The authors have presented a methodology to construct FDCs for missing portions of the hydrologic record at ungaged sites. The topic is very interesting and relevant. The authors have deeply considered all comments from the reviewers. To me, there is still a significant amount of work that needs to be completed before this manuscript can be published. There is nothing factually incorrect about the method, but I believe that it over-complicates itself and misses a few major elements. In particular, the derivation I have provided shows how this method is almost identical to previous publications. The revision should seek to explore the novelty of what is presented here.

We acknowledge Dr. Farmer for the time he dedicated to our manuscript and we are thankful for his suggestions and comments that helped us to improve our paper. The main changes made in the manuscript based on Referee's comments:

1. *We formulated the methodology and the underlying assumptions in a clearer manner in the Methodology section.*
2. *The German case study was removed to focus on the U.S.A. one.*

In the following, we are reporting a detailed point-by-point reply to Dr. Farmer's comments. Authors reply to Reviewer's comments are in red Italic.

As you can tell from the length of my review, I found this manuscript particularly engaging. (Sorry!) The main contribution of this work, in my opinion, is the hypothesis that the API can be used to produce the temporal sequence of non-exceedance probabilities of streamflow at a partially gaged site. In a related way, the authors hypothesize the existence of a unique, functional relationship between API and Q. This argues that the FDC is a function of weather, not just a function of the basin. These three points are phenomenally interesting to me, but the manuscript seems to get bogged down in the overly complex methodology rather than these hypotheses, which, when demonstrated, could be used to develop simplified methods. I would focus my revision on the novelty. I see the novelty as: (1) The FDC is a function of weather and location, not just location. (2) There exists a unique, one-to-one function relating API and streamflow. (3) For a given period, the temporal sequence of non-exceedance probabilities of streamflow is identical to the temporal sequence of non-exceedance probabilities of API. [NOTE: This is (I) on page 10.] (4) The temporal sequence of non-exceedance probabilities of API is identical at all sites in a region. [Note: This is (III) on page 10.] [NOTE FURTHER: Novelties 3 and 4 combine for an interesting result: QPPQ, described below.]

My main concerns are included below and pertain to the presentation and analysis of the methods. While these is nothing factually incorrect, the underpinnings and the implications of the method need to be explored and analyzed more fully.

To discuss the method, I'll begin by rephrasing the assumptions on page 10 in probabilistic terms, as they are central to this method:

I. $P(Q_{XY} < Q_{i,XY}) = P(API_{XY} < API_{i,XY})$, where X and Y designate the basin and time period, respectively. X and Y can therefore take a value of either D (donor) or T (target). The i indicates an arbitrary day of the record at the designated location and in the designated period.

II. $P(API_{DY} < API_{i,DY}) = P(Q_{TY} < Q_{i,TY})$ where all variables have been previously defined.

III. $P(API_{DY} < API_{i,DY}) = P(API_{TY} < API_{i,TY})$ where all variables have been defined previously.

$P(APIDY < APIi,DY) = P(APITY < APIi, TY)$.

As an aside: It should be noted that (II) is an implication of (I) and (III). That is, if (I) and (III) are true, then II must be true. To be explicit: $P(Q_{TT} < Q_{i,TT}) = P(API_{TT} < API_{i,TT})$ by (I), and $P(API_{TT} < API_{i,TT}) = P(API_{DT} < API_{i,DT})$ by (III), so $P(Q_{TT} < Q_{i,TT}) = P(API_{DT} < API_{i,DT})$. This last equality is (III).

The method proposed by the authors relies on (I) and (III) and the assumption that there exists a one-to-one function H_X at each site that maps API to Q.

The method proposed by the authors is as follows: (The method is accurate, though I'm repeating it to put us in the same language.)

Given, $P(Q_{TT} < X) = 10.11\%$, find X.

By (I), $P(Q_{TT} < X) = P(API_{TT} < Y)$.

By (III), $P(API_{TT} < Y) = P(API_{DT} < Z)$.

So we know that $10.11\% = P(Q_{TT} < X) = P(API_{TT} < Y) = P(API_{DT} < Z)$, and Y and Z are knowable by observation. The method presented here ignores Y and focuses on Z.

From the appendix example, Z = 37.72 mm.

Because $H_D(\ldots)$ is a unique, one-to-one function that ingests API and transforms it to Q we can say that we are curious to know what Q is produced by the API Z at the donor site.

The method answers this question by looking at the donor basin in the donor period. That is, by (III), $P(API_{DD} < Z) = P(API_{TD} < Y)$. (Note, because $H_D(\ldots)$ and $H_T(\ldots)$ are one-to-one, unique, time-invariant functions there exists a constant mapping from Y to Z based on (III).)

By (I) we can then say further that $P(API_{TD} < Y) = P(Q_{TD} < X)$. Now, because we are in the donor period, X is knowable by observation given that we know $P(API_{DD} < Z)$.

From the appendix example, $P(API_{DD} < 37.72mm) = [7.52\%, 7.54\%]$.

The value X is therefore knowable by observation such that, $P(Q_{TD} < X) = [7.52\%, 7.54\%]$.

Because $H_D(\ldots)$ and $H_T(\ldots)$ are one-to-one, unique, time-invariant functions, Y-->X regardless of time period. So X is the value we were seeking. By observation, the value of X is 3.21mm.

There is nothing factually inaccurate about this method, but it does over-complicate the problem at a partially gaged site. If the target site (T) is partially gauged, then the assumptions provided above make the use of a donor unnecessary. This would proceed as:

Given, $P(Q_{TT} < X) = 10.11\%$, find X.

By (I), $P(Q_{TT} < X) = P(API_{TT} < Y)$.

Because all APIs at the target site in the target period, we can observe the value of Y.

We can then go to the donor period at the target basin with a known Y.

Because we assumed there existed a one-to-one, unique, time-invariant function $H_T(\ldots)$, and all the APIs and Qs of the donor period at the target basin are known, we can interpolate the value Y along $H_T(\ldots)$ as observed in the donor period to determine X. So there is no need for the donor basin when the target is partially (and sufficiently) gaged.

The need for a donor basin will be essential when you move to a completely ungaged basin. However, with the ungaged basin, it is impossible to apply this method. This method relies on the donor period for an approximation of the FDC, though embodied through the function $H_T(\ldots)$, the proposed transferability of API across sites, and the proposed equality of probabilities for API and Q.

The reason that this method won't work at the fully ungaged site is because it relies on the donor period at the target site to estimate the FDC at the target site and uses the assumed $H_T(\ldots)$ to translate that across time periods (from donor to target).

To apply this method at the ungaged site, we will need an alternative method to approximate $H_T(\ldots)$. Arriving at this conclusion, it is obvious that this method, when applied to ungaged basins, is no different than the method presented by Smakhtin and his team in the mid to late

1990s. I will show this now, abbreviating their method as QPPQ:
Let's begin as we did before:
Given, $P(Q_{TT} < X) = 10.11\%$, find X.
By (I), $P(Q_{TT} < X) = P(API_{TT} < Y)$.
By (III), $P(API_{TT} < Y) = P(API_{DT} < Z)$.
Whereas earlier we stopped here, we can use (I) again to say that $P(API_{DT} < Z) = P(Q_{DT} < W)$.
We can then summarize this into a new implication of (I) and (III), namely $P(Q_{TT} < X) = P(Q_{DT} < W)$.
More generally, $P(Q_{DY} < Q_{i,DY}) = P(Q_{TY} < Q_{i,TY})$.
Now, if we had a first-order approximation of the relationship between Q and P at the target site, we could estimate X as a day in a simulated hydrograph.
The method presented by the present authors, uses API and $H_T(\ldots)$ in the donor site to make this approximation. Smakhtin et al., and following work, use something like a regional-regression FDC. The QPPQ school uses this first-order FDC to produce a sequence of daily flows based on the probabilistic implication described here ($P(Q_{DY} < Q_{i,DY}) = P(Q_{TY} < Q_{i,TY})$). The QPPQ stops here.
You could move further to take the simulated hydrograph and compute a new FDC.
Because of inaccuracies used to produce the first-order FDC at the target site, inaccuracies stemming from estimation error, for example, and the evolving sequence of probabilities in the target period (for example, it could be overly weighted toward high flows). There is no guarantee that the first-order FDC used in QPPQ will be the same as the FDC that would be built from the resulting simulated hydrograph.
Moving to this second-order FDC would produce a time-dependent FDC at the ungauged site, much in the spirit of the methods presented by these authors.
These two processes, that of simplifying the method presented for partially ungaged sites and the extension of the QPPQ method, show that the work presented by these authors misses the novelty of their own method. I believe that the work of these authors has merit and the potential to advance the field. With further consideration of the implications and theoretical underpinnings of the proposed methodology, this manuscript could prove impactful. I will follow now with a few less-major concerns I have with the manuscript.
At the opening, it was necessary for me to rephrase the assumptions outlined on page 10. This is because those assumptions are not clear, and the tests provided to demonstrate their veracity test something different. In (I) it is assumed that the CDFs "correlate", and the test shows a temporal correlation between nonexceedance probabilities. While this isn't incorrect, it is ambiguous. For example, one could equally compute the correlation of API and Q quantiles to say that the CDFs correlate, but that would be very different than what the authors are seeking to assume or prove. Similarly, (III) states that the CDFs of API are identical across sites, but the proof shows only that the exceedance probabilities correlate in time. Identical CDFs would have the same probabilities associated with the same values of the proxy. These assumptions need to be less ambiguous.
As notes above, these assumptions are also redundant. Really, only (I) and (III) are needed, as (II) is a natural extension. Another natural extension, as shown above, is $P(Q_{DY} < Q_{i,DY}) = P(Q_{TY} < Q_{i,TY})$. This last one raises a point of confusion in the manuscript.

We thank the reviewer for his interest in our paper. We have substantially modified the Methodology section to provide a clearer explanation of the approach and of the underlying assumptions. In the following, between quotes, we report the description that is also in the manuscript.

"Formally the basic hypotheses of this paper are:

- Flow duration curves are not invariant properties of basins but are the product of basin, weather and human interactions. In this investigation we do not consider the human interactions.
- Precipitation is the most important influencing factor on discharge.
- Basins delay the reaction on precipitation therefore the API is a better indicator for the influence of precipitation on discharge.
- We assume that discharge and API are changing in a similar way for longer time periods.

