchment variability using the relative error as a percentage:.

Evaluation framework

FDC estimation deals with a double variability: the variability between quantiles and the variability between catchments. It therefore seems logical to assess the overall model performance on these two aspects. We evaluated the catchment variability using the relative error as a percentage:

$$RE_{i,j} = 1 - (Q^{i}_{j})/(^{A}Q^{i}_{j})$$
(7)

fusing the same notations as above. For a perfect model, the relative error is equal to 0. RE is negative when the model underestimates the quantile, positive when it is overestimated. Since RE is normalized using the flow quantile, it tends to have higher values when the flow quantiles are low. The distribution of $RE_{i,j}$ values describes, for a given quantile i, the error distribution in the dataset. The median and quartile errors for quantile i give a measure of the bias and the accuracy of the model, respectively.

The model performance regarding quantile variability for each validation catchment was evaluated using the Nash-Sutcliffe efficiency [34]:

$$NSE_{j} = 1 - \left(\sum_{i} (^{Q}Q_{j}^{i} - Q_{j}^{i})^{2}\right) / \left(\sum_{i} (Q_{j}^{i} - Q_{j}^{i})^{2}\right)$$
(7)

using the same notations as above. For a perfect model, NSE is equal to 1. We selected this criterion over the KGE used to select the features because NSE allows us to compare this approach with other related studies more easily. Since NSE puts greater emphasis on high flows, the NSE and the relative error criteria are complementary. Table 2 summarizes the different models tested and evaluated in this study.

Results and discussion

Physical relevance of the selected features

Hydrological interpretation of the features retained

Fig. 4 shows the parameters obtained by calibration for the CM396 model. As stated before, the best combination which we retained for the FDC estimation model involved three catchments features: the Drainage Support Area (DSA), the mean daily precipitation (P) and the mean daily potential evapotranspiration (*E0*). It is interesting to discuss the physical meaning of the curves shown in Fig 4 (here, we only examine the individual quantile points and will discuss the parametric simplification of the curves later):

- The DSA parameter is positive for the first 30 quantiles (low flows) and negative for the last quantiles (medium and high flows). In other words, low flows are positively correlated with DSA values (large DSA catchments will have a tendency to have more sustained low flows), while medium and high-flows are negatively correlated with them (small DSA catchments will have larger high-flows). As shown in Fig. 3, a large DSA is equivalent to a catchment of low drainage density, which in France is likely to be associated with aquifers that support low flows. In flood conditions, a low drainage density catchment will have a longer transit time to the outlet, which means a buffering impact on high flows [11].
- The *P* parameter is positive whatever the quantile, meaning that the larger the average precipitation, the larger the streamflow quantile. This result was obviously expected since precipitation is the main forcing of streamflow generation [<u>36</u>, <u>5</u>].
- The *E0* parameter is always negative: the larger the potential evapotranspiration, the lower the streamflow quantile. This result was also expected. Note that the E0 parameter is close to zero for flood flows, meaning that potential evaporation has a very limited impact on the highest flow quantiles.

Why did some features not appear? The case of the catchment area

While it is extremely instructive to discuss why the selected features were retained, it

can also be informative to discuss why some expected features were not. The case of the catchment area is particularly interesting because many FDC studies use it [<u>17</u>, <u>46</u>, <u>8</u>, <u>32</u>]. Catchment area usually appears in the role of the *DSA*: it is negatively correlated with high flows (larger catchments have attenuated flood flows) and positively correlated with low flows (larger catchments have more sustained low-flows). Moreover, the Spearman correlation between A and DSA is 0.38. It should be remembered that the 887 combinations of three and four features tested as a prelude for this study are presented in the supplementary material. These results show that (i) in the case of a three-feature combination adding area adds only little explanatory power to the three features selected (see Fig. 5). To conclude on the catchment area, we can only state that (contrary to our expectations) its impact on specific flows is very limited, and that the *DSA* used here has a greater explanatory power.