Let $F_{A,Ti}(q)$ be the distribution of daily discharges at basin A and time period T_i (flow duration curve for the selected time period) and $G_{A,Ti}(a)$ be the distribution of daily API at basin A and time period T_i.

The transformation from T_i to T_j provides an estimated $F^*_{A,Tj}(q)$:

$$F^*_{A,T_j}(q) = G_{A,T_j}(G^{-1}_{A,T_i}(F_{A,T_i}(q))) \tag{7}$$

This is a quantile/quantile transformation.

$$F_{A,T_j}(q) = F_{A,T_j}(F^{-1}_{A,T_i}(F_{A,T_i}(q))) \tag{8}$$

The basic question can be written in the form of the following equation

$$G_{A,T_j}(G^{-1}_{A,T_i}(p)) \approx F_{A,T_j}(F^{-1}_{A,T_i}(p)) \tag{9}$$

That can be summarized in the following question: do the percentiles of the API change in the same way as those of the discharge? Note that if the relationship between API and discharge is a good one then the two sides are nearly equal. Even a weak relationship can do a good job if the errors are independent and the sign of the change is correct. Figures 4 and 5 show the difference between the real change in percentiles and that obtained by using the API for different time periods according Eq.9. Note that the assumption that the FDC is time invariant would imply that the lines for the discharge are on the diagonal.

The correlation between API and discharge is around 0.6, but the transformations are quite similar and the API based transformation delivers good FDCs.

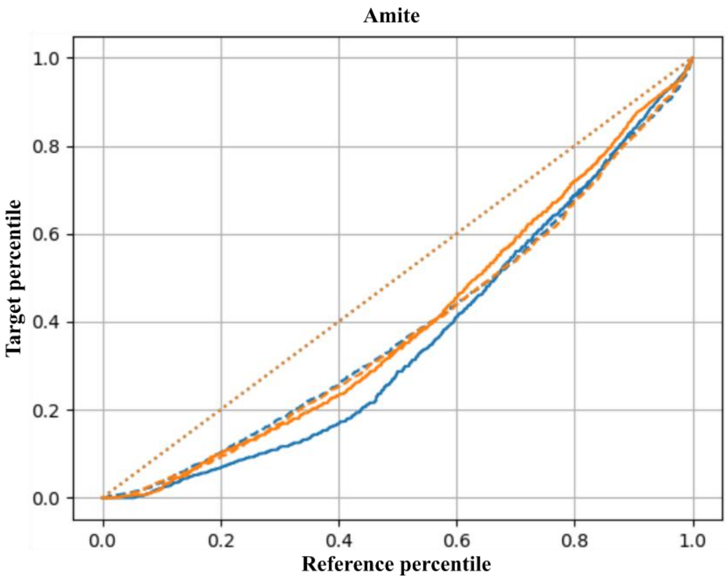


Figure 4. The transformation functions of the 1951-1960 FDC (solid) and API (dashed) to the target period 1971-1980 (blue) and 1981-1990 (orange) for Amite. For the sake of comparison, the diagonal is dotted in orange.

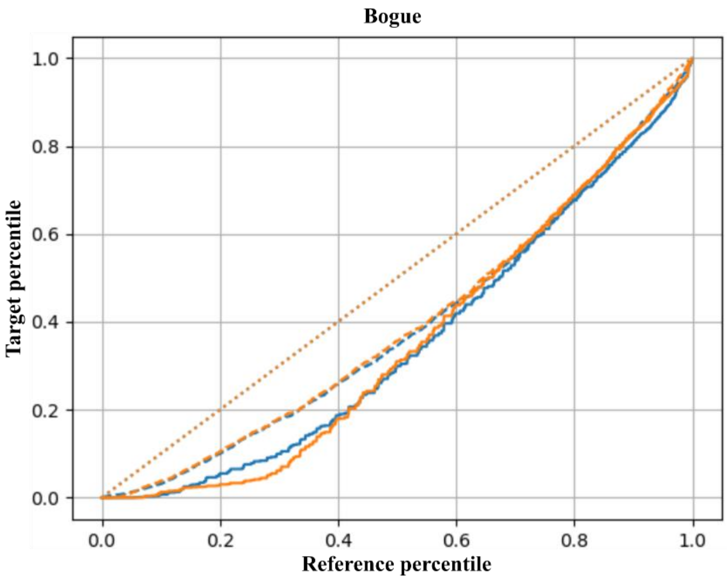


Figure 5. The transformation functions of the 1951-1960 FDC (solid) and API (dashed) to the target period 1971-1980 (blue) and 1981-1990 (orange) for Bogue. For the sake of comparison, the diagonal is dotted in orange.

If API is changing continuously in space then one can use the change of the FDC of a different location B for the estimation:

$$F^{**}_{A,T_j}(q) = F_{B,T_j}(F^{-1}_{B,T_i}(F_{A,T_i}(q))) \qquad (10)”$$

On page 13, line 34, it is stated that “the FDC of a donor site cannot be transferred to another site”. First, it is unclear what is mean by “the FDC”: Do you mean the FDC in its entirety, including probabilities and quantiles, or the correlation as “proven” for (I), (II) and (III). In any case, the

extension that $P(Q_{DY} < Q_{i,DY}) = P(Q_{TY} < Q_{i,TY})$ contradicts this statement. Furthermore, the authors then use this extension to transfer the FDC of a donor to a target in the German catchments. This confusion should be clarified. (In point of fact, I recommend removing the German catchments altogether.)
We acknowledge the reviewer for his comment. We deleted the sentence as it was leading to confusion. We also removed the German case study for the sake of simplicity as suggested.

A minor aside, the authors have no clearly defined that a partially gaged basin is. Is it any site with at least one value of streamflow? Is there some threshold for sufficiency? The method, whether as presented by the author or simplified here, requires a substantial donor period at the target catchment, but it might be important to define some length of record that would be required.
A site is defined partially gauged when a limited amount of flow data is available. As we want to assess the performance of the method depending on the extension of the period with missing data, we have chosen time windows with different length as target periods and we have shown the performance of the method depending on this length. The minimum tested length is one-year long as it was used to retrieve the annual FDC. We clarified it in the Conclusion section.

Stepping away from the methodology for a moment, I still struggled with the structure of the manuscript. Here, I have two recommendations. First, I suggest removing the section on energy and water limited catchments (2.3). We know that there are different limiters, and we know this is important, but I don't think this section adds to the manuscript. This is evidenced by the fact that the information in this section is never mentioned again in the manuscript: Not to explain the results or adapt the methodology. In a similar vein, I suggest removing the presentation and analysis of the German catchments. The results for the German catchments, starting on page 18, provide little value. The methodology is substantially different (confusingly using a streamflow proxy, as described above). The authors provide little discussion as to what new information is presented by these catchments. Obviously, with some additional discussion, both sections could prove useful, but it may be simpler just to remove them from the manuscript.
We thank the reviewer for the comments, we deleted the sections related to the German catchments to keep the focus on the US catchments. On the other hand, we decided to keep the section on energy and water limited catchments as it involves a discussion on the role of the weather on annual runoff variability which is the leading message of the paper. We have added a discussion about energy and water limited catchments in the Conclusions section.

While this revision did a great job of supporting its assumptions, I feel that a basic one was left out. The assumption that these exists the function $H_T(\ldots)$ (page 9, line 7). As shown above, this is the crux of the methodology. I think it would be worthwhile to show that the interpolation of values implies that this relationship (between Q and API) is site dependent and time independent through a cross-validation exercise.
Finally, before considering minor comments, I was quite surprised that the authors did not address the large literature on the appropriateness of donor catchment selection. Instead, the examples presented here use the same donor for all gages and treat all gages as donors. Could it be that the method is sensitive to donor selection? In truth, the partially gaged method is probably not sensitive to donor selection, as demonstrated by the simplification where a donor is not needed, but the ungaged approach (QPPQ) is certainly dependent on donor selection.
We thank the reviewer for the comment, the most important part is that both API and discharge change into the same direction, and the errors of the relationship are more or less independent so that they cancel out over time. The assumptions are now formulated in a clearer manner in the equations.

MINOR COMMENTS:
Page 1, line 10: It seems odd to say that FDCs are “set up”. They are a product of the record, they aren’t established. Maybe something like, “The FDC of streamflow at a specific site provides knowledge on the distribution and characteristics of streamflow at that site.”
We rephrased as suggested.

Page 1, line 12: “In spite of its importance,…”
We rephrased as suggested.

Page 1, line 14: What do you mean by partially gaged? The ability to build an FDC depends on this definition. Certainly, a site that has a record from 1980-2000 and 2002-2018 is partially gauged, but it still has more than enough information to build an FDC. I see your point that it is weather dependent, but it still leaves the definition of partially gaged ambiguous.
The idea of this paper is that it is possible to retrieve the FDC of a basin using the precipitation; the unknown FDC could be annual or could be built with a longer data set. A site is defined partially gauged here if there is a lack of records in a given amount of time, this time could be a year or more. In the paper we show how the method performs choosing time windows with different length as target periods. Please, see also comment above regarding “partially gauged site” definition.

Page 1, line 14: “…among the other streamgages.”
We were referring to other methods, it is clearer now.

Page 2, line 3: I'm not sure that FDCs provide information on the severity because they don't talk about duration (length) of drought. So two sites could have a similar frequency of low flows, but one might see all low-flows consecutively (long drought), while the other might see intermittent lows (episodic drought). One could argue that drought is more severe in the former case.
We removed the sentence.

Page 2, line 12: “…to meet intake requests…”
Page 2, line 15, “As the FDC…”
We rephrased as suggested.

Page 2, line 27: “In spite of its importance, the FDC…”
We rephrased as suggested.

Page 3, line 27, “In the following…”
We removed the sentence.

Page 3, line 28, “…divided into water…”
We removed the sentence.

Page 3, line 28: Are you making a meaningful distinction between basins and catchments? If so, explain. If not, please use consistent terminology throughout.
We checked the consistency and we are using basin now.

Page 3, line 40: Check citation style.
Page 3, line 40: Revise to “The soil type can strongly affect the impact of climate on the water balance.”
Page 3, line 41: I think the terms are karst and non-karst and should be throughout.

Page 4, line 2: "…karst regions makes it difficult to transfer information from ..."
Page 4, line 13: Please provide citation for degree-day approach.
As we removed the German case study, we removed these sentences.