Explanatory power of the relationship retained

Fig. 5 shows the percentage of variability (R²) explained by the combination of DSA, P and E0. Please note that the percentage of explained variability is computed for the CM396 model, i.e., the linear shape function used to constrain the parameters of the model along the quantiles does not impact the explanatory power of the features used.

Fig. 5 shows that the catchment features only provide enough information to explain 70 percent or more of the variability for quantiles between 50 and 99, i.e., the high-flow quantiles. For the low flows however, the percentage of variability explained by the features is far below 70%: in this case, adding residual information based on spatial proximity can be particularly valuable. We will propose relevant features to improve the model explanatory power based on the analysis of the spatial structure of errors in section 4.2.2.

Performance of the QS approach

Which shape should be used to link parameters with features?

The idea of the Quantile Solidarity approach came from the analysis of the graphs presented in Fig. 4. The parameters calibrated independently for each quantile show a continuous evolution pattern, not strictly linear but linear in a first approximation. We are aware that this is a rough approximation and we did attempt to improve the match between calibrated and constrained parameters by adapting the function to the shapes observed in Fig. 4. We found that the gain was counterbalanced by: (i) a loss on the identifiability of the error term (K) and its constraint (equation 4) and (ii) the loss of FDC monotonicity (see section 4.3). For this reason and for the sake of simplicity, we chose to retain the linear approximation to compute the constrained parameters as a function of the quantile. Fig. 4 provides qualitative insights into this match while Table 3 provides numerical evidence.

Table 3 shows the metaparameter values to compute the constrained parameters from the quantiles. We see that the quality of the adjustments is rather satisfying: the linear regression captures most of the variability of the calibrated parameters, the greatest errors are restricted to the extremes (low and high flows), i.e., for a few quantiles. It should be remembered that the objective here was to estimate the FDC (quantiles 0.01 to 0.99). If we were interested only in the low-flow or high-flow section, we would probably have made different choices.

Spatial structure of the model errors

In this section, we discuss the spatial structure of the errors of the PM10 model, i.e. just before the residuals regionalization. Errors are estimated on the calibration set as relative errors (see equationv 10) between the FDC computed with the PM10 model and the observed FDC. Fig. 6 presents a map of the error for three quantiles of interest: Q5 (low flows), Q50 (median flows) and Q95 (high flows). A negative relative error corresponds to an underestimation, while a positive error means the model has overestimated the given flow quantile.

Fig. 6 shows regional error patterns since neighboring catchments have similar RE values for all quantiles. The spatial error patterns emerge more clearly for the extreme quantiles than for the mid flow quantile. The PM10 model underestimates the Q5 and Q50 quantiles and overestimates the Q95 quantile for catchments located in northern France. These catchments are located in highly permeable sedimentary basins and fed by groundwater aquifers that buffer high flow and regulate low flows. In Brittany (North-West of France), the model overestimates low and medium flows and underestimates high flows. These catchment have a crystalline (impermeable) bedrock: there is almost no groundwater to sustain the low and mid flows. The mountainous catchments located in the Alps, the Pyrenees and the Massif Central express similar error patterns for the three quantiles. The PM10 model underestimates all flow quantiles, and particularly the low flows. In mountainous catchments low flows occur during winter, when precipitations are stored as snow. As a result, the low flow are less pronounced and underestimated by the PM10 model. During the mid and high flow, the errors are also explained by the altitude pattern and related to the underestimation of precipitation inputs in mountainous regions [12].

The coherent spatial patterns of model errors justify the regionalization of the model residuals based on spatial proximity. The analysis indicates that the errors can be related to dominant flow regimes at the regional scale that PM10 does not take into account since we made the choice to use the same features for all catchment. PM10 estimates the FDC based only on the hydrological processes which are dominant at the national scale. Likewise, the same features are used to estimate all quantiles whereas the dominant drivers are often different between high and low flows.

Can the catchment variability of FDCs be captured?