Page 4, line 15: Keep in mind for formatting: This is an odd page break for a header.
It is a mistaken format indeed.

Page 5, line 21: "…the catchment allows for infiltration and stores…"
Page 5, line 21: "…runoff production declines since…"
Page 5, line 23: "However, the climatic timing…"
Page 5, line 24: "They show that the difference…"
We rephrased as suggested.

Page 8, line 30: I think maybe an extra sentence or two of explanation is needed here. Why would you expect two different time periods at the same site to be independent (as assumed by the KS test)? I see the point about auto-correlation, but I find it hard to imagine that the distribution of Fall 2012 streamflow and Fall 2013 stream flows are wholly independent at the same site. I think I follow your logic, and I would love to see it spelled out more fully. It may be tied to ambiguous definition of partial gaging.

The KS test is designed to compare two distribution functions. The question is not whether the sample corresponding to different time periods are independent but whether they correspond to a different distribution. Anyhow the long memory is relatively low, and we consider full years thus annual cycles do not have an influence on our results. We clarified it in the text.

Page 8, line 38: "…Figure 4 shows the magnitude of the difference between…"
Page 9, line 10: "…hydrological modelling as modelling often introduces additional errors…"
Page 9, line 11: "…errors and may be…"
We rephrased as suggested.

Page 9, line 15: Ask, defined above, denotes location, what location is being addressed here?
We rephrased to introduce the location.

Page 10, line 10: This paragraph cites discussion about why the alpha should decay with time. Does it not decay here? It's fixed at 0.85?

We have deleted the sentence that was leading to confusion. In this paragraph we describe that alpha may assume values ranging from 0 to 1, depending on the type of API's memory one wants to capture. Once that the value of alpha is chosen, this is constant throughout the time. Here it is set equal to 0.85.

Page 10, line 17: See discussion above. Here you need to clarify that you are talking about the correlation between $p(Q<Qi)$ and $P(API < APIi)$ for the day i. This is different than the correlation of Qi and APIi of rank i.
Page 10, line 27: Are these Pearson correlations? Please specify?
We consistently revised the methodology section.

Page 13, line 5: Here you are using "," to denote a decimal point and "." to denote figures to the right of the decimal. This is not the style of the journal. I suggest, "X.yz, where y and z are the digits to the right to the decimal" Perhaps still better: In binning by percentiles, all percentages

were rounded down to the nearest whole number.
We rephrased as suggested.

Page 13, line 9: Shouldn't there be an absolute value for this statement to be true? Bias is defined here as an average. If, for example, we looked at three days producing these values of the thing in parenthesis: -3, 0, 3. In this case, the BIAS would be 0, but this is certainly not a perfect fit.
We have rephrased the description of BIAS.

Page 13, line 13: "…and so is the MAE. It …"
We rephrased as suggested.

Page 13, line 29: I think it would be better to show the p-value of D*. Looking at Figure 8 (page 14, line 12), the D may be bigger, but it may not be as significant because of changing sample size.

There was a typo regarding the D values that are now correct. The p-value is always zero but for a few cases that are reported in the captions and in the text.*

Page 17, line 13: I find it very confusing that the color scale changes from figure to figure (e.g. 9 and 10). The change makes it hard to compare across figures.

We have corrected the scale so that it is easier to compare now, we also carefully checked the figures since there was an issue regarding the colors of Figure 10 (now Figure 11). Now it is fixed.

Page 18, line 6: Here the proxy is donor streamflow. I discuss this above, but this strikes me as odd when page 13, line 34 seems to say you should not do this. (As you know, I think it might be best to just remove this section.)
We removed it as suggested.

Page 19, line 2: The correlation between the donor and the target is problematic and cannot be known a priori (line 12). This needs to be contextualized in an ungaged example that is not discussed here. The ambiguity here makes the conclusion on line 39 of page 40 rather tenuous.

We removed the German case study as suggested and so this paragraph, besides the methodology is now presented in a clearer manner.

Page 23, line 18: The probability statement should be in terms of API, right?
Yes, it is. Thank you, it was a typo.

A methodology to estimate flow duration curves at partially ungauged basins

Elena Ridolfi[1,2], Hemendra Kumar[3], and András Bárdossy[4]

[1]Department of Earth Sciences, Uppsala University, Uppsala, Sweden

[2]Centre of Natural Hazards and Disaster Science, CNDS, Sweden

[3]Department of Biosystems Engineering Auburn University, Alabama, USA

[4]Institute of Hydraulic Engineering, University of Stuttgart, Stuttgart, Germany

Correspondence: András Bárdossy (andras.bardossy@iws.uni-stuttgart.de)

Abstract. The Flow Duration Curve (FDC) ~~set up~~ of streamflow at a specific site has a key role to the knowledge ~~of~~on the distribution and characteristics of streamflow ~~characteristic~~ at that site. The FDC gives information on the water regime providing information to optimally manage the water resources of the river. ~~Spite~~In spite of its importance, because of the lack of streamflow gauging stations, the FDC construction can be a not straightforward task. In partially gauged ~~catchments~~basins, FDCs are usually built using regionalization ~~methods~~ among the other~~s~~ methods. In this paper we show that the FDC is not a characteristic of the basin only, but of both the basin and the weather. Different weather conditions lead to different FDC for the same catchment. The differences can often be significant. Similarly, the FDC built at a site for a specific period cannot be used to retrieve the FDC at a different site for the same time window. In this paper, we propose a new methodology to estimate FDCs at partially gauged basins (i.e., target sites) using ~~discharge and~~ precipitation data gauged at another ~~catchment~~basin (i.e., donor ~~catchment~~site). The main idea is that it is possible to retrieve the FDC of a target period of time using the data gauged during a given donor time period for which data are available at both target and donor sites. To test the methodology, several donor and target time periods are analyzed and results are shown for ~~two~~ different ~~case study areas~~sites in the USA. The comparison between estimated and actually observed FDCs show the reasonability of the approach especially for intermediate percentiles.

1 Introduction

A duration curve is a function that associates to a specific variable its exceedance frequency. Specifically, in hydrology a Flow Duration Curve (FDC) is a function describing the flow variability at a specific site during a period of interest. It represents the streamflow values, gauged at a site, against their relative exceedance frequency. An empirical long-term FDC is the complement of the empirical cumulative distribution function of streamflow values at a given time resolution based on the complete streamflow record available for the basin of interest (Castellarin et al., 2007). FDCs are built as explained in the followings:

– rank the streamflow values in descending order;

– plot the sorted values against their corresponding frequency of exceedance.

The duration d_i of the *i-th* sorted observation is its exceedance probability P_i. If P_i is estimated using a Weibull plotting position (Weibull, 1939), the duration d_i for any q_i (with $i = 1;\ldots;N$) is

$$d_i = P(Q < q_i) = P_i = \frac{i}{N+1}, \qquad (1)$$

where N is the length of the streamflow series and q_i is the *i-th* sorted streamflow value.

The FDC provides historical information on the water regime~~: on the severity of the droughts and on the magnitude of high flows~~. Several time resolutions of streamflow data can be used to build the FDC: annual, monthly or daily. However, the finer is the resolution, the higher is the information provided by the FDC about the hydrological characteristics of the river (Smakhtin, 2001). FDCs may be built either on the basis of the whole available record period (Vogel, 1994); or on the basis of all similar months (Smakhtin et al., 1997); or on the basis of a specific month.

In one curve, the FDC condenses a wealth of hydrologic information that can be easily accessed. Because of the key role of runoff variability to both water resources management and environmental health maintenance, FDC is used in a large variety of applications as reported by Vogel (1994). For instance, FDC can quantify the capacity of the river to meet intake requests as it provides information about the reliability of the water resource for water abstraction activities (Dingman, 1981). It is at the base of hydropower plants design as they are used to determine the hydropower energy potential, especially for run-of-river plants (Hänggi and Weingartner, 2012; Blöschl et al., 2013). As the FDC is a key signature of runoff variability, it can be used to assess the impact of changes in a catchment. To this end, through the FDC, Vogel et al. (2007) introduced the indicators of the eco-deficit and eco-surplus. Moreover, the FDC can be used to define and investigate low flows (Smakhtin, 2001). The knowledge of the streamflow characteristics is also relevant for stream water quality studies, for instance, to regulate the proper threshold for chemical concentration and load (Bonta and Cleland, 2003). FDC has a further application in model calibration. This application is based on the replication of the flow frequency distribution rather than of the simulation of the hydrograph (Yu and Yang, 2000; Westerberg et al., 2011). Other applications are related to irrigation planning (Chow, 1964); schedule optimal flow release from reservoirs (Alaouze, 1991); basins afforestation (Scott et al., 2000); investigation of the effects on flows regime due to ~~catchments~~basins vegetation change (Brown et al., 2005).

~~Spite~~In spite of ~~FDC~~its importance, the FDC is affected by the lack of data in ungauged and poorly gauged basins. Many authors dealt with the issue of FDC prediction at ungauged or partially gauged locations through regional regression (e.g., Fennessey and Vogel, 1990; Mohamoud, 2008; Rianna et al., 2011, 2013; Castellarin et al., 2013; Pugliese et al., 2016) and geostatistical interpolation (e.g., Pugliese et al., 2014). Ganora et al. (2009) developed a methodology to estimate FDC at ungauged sites based on distance measures that can be related to the catchment and the climatic characteristics. Spatial non-linear interpolation methods were developed by several scholars (e.g., Archfield and Vogel, 2010; Mohamoud, 2008; Hughes and Smakhtin, 1996; Farmer et al., 2015). Worland et al. (2019) presented a method involving the use of the copula function. Hughes and Smakhtin (1996) proposed a method to extend and/or filling in daily flow time series at a site using monthly FDCs of the target site itself. These monthly FDCs should be recorded during a donor period or retrieved using different methods such as (i) regionalization of FDCs based on available observed records from several adjacent gauges (Smakhtin et al., 1997) or (ii) conversion of FDCs calculated from monthly data into 1-day FDCs (Smakhtin, 1999). Since the main limitation of the approach proposed by Hughes and Smakhtin (1996) is that it is based entirely on observed flow records, later, Smakhtin and Masse (2000) proposed a further development, which uses the current precipitation index (CPI) of the donor site to extend the daily hydrograph at the target site. The major assumption is that both the CPIs occurring at donor sites in a reasonably close proximity to the target site and target site's flows themselves correspond to similar percentage points on their respective duration curves. On the other hand, the basic assumption of the spatial interpolation algorithm proposed by Hughes and Smakhtin (1996) is that flows occurring simultaneously at sites in reasonably close proximity to each other correspond to similar probabilities on their respective flow duration curves. On the contrary, one important message of our paper is that FDCs can be very different from time period to time period both at the site itself and at pairs of sites as a long term change in the weather effects the FDCs. Therefore, our approach is based on the concept that proximal sites do not share similar FDCs. This will be demonstrated in the paper applying a two-sample Kolmogorov-Smirnov test to pairs of stations. The usual assumption that they and the related indices are characteristic for the ~~catchment~~basin is not true. Therefore, the FDCs built at a given location for different periods cannot be regarded as the same distribution. It is not possible to determine a unique distribution and therefore a unique set of parameters. The same results from the analysis of FDCs built in two different ~~catchments~~basins. It is not possible to develop relations between parameters of the basin and characteristics of the FDC to yield synthesized FDCs in locations where flow data are not available, as done for instance by Quimpo et al. (1983). These issues have a key role especially when dealing with ungauged basins.