This section deals with what we called "catchment variability" of FDC because, at a given quantile, it allows to discuss the variability of FDC estimations performance between the catchments. Fig. 7 shows how model performance, measured by the relative error, evolves along the quantiles. A positive relative error corresponds to an underestimation, while a negative error means the model has overestimated the flow quantiles. Note that all errors are computed for the catchments belonging to the validation set.

By comparing Fig. 7(a) and Fig. 7(b), we assess the impact of reducing the number of parameters. Both models tend to slightly overestimate most of the flow quantiles. The median performance of each of the two models is equivalent, except for the last quantiles (high flow) where the PM10 model performs worse and overestimates these quantiles, because the constrained constant term K is further away from the calibrated K for these few quantiles. However, compared to the CM396, the extent of the boxplots and of the whiskers of the PM10 are narrower. This means that the QS approach slightly degrades the median performance of the FDC estimation but limits poor estimations, i.e., it is more robust.

Comparing Fig. 7(b) and Fig. 7(c) assesses the impact of the residuals regionalization. This step has a different impact on low flows and high flows. For the former, the median performance is slightly degraded and the flow quantiles are overestimated. For the latter, the residuals regionalization corrects the bias observed on high flows and improves the model's median behaviour. This improvement is mostly due to the fact that the regionalization step tends to correct the model bias for most quantiles. The variability of the relative error is narrower after the residuals regionalization for both low and high flows. However, the larger errors remaining for the low flow quantiles may be the reflection of the constant exponent chosen for the inverse distance scheme (equation 7). Indeed keeping the same exponent for all quantiles means that the same weight is given to the donor catchments throughout the quantiles and we could as a perspective adjust the donor group (defined by the spatial proximity) to each quantile by using a variable exponent. As shown by Fig. 6 the RE spatial structure is different between the quantiles. As a result, adjusting the residuals regionalization procedure to the quantile value might provide further improvement to the FDC estimation.

Finally, by comparing Fig. 7(c) and 7(d), we evaluate the performance of the FDC estimation model compared to the reference method. The median performance of the reference method is closer to the optimal value of zero for the first 30 quantiles (Q1 to Q30), but further away for the last 60 quantiles (Q40 to Q99). For all quantiles, the variability of the relative error is larger for the reference model. These observations allow us to conclude that the PM10+REG model is robust and performs at least as well for most quantiles as the reference model used.

Are we able to capture the quantile variability?

Fig. 8 shows the performance of the models in terms of relative errors (graphs a, b and

c) and Nash-Sutcliffe efficiency. The NSE distributions are made with one value per catchment, which summarizes the errors on the 99 quantiles albeit with a measure that is more sensitive to high-flow errors. To assess whether the performance differed between the regionalization methods used, we performed a Friedman test [<u>18</u>], as shown in the legend of Fig. 8.

Fig. 8(d) shows that the parsimonious model without residuals regionalization performs significantly worse than the other FDC estimation with a median NSE of 0.85 when the median NSE of CM396 and REF are 0.91 and 0.90 respectively. Note that 12% of the catchments have a NSE less than 0.5 using the PM10 model, while this number increases to 14% and 15% for CM396 and REF, respectively. These results confirm those shown in Fig. 7, namely that the parsimonious model (PM10) is more robust than REF and CM396.

The parsimonious model with residuals regionalization (PM10+REG) performs significantly better than all models with a median NSE of 0.95 and only 10% of catchments with a NSE less than 0.5. The residuals regionalization step is highly beneficiary for the estimation of the FDC. Fig. 7 a-c shows that the improvement of the NSE for the PM10+REG model mostly stems from reducing the errors for the high-flow quantiles and avoiding larger errors for all quantiles. The fact that PM10+REG outperforms the reference model confirms that regionalizing residual information instead of the flow quantiles themselves improves the performance of the regionalization step [<u>33</u>].

Monotonicity of the estimated FDC

Regression-based methods can produce FDC estimations that do not increase monotonically with the quantiles [7]. This section aims at demonstrating that FDCs estimated with the QS approach are strictly increasing, at least on the catchment set tested herein. The condition for a strictly increasing estimated FDC is that its derivative is strictly positive. Equation 12 is obtained by taking the derivative of equation 5:

$$d\ln(^{Q})/di = d[\alpha_{K} + \beta_{K} G^{-1}(i;\mu_{K},\sigma_{K})]/di + d[(\alpha_{a} + \beta_{a} i) \ln(DSA)])/di + d[(\alpha_{b} + \beta_{b} i)]/di + d[(\alpha_{c} + \beta_{c} i) \ln(E0)]/di$$
(12)

By translating equation 12 into the terms of this study, we obtain the following derivatives:

$$d[\alpha_{K} + \beta_{K} G^{-1}(i;\mu_{K},\sigma_{K})]/di = \beta_{K}/G'(G^{-1}(i;\mu_{K},\sigma_{K})) = 0.99/G'(G^{-1}(i;-1.59,0.60))$$

$$d[(\alpha_{a} + \beta_{a} i) \ln(DSA)]/di = \beta_{a} \ln(DSA) = -0.12 \ln(DSA)$$

$$d[(\alpha_{b} + \beta_{b} i) \ln(P)]/di = \beta_{b} \ln(P) = 0.95 \ln(P)$$

$$d[(\alpha_{c} + \beta_{c} i) \ln(E0)]/di = \beta_{c} \ln(E0) = 0.57 \ln(E0)$$

(13)

using the same notations as above and with G' being the Gaussian density

function. Hence the estimated FDC increases if the sum of the terms of equation 13 is positive. To demonstrate this for the conditions of the study, we considered the worst case scenario over the catchment set, i.e., the numerical values that tend to maximize the negative terms and minimize the positive terms of equation 13. This worst case scenario is reached on our catchment set for E0=0.61 mm/d, P=1.80 mm/d and DSA=195.36 km².

 $0.95 \ln(P)=0.56$

$0.57 \ln(E0) = -0.28$

(13)

With this scenario, equation 12 equals 2.18, which guarantees strictly increasing FDCs in the conditions of this study. The next step is to assess whether this numerical application is valid for all catchment in France. The worst case scenario for France is reached for E0=0.10mm/d, P=1.34mm/d and DSA=200km². With these numerical values, equation 12 equals 0.86 which means that constraining the parameters' value as a function of the quantile is sufficient to guarantee increasing estimated FDCs for France.

Conclusions

In this study, we have presented a new and parsimonious way to estimate flow duration curves for ungauged catchments. The parsimonious model was made possible by describing the coherence of the model parameters in the frequency spectrum. We show that this new approach guarantees the estimation of strictly increasing FDC, at least on the conditions found in France. Last, we were able to quantify the added value of including regional information by obtaining an additional improvement in model performance by regionalizing models' residuals obtained for the calibration set. We see this three-step approach as the most satisfying from a hydrological point of view: the physical-climatic determinants of flow quantiles are first fully exploited, and then the remaining unexplained residual information is exploited through a simple neighbourbased (regional) transfer scheme. The final comparison with a reference method demonstrates the performance of the approach presented here.

However, the approach has its limits and possible ways to further improve it can be mentioned. A first limitation of this study is that we are not able to estimate the flow duration curve of intermittent catchments. To address this issue, the regression-based model that is multiplicative may need to be replaced with an additive model that allows null flow quantiles, see e.g. the work by [32, 41, 54]. A second limitation is that for low flow quantiles, the percentage of explained variance remains low. We do not think that the low efficiency in reconstituting low flows stems from retaining only three descriptors (four-descriptor combinations were tested without noticeable improvement). We believe that the problem lies in the information content of the available descriptors: further work is needed to identify more relevant physical descriptors for low flows. It should also be mentioned, that although all the catchments are unregulated, they are not uninhabited, and some human influence could affect the lowest flows [55], making residual transfer even more difficult. Along with the use of different descriptors, the fit between calibrated parameters and constrained parameters could also be improved.