The main idea underlying our work is to build the FDC at a target site using a filter, which relates the distributions of the discharge and the precipitation. As the weather is the main driver of annual runoff variability, we propose a transformation driven by the weather. The paper is organized as follows. First, the case study~~ies~~ ~~are~~ is presented and ~~catchments~~basins are grouped into energy- and water-limited ones. Then, the Kolmogorov-Smirnov test is carried out on pairs of FDCs to assess whether these curves can be regarded as the same distribution. Second, the methodology is presented ~~and~~together with the underlying assumptions. Then, the approach is applied to a set of ~~catchments~~basins located in ~~Germany and in U.S~~the case study area. Finally, results are shown and discussed.

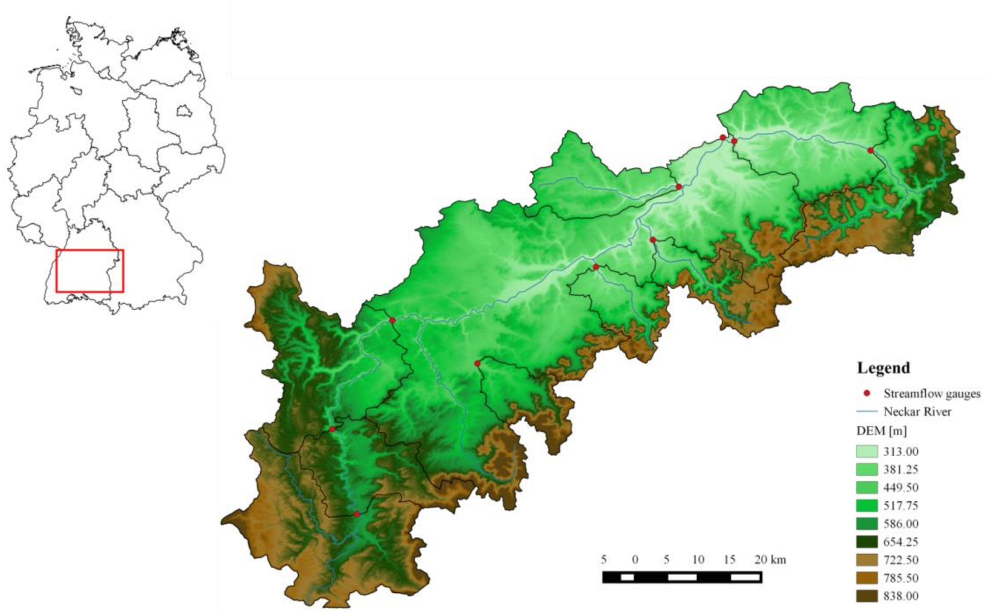
2 Case study area

~~The methodology was applied to several catchments located in two different areas. Ten basins are located in the upper Neckar River basin (Germany), while ten basins are located on the Gulf coast of the USA. In the followings, the two case study areas are presented. Since the procedure is based on the climatological characteristics of basins, catchments will be divided in water and energy limited ones.~~

~~2.1 Upper Neckar catchments (Germany)~~

~~This study uses data from ten sub-catchments belonging to the Upper Neckar River basin, south-west Germany. The Neckar is a tributary of the Rhine, it springs at an altitude of 706 m a.s.l. and it is 367 km long, Figure 1. The Upper Neckar catchment lies in between the Black Forest and Schwäbische Alb in the Baden-Württemberg region. The basin has an area of 4000 km^2, its elevation ranges from~~

~~about 240 m a.s.l. to around 1010 m a.s.l., with a mean elevation of 548 m a.s.l. (Singh et al., 2012). The sub-catchments are characterized by a drainage area ranging from around 120 km^2 to about 4000 km^2. The region is characterized by warm summers and mild winters (Samaniego, 2003). In the Upper Neckar catchments, the main geological formations originated in the Triassic and Jurassic periods. The main formations are composed of altered keuper, claystone-jura, claystone-keuper, limestone-jura, loess, sandstone and shelly limestone (Muschelkalk), Samaniego (2003). The effect of soil type can strongly modify the impact of climate on the water balance. For instance, karstic and non-karstic catchments are characterized by very different water balances, since an underground karstic catchment is very different from its overground catchment. The presence of karstic regions makes difficult the transfer of information from precipitation to discharge data in the same basin and from a karstic basin to a not karstic one. Approximately 35% of the basin has karstic formations (Samaniego et al., 2010).~~



~~**Figure 1**. Streamflow gauges (red circles) used to test the methodology in the corresponding catchments located on the Upper Neckar River, Germany.~~

~~The mean daily discharge, precipitation, and evapotranspiration, the minimum and maximum daily temperature are available for each sub-catchment for the period 1961-1990. Basins characteristics are presented in Table 1; for more details on this study area, please refer to Samaniego (2003) and Bárdossy et al. (2005). Snow effects are considered by a simple snow accumulation and snowmelt model using a degree day approach. This allows to convert snow to a daily liquid.~~

~~**Table 1.** Study area: upper Neckar catchments in south-west Germany~~

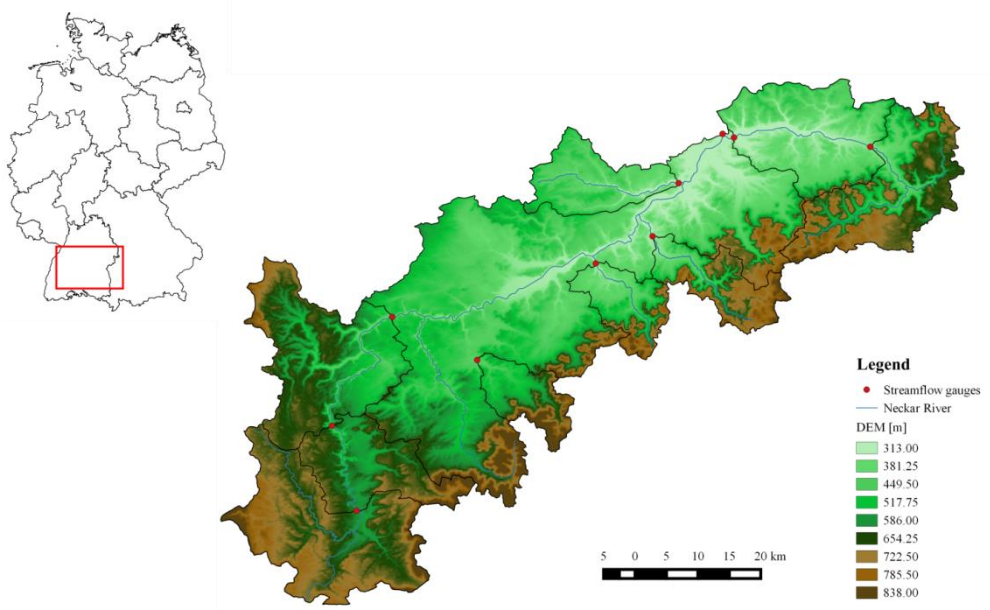
~~Catchment~~	~~Area~~	~~Drainage area~~	~~Elevation~~	~~Slope~~	~~Annual discharge~~	~~Annual precipitation~~
	~~km^2~~	~~km^2~~	~~m~~	~~degree~~	~~mm~~	~~mm~~
~~Rottweil~~	~~456~~	~~456~~	~~555–1010~~	~~0–34.2~~	~~352.7~~	~~976~~
~~Obendorf~~	~~235~~	~~691~~	~~460–1004~~	~~0–44.2~~	~~360.5~~	~~953~~
~~Horb~~	~~427~~	~~1118~~	~~383–841~~	~~0–48.9~~	~~417.5~~	~~1158~~

~~Rangendingen, Starzel~~	~~118~~	~~118~~	~~421–954~~	~~0–36.9~~	~~347.4~~	~~905~~
~~Wannweil, Echaz~~	~~135~~	~~135~~	~~309–862~~	~~0–45.9~~	~~654.1~~	~~877~~
~~Riederich, Erms~~	~~170~~	~~170~~	~~317–865~~	~~0–49.4~~	~~556.5~~	~~956~~
~~Oberensingen, Aich~~	~~175~~	~~175~~	~~278–601~~	~~0–27.1~~	~~234.3~~	~~762~~
~~Suessen, Fils~~	~~340~~	~~340~~	~~360–860~~	~~0–49.3~~	~~547.2~~	~~1003~~
~~Plochingen, Fils~~	~~352~~	~~692~~	~~252–785~~	~~0–39.7~~	~~446.6~~	~~936~~
~~Plochingen, Neckar~~	~~473~~	~~3962~~	~~241–871~~	~~0–45.8~~	~~397.2~~	~~863~~

The methodology was applied to several ~~catchments located in two different areas. Ten~~ basins ~~are located in the upper Neckar River basin (Germany), while ten basins are located on the Gulf coast of the USA. In the followings, the two case study areas are presented. Since the procedure is based on the climatological characteristics of basins, catchments will be divided in water and energy limited ones.~~

~~2.1 Upper Neckar catchments (Germany)~~

~~This study uses data from ten sub-catchments belonging to the Upper Neckar River basin, south-west Germany. The Neckar is a tributary of the Rhine, it springs at an altitude of 706 m a.s.l. and it is 367 km long, Figure 1. The Upper Neckar catchment lies in between the Black Forest and Schwäbische Alb in the Baden-Württemberg region. The basin has an area of 4000 km^2, its elevation ranges from about 240 m a.s.l. to around 1010 m a.s.l., with a mean elevation of 548 m a.s.l. (Singh et al., 2012). The sub-catchments are characterized by a drainage area ranging from around 120 km^2 to about 4000 km^2. The region is characterized by warm summers and mild winters (Samaniego, 2003). In the Upper Neckar catchments, the main geological formations originated in the Triassic and Jurassic periods. The main formations are composed of altered keuper, claystone-jura, claystone-keuper, limestone-jura, loess, sandstone and shelly limestone (Muschelkalk), Samaniego (2003). The effect of soil type can strongly modify the impact of climate on the water balance. For instance, karstic and non-karstic catchments are characterized by very different water balances, since an underground karstic catchment is very different from its overground catchment. The presence of karstic regions makes difficult the transfer of information from precipitation to discharge data in the same basin and from a karstic basin to a not karstic one. Approximately 35% of the basin has karstic formations (Samaniego et al., 2010).~~



~~**Figure 1**. Streamflow gauges (red circles) used to test the methodology in the corresponding catchments located on the Upper Neckar River, Germany.~~

~~The mean daily discharge, precipitation, and evapotranspiration, the minimum and maximum daily temperature are available for each sub-catchment for the period 1961-1990. Basins characteristics are presented in Table 1; for more details on this study area, please refer to Samaniego (2003) and Bárdossy et al. (2005). Snow effects are considered by a simple snow accumulation and snowmelt model using a degree day approach. This allows to convert snow to a daily liquid.~~

~~**Table 1.** Study area: upper Neckar catchments in south-west Germany~~

~~**2.2 USA catchments**~~

~~The catchments on the Gulf coast of the USA are~~ located in three different States on the Gulf coast of the USA: Florida, Louisiana and Texas, Figure ~~2~~1. These basins were selected because they are characterized by a mild climate and therefore, no snow events have been recorded, allowing us to neglect the snow melting effect. Daily streamflow discharge and precipitation values are available for each ~~catchment~~basin for different time windows, Table ~~2~~1.

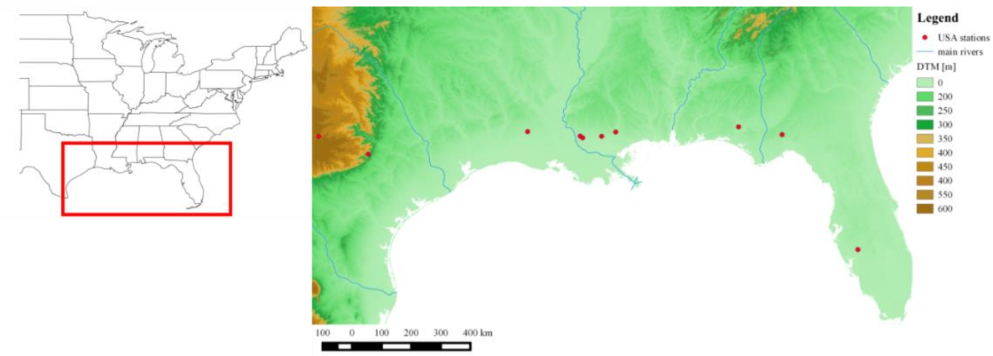


Figure ~~2~~1. Streamflow gauges (red circles) used to test the methodology in the corresponding USA ~~catchments~~basins.

Daily streamflow discharge data were originally provided by the United States Geological Survey (USGS) gauges, while mean areal precipitation and climatic potential evaporation were supplied by the National Climate Data Center (NCDC) at daily resolution. The data set is a subset of the Model Parameter Estimation Experiment (MOPEX) database, used for hydrological model comparison studies (Duan et al., 2006) and for simultaneous calibration of hydrological models (Bárdossy et al., 2016).

Table ~~2~~1. US case study area: streamflow gauges and corresponding ~~catchments~~basins characteristics

Station name	Drainage Area *km²*	Mean elevation *m*	Mean slope -	Mean discharge *mm*	Mean annual precipitation *mm*	Available record -
Peace River At Arcadia, FL	3540.53	32.3	0.3	257.4	1296.2	1948-2001
Ochlockonee River Nr Havana, FL	2952.6	75.6	1.8	322.6	1366.7	1948-2001
Choctawhatchee River at Caryville, FL	9062.41	92.2	3.2	540.8	1464.7	1948-1994
Bogue Chitto River near Bush, LA	3141.67	101.6	1.8	579.2	1637.1	1948-1999
Tangipahoa River at Robert, LA	1673.14	76.9	1.6	635.2	1682	1948-1999
Comite River near Comite, LA	735.56	59.6	1.1	595.9	1644.2	1948-1999
Amite River near Denham Springs, LA	3315.2	75.6	1.3	584.1	1647.9	1948-1999
Calcasieu River near Oberlin, LA	1950.27	62.2	1.1	502.9	1558.9	1948-1986
Llano Rv near Junction, TX	4807.04	670.9	3.4	34.8	645.8	1948-1988
Blanco Rv at Wimberley, TX	919.45	417.3	5.2	140.6	896.7	1948-2001

2.~~3~~ 1 Energy and water limited ~~catchments~~basins

Annual runoff variability is driven by the relative availability of water (i.e., precipitation) and energy (i.e., evaporation potential). Therefore, the weather is the most important driver of annual variability (Blöschl et al., 2013). Much of the annual runoff variability can be explained observing the different availability of water and energy. For instance, if more water arrives to the ~~catchment~~basin than energy can remove through evaporation, the annual runoff will be high. Moreover, in this case the relationship between runoff and precipitation will be more linear than when more energy is available to evaporate the water. On the other hand, in an arid region, the aridity of the climate determines a high inter-annual runoff variability because of the non-linear relationship between runoff and precipitation. Therefore, differences in water and energy availability cause differences in annual runoff variability. However, additional factors such as differences in seasonality and precipitation must be considered (Jothityangkoonad and Sivapalan, 2009). The relative availability of water and energy can be described through the Budyko curve (Budyko, 1974). The curve plots the ratio between mean annual actual evaporation and mean annual precipitation as a function of the ratio between mean annual potential evaporation and mean annual precipitation~~ (i.e., the aridity index)~~. Therefore, it defines a similarity index (i.e., the aridity index) to express the availability of water and energy, and thus bolsters the classification of hydrological sceneries into various degree of aridity. The Budyko curve represents the effects of water and energy availability on annual runoff variability. Moreover, it provides indication about the synchrony of evaporation and precipitation. For instance, where precipitation and evaporation are in phase, runoff production ~~reduces~~declines since the ~~catchment infiltrates~~basin allows for infiltration and stores water and vice versa. Many regions range from in phase to out of phase because of the strong seasonality of climate forcing. However,~~ also~~ the climatic timing can influence runoff variability as presented by Montanari et al. (2006). They ~~shown~~show that the difference in annual runoff between two years with equivalent annual precipitation was of 100% in a monsoonal area of Northern Australia because during the wet year the precipitation occurred during the wet season, i.e., when the potential evaporation was smaller. In this framework, it is important to understand the behavior of the ~~catchments~~basins under analysis. To this end, we analyzed the mean annual runoff coefficient, the annual precipitation and the annual evapotranspiration against the annual mean temperature. This analysis is essential to understand the causal processes leading to the long-term mean and variability of runoff as also described in McMahon et al. (2013). The mean annual runoff coefficient is defined as:

$$\mu_R = \frac{\overline{Q_{yr}}}{\overline{P_{yr}}}, \qquad (2)$$

where $\overline{Q_{yr}}$ is the annual discharge volume and $\overline{P_{yr}}$ is the annual precipitation volume.

Results show that ~~catchments~~basins have two different behaviors: precipitation, evapotranspiration and runoff have either a positive or a negative correlation with the air temperature. In the former case the evapotranspiration is limited by the available water, which happens in water-limited ~~catchments~~basins; in the latter the evapotranspiration is limited by the available energy which happens in energy-limited ~~catchments~~basins. For instance, measurements at Peace River (LA) suggest that the ~~catchment~~basin is balanced between energy and water limitation by the correlation criterion, Figure ~~3~~2 upper panel. While Ochlockonee River (FL), Amite River near Denham Springs (LA) and Bogue Chitto River (LA) are energy-limited. Results for Amite River are consistent with what found by Carrillo et al. (2011). Since it is not possible to infer discharge values of a water-limited ~~catchment~~basin from the data set of an energy-limited one, analysis have been carried out on climatically homogeneous sets of basins.