This work could be continued by further investigating the constant exponent chosen for the regionalization of the residuals. The power of the distance was set at 2 for all quantiles (equation 7). A larger exponent would put a larger weight on the closest catchment, while a smaller exponent would collect information from a larger number of catchments. From the analysis of the spatial structure of the model errors, it is likely that the "average" quantiles (around 0.7) could benefit from a smaller exponent, while the more extreme quantiles could benefit from a larger one. All these hypotheses remain to be confirmed.

Acknowledgement

The authors would like to thank the Knowledge and Innovation Center (Climate KIC program) for funding the first author.

References

- [1] Vazken Andréassian, A. Hall, N. Chahinian, J. Schaake: Large Sample Basin Experiments for Hydrological Model Parameterization: Results of the Model Parameter Experiment–MOPEX. 2006.
- [2] Stacey A. Archfield, Richard M. Vogel: "A decision-support system to assess surface-water resources in Massachusetts", World Environmental and Water Resources Congress 2008: Ahupua'A, pp. 1—10, 2008.
- [3] A. E. Best, Lu Zhang, T. A. McMahon, A. W. Western: "Development of a model for predicting the changes in flow duration curves due to altered land use conditions", MODSIM 2003 International Congress on Modelling and Simulation, pp. 861—866, 2003.
- [4] D.J. Booker, T.H. Snelder: "Comparing methods for estimating flow duration curves at ungauged sites", Journal of Hydrology, pp. 78—94, 2012.
- [5] M. I. Budyko: Climate and life. New-York: Academic Press, 1974.
- [6] Donald H. Burn: "An appraisal of the "region of influence" approach to flood frequency analysis", Hydrological Sciences Journal, pp. 149—165, 1990.
- [7] A. Castellarin, G. Botter, D. A. Hughes, S. Liu, T. B. M. J. Ouarda, J. Parajka, D. A. Post, M. Sivapalan, C. Spence, A. Viglione, R. M. Vogel: Prediction of flow duration curves in ungauged basins in Runoff Prediction in Ungauged Basins (Blöschl, G. and Sivapalan, Murugesu and Wagener, Thorsten and Viglione, Alberto and Savenije, Hubert, ed.). 2013.
- [8] A. Castellarin, G. Galeati, L. Brandimarte, A. Montanari, A. Brath: "Regional flowduration curves: reliability for ungauged basins", Advances in Water Resources, pp. 953—965, 2004.
- [9] S. Castiglioni, A. Castellarin, A. Montanari: "Prediction of low-flow indices in ungauged basins through physiographical space-based interpolation", Journal of Hydrology, pp. 272—280, 2009.
- [10] Laurent Coron: Les modèles hydrologiques conceptuels sont-ils robustes face à un climat en évolution? Diagnostic sur un échantillon de bassins versants français et australiens. 2013.
- [11] S. L. Dingman: "Drainage density and streamflow : a closer look", Water Resour. Res., pp. 1183—1187, 1978.