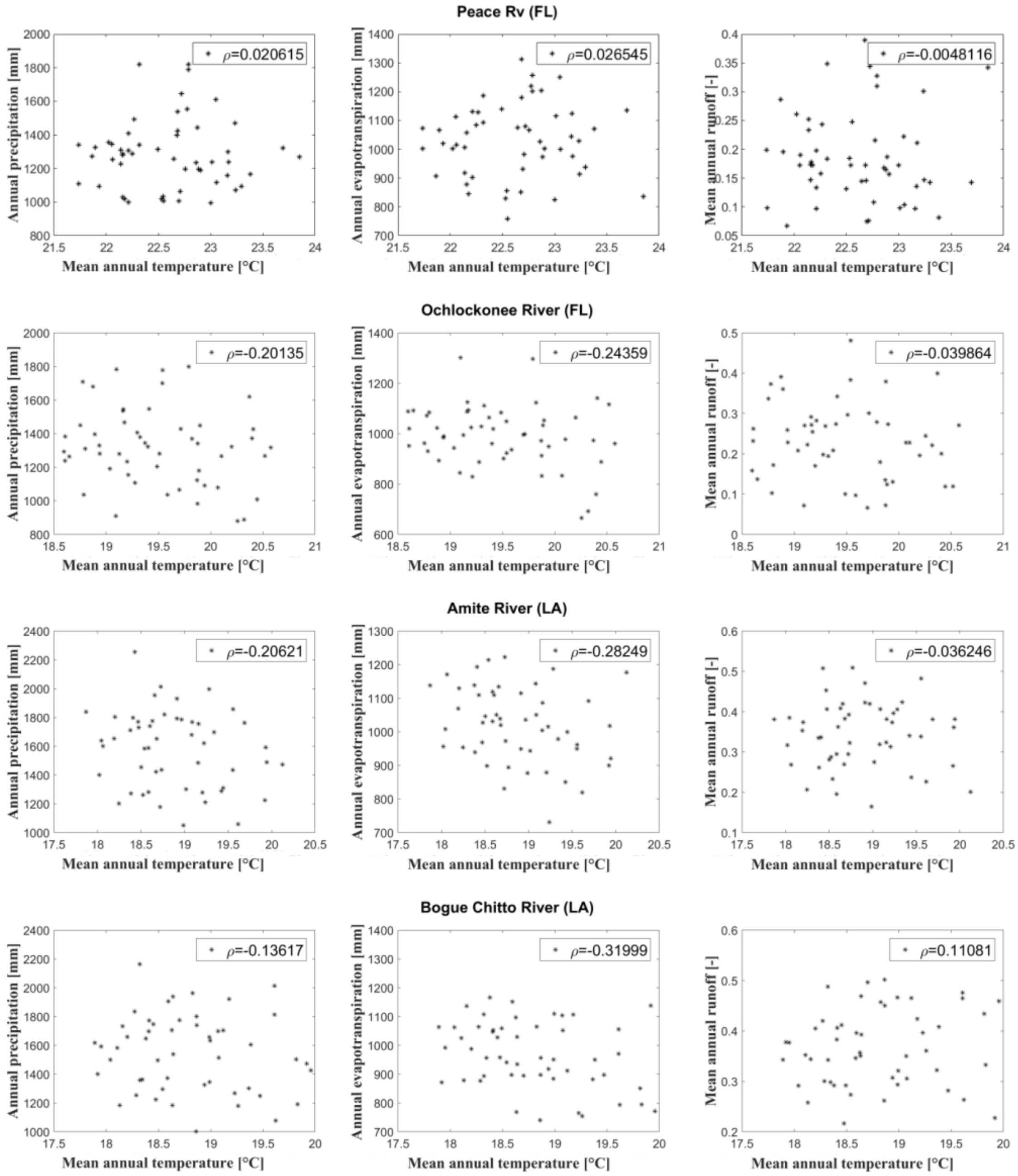


Figure 2. Annual precipitation against mean annual temperature (left panels), annual evapotranspiration against mean annual temperature (middle panels) and annual runoff coefficient against mean annual temperature (right panels) for four different basins: Peace River (FL), Ochlockonee River (FL), Amite River near Denham Springs (LA), Bogue Chitto River (LA). In each plot, the Pearson correlation coefficient ρ is reported in box.

2.~~4~~2 Preliminary analysis

The FDC can be interpreted as a distribution function of discharge over a given time period. To determine if samples are drawn from the same distribution, here the two-sample Kolomogorov-Smirnov test (KS; Massey, 1951) is carried out on each pair of samples. The KS statistic on two samples is a non-parametric test for the null hypothesis that the two independent samples are drawn from the same continuous distribution. The decision to reject the null hypothesis is based on comparing the p-value with the significance level set equal to 5%. Moreover, the test allows us to estimate the distance between couples of FDC:

$$D^{*} = \max_{x}(|F_1(x) - F_2(x)|), \tag{3}$$

where $F_1(x)$ is the proportion of x_1 values less than or equal to x and $F_2(x)$ is the proportion of x_2 values less than or equal to x. F_1 and F_2 are two FDCs. The KS statistic is applied on daily streamflow data sampled in several periods of record (e.g. 1 year, 10 years, 15 years). The long memory is relatively low, and we consider full years thus annual cycles do not have an influence on our results. The test is carried out both on pairs of samples gauged at the same location in two different years (or in two different decades) and on pairs sampled at two different sites. Since the streamflow data presents autocorrelation, the autocorrelation effects the KS test. Weiss (1978) proposed a methodology to account for modifying the KS test for autocorrelated data. Later, Xu (2014) suggested a method that can be applied to two samples test. The information contained in the data is (usually) less than an i.i.d. sample with the same size. In other words, the number of equivalent independent observations is fewer than the sample size. In the following, we explain how we ~~took into account~~accounted for the equivalent sample size. It is easier to implement and more importantly, it can be easily generalized to two samples test. We can assume that the autocorrelation effect attenuates after three days. For instance, let take as an example a 1~~-~~-year FDC. If the sample was three times smaller and for instance the length would equal 122 (i.e., 365 divided by 3), the null hypothesis would have been rejected anyway, leading to the same conclusion (i.e., the two samples cannot be regarded as the same distribution). This is due to the fact that, according to the two samples KS test, the length of the equivalent sample that could pass the test should be 22.

The application of the KS test to our samples is pivotal to the development of the methodology. Test results show that streamflow data gauged in different periods (e.g. years or decades) at a specific location do not have the same distribution. The consequence is that it is not possible to use the parameters and the distribution derived from a FDC built for a specific time window to build the FDC of another time window. The same results comparing streamflow data gauged in a specific year or decade at two different sites. Since the two data sets cannot be regarded as the same distribution, it is not possible to derive the FDC at one location using the parameters of ~~a~~the FDC sampled at another location. Therefore, it is necessary to develop a methodology that accounts for the weather, as it is the main driver of FDCs variability as shown in the following. Figure ~~4~~3 shows ~~how different can be~~the magnitude of the difference between FDCs built at the same location using streamflow data gauged during different time windows.

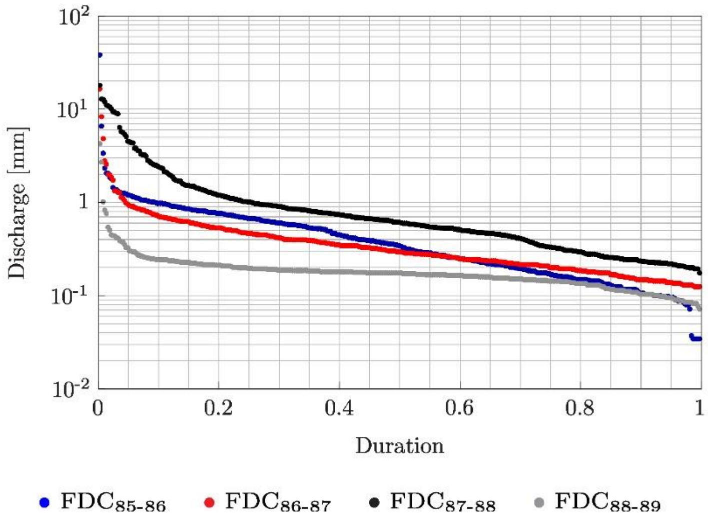


Figure ~~4~~3. FDCs built for Tangipahoa River (FL) for four different hydrological years. Every hydrological year starts in October and ends the following September.

3 Methodology

The aim of this paper is to find the distribution $Q_k(t)$ for a time period (T_1,T_2), that is a FDC. We assume that discharge is related to precipitation in the form:

$$Q_k(t) = h_k(P_k(t-\tau), \tau = 0, \dots, n, \dots, \beta_k), \tag{4}$$

where k ~~stands for the location~~is a generic site, h_k is the transformation, usually approximated by a hydrological model, P_k is the precipitation and β_k is the specific parameter of the hydrological model. The core of this work is to retrieve the discharge values without hydrological modelling as modelling ~~is~~ often ~~introducing~~introduces additional errors and ~~it~~ may be biased for long subperiods. Thus, the main idea is to get rid of a complicated non-linear ~~processes~~process and to find a filter which relates the distributions.

The main hypothesis underlying this work is that daily flow duration curves at a partially ungauged location can be found with knowledge of the precipitation record at a donor site. The most important descriptor of the weather characteristic is the rainfall~~,~~; however, we cannot use the distribution of P_k to assess the FDC directly as it will fail due to the lacking temporal structure and the many zeros. We can then use a transformation of P_k, the Antecedent Precipitation Index (API):

$$\mathrm{API}(t) = a_k(P_k(t-\tau), \tau = 0, \dots, n). \tag{5}$$

Both transformations reported in Eq. 4 and 5 can be regarded as filters acting on P_k. These filters do not necessarily produce highly correlated series, but may produce series with similar distributions.

The API is used to investigate precipitation data in a similar way to discharge data as it combines in a streamflow-like way the history of the precipitation. It represents the memory of a basin as it is related to the amount of water released by the soil to the river considering a given time window. Specifically, the API allows us to take into account the antecedent conditions, the duration of the rainfall events and gives an estimate of the portion of rainfall contributing to storm runoff (Linsley et al., 1949). It is a sequence of linear combination of rainfall events in the period preceding a specific storm (Kohler and Linsley, 1951). For a resolution of one day and a time window of 30 days, the API at the *i-th* day is given by:

$$\mathrm{API}_i = \sum_{j=0}^{29} \alpha^j P_{i-j}, \tag{6}$$

where α is a constant and ranges from 0 to 1 and P_i is the daily precipitation occurred at the *i-th* day~~. Since a day-to-day value of the API is required, there is a considerable advantage in assuming α decreasing with the time as shown by~~ (Kohler and Linsley, ~~(~~1951). When α tends to zero, API keeps tracks of the precipitation occurred in the few previous days and it represents the short memory of the basin. When α tends to 1, the API represents the long memory of the basin as it includes the effect of precipitation occurred many days before. To capture ~~this~~the latter behavior, in this study α is chosen equal to 0.85~~,~~. T~~t~~his is in agreement with a previous study by Sugimoto (2014) who investigated a case study area whereby a preliminary analysis was performed (i.e. Neckar ~~who investigated one of the twoa case study areasarea whereby a preliminary analysis was performed (i.e. Neckar catchment~~basin); nevertheless, this value was found to be suitable also for the US ~~catchments~~basins. Here the API is calculated from areal precipitation instead of point precipitation.

Formally the ~~The~~ basic hypotheses of this paper are:

- Flow duration curves are not invariant properties of basins but are the product of basin, weather and human interactions. In this investigation we do not consider the human interactions.
- Precipitation is the most important influencing factor on discharge.
- Basins delay the reaction on precipitation therefore the API is a better indicator for the influence of precipitation on discharge.
- We assume that discharge and API are changing ~~the~~ in a similar way for longer time periods.