- [12] Y. Durand, E. Brun, L. Mérindol, G. Guyomarch, B. Lesaffre, E. Martin: "A meteorological estimation of relevant parameters for snow models", Annals of glaciology, pp. 65—71, 1993.
- [13] EEA: CLC2006 technical guidelines. Publications Office, 2007.
- [14] A. Efstratiadis, A. Tegos, A. Varveris, D. Koutsoyiannis: "Assessment of environmental flows under limited data availability: case study of the Acheloos River, Greece", Hydrological Sciences Journal, pp. 731—750, 2014.
- [15] Neil Fennessey, Richard M. Vogel: "Regional flow-duration curves for ungauged sites in Massachusetts", Journal of Water Resources Planning and Management, pp. 530—549, 1990.
- [16] P. Finke, R. Hartwich, R. Dudal, J. Ibanez, M. Jamagne, D. King, L. Montanarella, N. Yassoglou: Georeferenced soil database for Europe. European Soil Bureau Scientific Committee, 2001.
- [17] Marco Franchini, Mauro Suppo: "Regional analysis of flow duration curves for a limestone region", Water Resources Management, pp. 199—218, 1996.
- [18] M. Friedman: "The use of ranks to avoid the assumption of normality implicit in the analysis of variance", J. Am. Stat. Assoc., pp. 675—701, 1937.
- [19] D. Ganora, P. Claps, F. Laio, A. Viglione: "An approach to estimate nonparametric flow duration curves in ungauged basins", Water Resources Research, pp. 10, 2009.
- [20] H. V. Gupta, C. Perrin, G. Blöschl, A. Montanari, R. Kumar, M. Clark, V. Andréassian: "Large-sample hydrology: a need to balance depth with breadth", Hydrology and Earth System Sciences, pp. 463—477, 2014.
- [21] Hoshin V. Gupta, Harald Kling, Koray K. Yilmaz, Guillermo F. Martinez:
 "Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling", Journal of Hydrology, pp. 80—91, 2009.
- [22] Leroy F. Heitz, Shahram Khosrowpanah: Prediction of Flow Duration Curves for Use in Hydropower Analysis at Ungaged Sites in Pohnpei, FSM. Water and Environmental Research Institute of the Western Pacific, University of Guam, 2010.
- [23] Ladislav Holko, Juraj Parajka, Zdeno Kostka, Peter Škoda, Günter Blöschl:
 "Flashiness of mountain streams in Slovakia and Austria", Journal of Hydrology, pp. 392—401, 2011.

- [24] Allen Hope, Ryan Bart: "Evaluation of a Regionalization Approach for Daily Flow Duration Curves in Central and Southern California Watersheds: Evaluation of a Regionalization Approach for Daily Flow Duration Curves in Central and Southern California Watersheds", Journal of the American Water Resources Association, pp. 123—133, 2012.
- [25] J.R.M. Hosking: "L-Moments: Analysis and Extimation of Distribution Using Linear Combination of Order Statistics", Journal of the Royal Statistical Society, pp. 105—124, 1990.
- [26] C. Jothityangkoon, M. Sivapalan, D. L. Farmer: "Process controls of water balance variability in a large semi-arid catchment: downward approach to hydrological model development", Journal of Hydrology, pp. 174—198, 2001.
- [27] Yann L'hôte, Pierre Chevallier, Anne Coudrain, Yves Lejeune, Pierre Etchevers: "Relationship between precipitation phase and air temperature: comparison between the Bolivian Andes and the Swiss Alps / Relation entre phase de précipitation et température de l'air: comparaison entre les Andes Boliviennes et les Alpes Suisses", Hydrological Sciences Journal, 2005.
- [28] Nicolas Le Moine: Le bassin versant de surface vu par le souterrain: une voie d'amélioration des performances et du réalisme des modèles pluie-débit?. 2008.
- [29] Laure Lebecherel, Vazken Andréassian, Charles Perrin: "On regionalizing the Turc-Mezentsev water balance formula", Water Resources Research, pp. 7508—7517, 2013.
- [30] I. Leleu, I. Tonnelier, R. Puechberty, P. Gouin, I. Viquendi, L. Cobos, A. Foray, M. Baillon, P. O. Ndima: "Re-founding the national information system designed to manage and give access to hydrometric data", La Houille Blanche, pp. 25—32, 2014.
- [31] R. Ley, M. C. Casper, H. Hellebrand, R. Merz: "Catchment classification by runoff behaviour with self-organizing maps (SOM)", Hydrology and Earth System Sciences, pp. 2947—2962, 2011.
- [32] Giuseppe Mendicino, Alfonso Senatore: "Evaluation of parHydrological role of the drainage support area (DSA).ametric and statistical approaches for the regionalization of flow duration curves in intermittent regimes", Journal of Hydrology, pp. 19—32, 2013.
- [33] R. Merz, G. Blöschl: "Flood frequency regionalisation—spatial proximity vs. catchment attributes", Journal of Hydrology, pp. 283—306, 2005.