Let $F_{A,Ti}(q)$ be the distribution of daily discharges at basin A and time period T_i (flow duration curve for the selected time period) and $G_{A,Ti}(a)$ be the distribution of daily API at basin A and time period T_i.

The transformation from T_i to T_j provides an estimated $F^*_{A,Tj}(q)$:

$$F^*_{A,T_j}(q) = G_{A,T_j}(G^{-1}_{A,T_i}(F_{A,T_i}(q))) \tag{7}$$

This is a quantile/quantile transformation.

$$F_{A,T_j}(q) = F_{A,T_j}(F^{-1}_{A,T_i}(F_{A,T_i}(q))) \tag{8}$$

The basic question can be written in the form of the following equation

$$G_{A,T_j}(G^{-1}_{A,T_i}(p)) \approx F_{A,T_j}(F^{-1}_{A,T_i}(p)) \tag{9}$$

That can be summarized in the following question: do the percentiles of the API change in the same way as those of the discharge? Note that if the relationship between API and discharge is a good one then the two sides are nearly equal. Even a weak relationship can do a good job if the errors are independent and the sign of the change is correct. ~~That can be summarized in the following question: do the percentiles of the API change in the same way as those of the discharge? Note that if the relationship between API and discharge is a good one to one relationship then the two sides are nearly equal. Even a weak relationship can do a good job if the errors are independent and the signal of the change is correct.~~

Figures 4 and 5 show the difference between the real change in percentiles and that obtained by using the API for different time periods according Eq.9. Note that the assumption that the FDC is time invariant would imply that the lines for the discharge are on the diagonal.

The correlation between API and discharge is around 0.6, but the transformations are quite similar and the API based transformation delivers good FDCs.

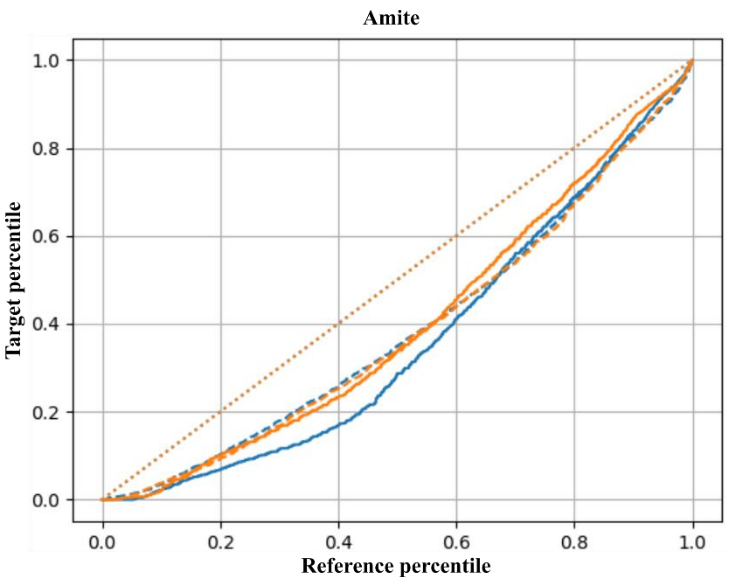


Figure 4. The transformation functions of the 1951-1960 FDC (solid) and API (dashed) to the target period 1971-1980 (blue) and 1981-1990 (orange) for Amite. For the sake of comparison, the diagonal is dotted in orange.

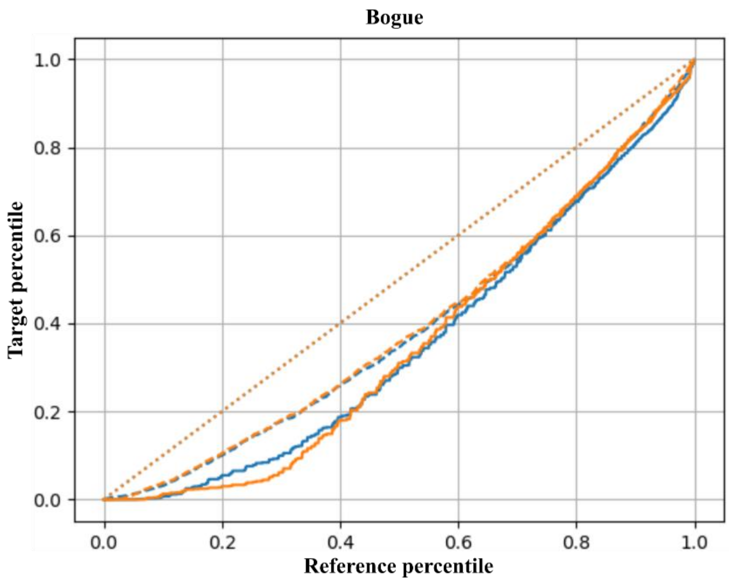


Figure 5. The transformation functions of the 1951-1960 FDC (solid) and API (dashed) to the target period 1971-1980 (blue) and 1981-1990 (orange) for Bogue. For the sake of comparison, the diagonal is dotted in orange.

If API is changing continuously in space then one can use the change of the FDC of a different location B for the estimation:

~~Is the API changing continuously in space? We assume that it is. If API is changing continuously in space then~~

$$F^{**}_{A,T_j}(q) = F_{B,T_j}(F^{-1}_{B,T_i}(F_{A,T_i}(q))) \qquad (10)$$

~~which means that FDC changes can be estimated from FDC changes of other basins (e.g., B in Eq.10).~~

In the following, the methodology is reported step by step~~ together with the underlying assumptions. Then~~, then, the performance criteria used to estimate the goodness of the methodology are presented.

3.~~2~~ 1 ~~How to determine the FDC at a partially gauged basin~~Procedure step-by-step

~~The assumptions underlying this work are the followings:~~

I. ~~The cumulative distributions of streamflow and the proxy correlate at a single site over the same period.~~
II. ~~The exceedance probability of the proxy on a specific day at the donor site is equivalent to the exceedance probability of streamflow on that same day at the target site.~~
III. ~~The cumulative distribution function of the proxy is identical across sites for both the index site and the target site in the same period.~~

~~Where the proxy variable is the variable used to retrieve the FDC at the target site. As the API was used as proxy for the U.S. case area, the first assumption is that the temporal sequence of API exceedance probabilities is highly correlated with the temporal sequence of streamflow exceedance probabilities at a single site over the same period, Table 3.~~

~~**Table 3.** Correlation between temporal sequence of API exceedance probabilities and the temporal sequence of streamflow exceedance probabilities estimated for different sites and different periods.~~

		~~**Correlation**~~	
~~**Period\Site**~~	~~Blanco~~	~~Tangipahoa~~	~~Choctawthachee~~
~~**1948-1968**~~	~~0.978~~	~~0.996~~	~~1~~
~~**1968-1988**~~	~~0.995~~	~~0.997~~	~~1~~
~~**1948-1963**~~	~~0.998~~	~~0.993~~	~~0.998~~
~~**1948-1958**~~	~~0.970~~	~~0.995~~	~~0.998~~

~~The second assumption verifies if, for the donor period, the temporal sequence of API exceedance probabilities at the donor site is highly correlated with the temporal sequence of streamflow exceedance probabilities at the target site. This assumption applies for all donor periods; Table 4 shows correlation values for some donor periods.~~

~~**Table 4.** Correlation between temporal sequence of API exceedance probabilities at the donor site (i.e., Blanco River, USA) and the temporal sequence of streamflow exceedance probabilities at three target sites for four different donor periods.~~

	~~**Correlation**~~			
~~**Sites\Donor period**~~	~~1948-1968~~	~~1968-1988~~	~~1948-1963~~	~~1948-1958~~
~~**Tangipahoa**~~	~~0.996~~	~~0.997~~	~~0.994~~	~~0.997~~
~~**Choctwhatchee**~~	~~1~~	~~1~~	~~0.999~~	~~0.999~~
~~**Bogue**~~	~~0.990~~	~~0.992~~	~~0.995~~	~~1~~

~~The third assumption is that cumulative distribution function of API is identical across sites for both the index site and the target site in the same period. This assumption was verified performing a Kolmogorov-Smirnov test on API at different sites for different periods. For instance, the Weibull distribution is accepted for the API at Tangipahoa, Choctawhatchee and Bogue (USA) for the periods shown in Table 5. However, the distribution parameters may differ from site to site and from time period to time period. The correlation between temporal sequence of API exceedance probabilities at the donor site and at each target site is found to be high, thus the assumption was further verified, Table 5.~~

~~**Table 5.** Correlation between temporal sequence of API exceedance probabilities at the donor site (Blanco, USA) and three other sites is reported for four different periods.~~

	~~**Correlation**~~			
~~**Sites\Donor period**~~	~~1948-1968~~	~~1968-1988~~	~~1948-1963~~	~~1948-1958~~
~~**Tangipahoa**~~	~~0.98~~	~~0.99~~	~~0.97~~	~~0.99~~
~~**Choctwhatchee**~~	~~0.98~~	~~0.98~~	~~0.98~~	~~0.98~~
~~**Bogue**~~	~~0.99~~	~~0.99~~	~~0.98~~	~~0.99~~

~~In the following, the procedure is explained step by step using the API as a proxy. However, similarly, it is possible to use the streamflow values recorded at the donor site.~~

Let consider two ~~catchments~~basins, A and B. We want to determine the Flow Duration Curve at ~~catchment~~basin B from data available at A. Therefore, A is the donor ~~catchment~~basin, while B is the target ~~catchment~~basin. Let suppose that in a given number of years, discharge is available at both sites A and B, named donor years, while for another number of years, i.e. the target years, data are available for A only.

1. *Donor years selection*. Select a number of years for which precipitation and discharge values are available at daily resolution for ~~catchment~~basin A and B, respectively. These will be named donor years (e.g. with duration of 1 year, 10, 15, 20 years).
2. *Generation of empirical distribution of API values.* Empirical distributions of API values are calculated for site A for donor and target years: sort API values and assign to each sorted value the corresponding rank and estimate the corresponding frequency of exceedance using the Weibull plotting position.
3. *Generation of empirical distribution of streamflow values.* Empirical distributions of streamflow values are calculated for site B for donor years only.
4. *Data transfer from donor site.*

i. Select the *i-th* frequency p_i, with i=1,…,N_t where N_t is the length of the target sample, and the corresponding API value recorded at the donor site during the target years, Figure ~~5a~~6a.

ii. Search for this API value among those recorded at the donor site during the donor years and estimate the corresponding frequency, Figure ~~5b~~6b.

iii. This frequency is then used to retrieve the corresponding streamflow value recorded at site B during the donor years, Figure ~~5c~~6c.

iv. This streamflow value is the missing value at site B corresponding to the *i-th* frequency p_i, Figure ~~5d~~6d.