- [34] JEa Nash, J. V. Sutcliffe: "River flow forecasting through conceptual models part I—A discussion of principles", Journal of hydrology, pp. 282—290, 1970.
- [35] A.D. Nobre, L.A. Cuartas, M. Hodnett, C.D. Rennó, G. Rodrigues, A. Silveira, M. Waterloo, S. Saleska: "Height Above the Nearest Drainage a hydrologically relevant new terrain model", Journal of Hydrology, pp. 13—29, 2011.
- [36] E. M. Oldekop: On the evaporation from the surface of river basins. Trans. Meteorol. Obs. Univ., 1911.
- [37] T. B. M. J. Ouarda, C. Girard, G. S. Cavadias, Model parameter values along the quantiles. The vertical bars represent the 95% confidence interval computed for each parameter. The dots represent the calibrated parameters (equation <u>21</u>), the curves represent the constrained parameter after reduction of the number of degrees of freedom (equation <u>31</u>). B. Bobée: "Regional flood frequency estimation with canonical correlation analysis", Journal of Hydrology, pp. 157–173, 2001.
- [38] Ludovic Oudin, Vazken Andréassian, Charles Perrin, Claude Michel, Nicolas Le Moine: "Spatial proximity, physical similarity, regression and ungaged catchments: A comparison of regionalization approaches based on 913 French catchments", Water Resources Research, 2008.
- [39] Ludovic Oudin, Claude Michel, François Anctil: "Which potential evapotranspiration input for a lumped rainfall-runoff model?", Journal of Hydrology, pp. 275—289, 2005.
- [40] Ludovic Oudin: Recherche d'un modèle d'évapotranspiration potentielle pertinent comme entrée d'un modèle pluie-débit global. 2004.
- [41] Thomas Over, J. D. Riley, J. B. Sharpe, D. Arvin: Scientific InvestigFDC estimation models used in the study.ations Report, 2014.
- [42] M. Rianna, F. Russo, F. Napolitano: "Stochastic index model for intermittent regimes: from preliminary analysis to regionalisation", Natural Hazards and Earth System Science, pp. 1189—1203, 2011.
- [43] Ernesto Rodriguez, Charles S. Morris, J. Eric Belz: "A global assessment of the SRTM performance", Photogrammetric Engineering & Remote Sensing, pp. 249—260, 2006.
- [44] E. Sauquet, C. Catalogne: "Comparison of catchment grouping methods for flow duration curve estimation at ungauged sites in France", Hydrology and Earth System Sciences, pp. 2421—2435, 2011.

- [45] James Kincheon Searcy: Flow-duration curves. USGS Water Supply paper 1542-A, 1959.
- [46] R. D. Singh, S. K. Mishra, H. Chowdhary: "Regional flow-duration models for large number of ungauged Himalayan catchments for planning microhydro projects", Journal of Hydrologic Engineering, pp. 310—316, 2001.
- [47] Jon Olav Skøien, Günter Blöschl: "Spatiotemporal topological kriging of runoff time series", Water Resources Research, pp. 21, 2007.
- [48] David G. Tarboton, Daniel P. Ames: "Advances in the mapping of flow networks from digital elevation data", World water and environmental resources congress, pp. 20—24, 2001.
- [49] L. Turc: "Le bilan en eau des sols: relation entre les précipitations, l'évaporation et l'écoulement", Ann. Agron., pp. 491—595, 1954.
- [50] V. N. Vapnik, A. YA. Chervonenkis: "On the uniform convergence of relative frequencies of events to their probabilities", Theory of Probability and its Application, pp. 264—280, 1971.
- [51] Jean-Philippe Vidal, Eric Martin, Laurent Franchistéguy, Martine Baillon, Jean-Michel Soubeyroux: "A 50-year high-resolution atmospheric reanalysis over France with the Safran system", International Journal of Climatology, pp. 1627—1644, 2010.
- [52] R. M. Vogel, N. M. Fennessey: "Flow Duration Curves II: a review of Applications in Water resources planning", Journal of Water Resources Planning and Management, pp. 11, 1995.
- [53] I. K. Westerberg, J.-L. Guerrero, P. M. Younger, K. J. Beven, J. Seibert, S. Halldin, J. E. Freer, C.-Y. Xu: "Calibration of hydrological models using flow-duration curves", Hydrology and Earth System Sciences, pp. 2205—2227, 2011.
- [54] J. R. Ziegeweid, D. L. Lorenz, C. A. Sanocki, C. R. Czuba: Scientific Investigations Report, 2015.
- [55] W. R. van Esse, C. Perrin, M. J. Booij, D. C. M. Augustijn, F. Fenicia, D. Kavetski, F. Lobligeois: "The influence of conceptual model structure on model performance: a comparative study for 237 French catchments.", Hydrology and Earth System Sciences, pp. 4227—4239, 2013.

Table 1. Variability of hydro-climatic conditions observed on the catchment set on the overall period-of-record.

	Min.	Med.	Mean	Max.
Drainage support area - DSA [km ²]	0.04	0.57	3.11	195.4
Mean annual precipitation - P [mm/d]	1.80	2.86	3.04	6.08
Mean annual potential evapotranspiration - E0 [mm/d]	0.61	1.80	1.79	2.43
Area - A [km ²]	5	220	763	26940
Mean Altitude - Z [m]	43	392	562	2871
Q0.05 [mm/d]	0.01	0.16	0.22	1.29
Q0.25 [mm/d]	0.02	0.33	0.43	2.24
Q0.50 [mm/d]	0.08	0.68	0.84	4.03
Q0.75 [mm/d]	0.21	1.41	1.79	8.21

Table 2. FDC estimation models used in the study.

Name	Number of parameters	Abbreviation	Class of FDC estimation
Reference model	0	REF	Spatial-proximity
Complete model	396	CM396	Regression-based
Parsimonious model	10	PM10	Regression-based with the QS approach
Parsimonious model with residuals regionalization	10	PM10+REG	Regression-based with the QS approach and spatial proximity for model residuals

Table 3. Application of the QS approach to the catchments of the calibration set.

	α	β	Shape function	\mathbb{R}^2
Κ	-0.03	0.99	Inverse Gaussian	0.99
DSA	0.05	-0.12	Linear	0.85
Р	1.28	0.95	Linear	0.95
E0	-1.43	0.57	Linear	0.54

Figure 1. Location of the 521 catchments used in this study. Some of the catchments in our set are nested: the smaller catchments are represented on the top of the larger ones.

Figure 2. Use of the 17 features tested as prelude for this study. The plot presents the mean performance of all three-features combinations containing each feature.

Figure 3. Hydrological role of the drainage support area (DSA).

Figure 4. Model parameter values along the quantiles. The vertical bars represent the 95% confidence interval computed for each parameter. The dots represent the calibrated parameters (equation 2), the curves represent the constrained parameter after reduction of the number of degrees of freedom (equation 3).

Figure 5. Percentage of variance explained by the three features for the calibration set (R² of the CM396 model, equation 2).

Figure 6. Relative errors (in %) of the PM10 model on the calibrated set for: a) the 5th quantile (low flows), b) the 50th quantile (medium flows) and c) the 95th quantile (high flows).

Figure 7. Relative error distributions over the 260 catchments from the validation set for: a) CM396, b) PM10, c) PM10+REG and d) the reference model. The boxplots are defined by the first, second and third quartiles. The whiskers extend from the 10th percentile to the 90th percentile. The crosses indicate the mean relative error.

Figure 8. Cumulative distribution of the NSE values for the different models over the 260 catchments from the validation set (d) and corresponding density of the relative errors (a, b and c) as shown in Fig. 7. For d), the Friedman test showed that the distributions are significantly different except for REF and CM396.

Figure 1.



Figure 2



Figure 3







Figure 5



Figure 6







Figure 8