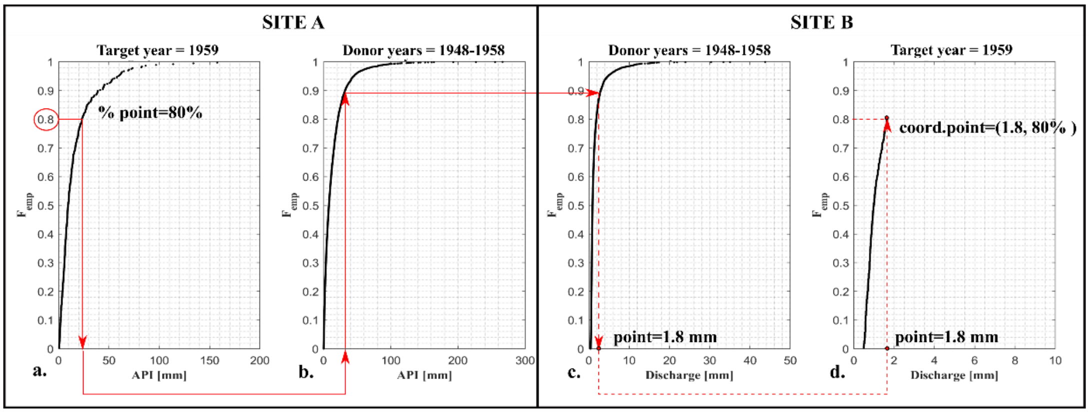


Figure ~~5~~6. Illustration of FDC generation using the interpolation with the API of the donor site as a proxy.

Steps from 1 to 4 are repeated for every frequency and then for different target periods and ~~for different~~ target ~~catchments~~basins. The FDC is expressed in millimeter, thus the area of the ~~catchment~~basin is not an issue using data of another ~~catchment~~basin.

An example of the procedure is reported step by step in Appendix A.

3.~~3~~ 2 Performance criteria

To determine the performance of the procedure proposed in this paper, different criteria are selected: the Nash-Sutcliff efficiency index (NSE; Nash and Sutcliff, 1970), the BIAS and the Mean Absolute Error (MAE).

The Nash-Sutcliffe efficiency between the interpolated and the observed flow value is the most widespread performance criterion:

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{N}\left(Q_{obs,i} - Q_{intrpl,i}\right)^2}{\sum_{i=1}^{N}\left(Q_{obs,i} - \bar{Q}\right)^2}, \qquad (7)$$

where Q_{obs} is the observed discharge value at the target ~~catchment~~basin during the target period; $\bar{Q}$ is the mean value of the observed discharge during the target period in the target ~~catchment~~basin; Q_{intrpl} is the interpolated discharge value. ~~The NSE is evaluated here for a specific set of percentiles, thus, *N* is the number of discharge values related to a specific percentile. The **X** percentile is defined as the set containing all "**X**,…" numbers where the dots stand for the decimal points. For instance, the~~

~~1.09%, 1.36%, 1.63%, 1.91% belong with the 1rst percentileIn binning by percentiles, all percentages were rounded down to the nearest whole number.~~

The BIAS represents the mean difference between observed and interpolated values~~:~~ (Castellarin et al., 2001; Ridolfi et al., 2016):

$$\text{BIAS} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{Q_{intrpl,i}-Q_{obs,i}}{Q_{obs,i}}\right). \tag{8}$$

~~If~~This metric comprises the ~~BIAS equals zero there~~mean of the error made relative to the observed record. It is a ~~perfect fit~~signed and unbounded metric. It indicates as a ratio the level of overall agreement between the observed and interpolated values. ~~If the BIAS is negative, observed values are underestimated, while if the BIAS is positive, they are overestimated.~~

The mean absolute error is defined as:

$$\text{MAE} = \frac{\sum_{i=1}^{N}\left|Q_{obs,i}-Q_{intrpl,i}\right|}{N}. \tag{9}$$

Discharge values are in mm and so is the MAE ~~is~~. It measures the overall agreement between observed and interpolated values. It is a non-negative metric without upper or lower bounds. A perfect model would result in a MAE equals to zero. This estimation metric does not provide any information about under- or over-estimation, but it determines all deviations from the observed values regardless of the sign. All metrics are evaluated here for a specific set of percentiles, thus, *N* is the number of discharge values related to a specific percentile. In binning by percentiles, all percentages were rounded down to the nearest whole number.

4. Results

The procedure explained above was tested on several target ~~catchments~~basins varying both donor and target periods. ~~For the U.S. catchments, using a donor period of 20 years, we considered 10 years and 1 year as target periods. For donor periods equal to 15 and 10 years, we considered as target periods 15 and 10 years, respectively.~~

Results show a good agreement between observed and interpolated FDCs. For instance, the FDCs interpolated using 20 and 10 years as donor and target periods, respectively, have a good performance, as shown for Tangipahoa and Bogue ~~catchments~~basins, Figure ~~6~~7. The method performance is higher for intermediate durations, while it can be lower for the low flows, e.g. as at Bogue for target years 1988-1998 (Figure ~~6~~7 lower panels) and for the high flows. The good performance of the approach is also noticeable when the target period is 15 years, Figure ~~7~~8. On each panel, the two-sample Kolmogorov-Smirnov test distance between observed and interpolated ~~value~~values, D*, is reported. D* is characterized by small values showing a good performance of the method. Since usually the FDC of a donor site is used to retrieve the FDC of a target site for the same period, the FDC of the donor ~~catchment~~basin recorded during the target period ~~was~~is also plotted. It is noteworthy to observe that the difference between these two FDCs can be substantial. This implies that the FDCs can be substantially different at different sites in the same period of time~~, in turn it entails that the FDC of a donor site cannot be transferred to another site~~.

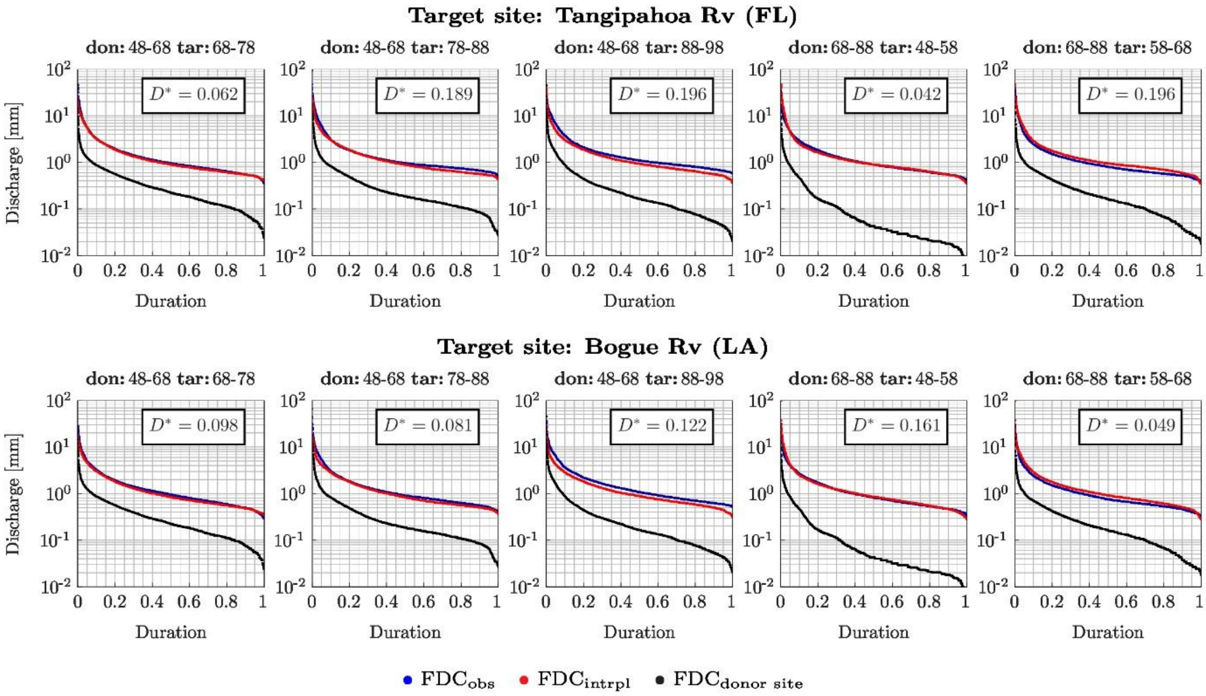


Figure ~~6~~7. Interpolated FDC at Tangipahoa River (FL) and Bogue River (LA), upper and lower panels, respectively. The donor ~~catchment~~basin is Blanco River (TX). The donor years are a 20 years time window from October 1948 to September 1968 and from October 1968 to September 1988. Target years are the decades shown above each panel. Blue ~~dots~~ and red dots are the observed and interpolated FDC at the target ~~catchment~~basin during the target period, respectively; the black dots are the observed FDC at the donor ~~catchment~~basin during the target period. In each box the KS distance between observed and interpolated values, D*, is reported. The p-value of D* is always around zero but for Tangipahoa target years 1948-1958.

Interpolated and observed FDCs almost perfectly match when obtained using long donor and target periods, Figures ~~6~~7 and ~~7~~8. On the other hand, when the target period is short, the performance decreases as also shown by the KS distance, D*, reported on each single panel of Figure ~~8~~9 where the target period equals one year. As a matter of fact, the donor period being constant, the KS distance is much higher when the target period is 1 year (Figure ~~8).~~9) and the p-value of D* is always zero but for hydrologic years 1972-1973 and 1976-1977. Nevertheless, the interpolated and observed FDCs have a high agreement in shape, as for instance at Tangipahoa River for all but one (i.e., 1969-1970) target years. In these cases, the difference between the two curves could be due to the different temperature values characterizing the donor and the target basins. This ~~effects~~affects the evapotranspiration in the two basins and therefore, the streamflow values.

Results suggest that the API gives effectively a good estimation of the memory of the basin and can be used to represent the precipitation similarly to the discharge.

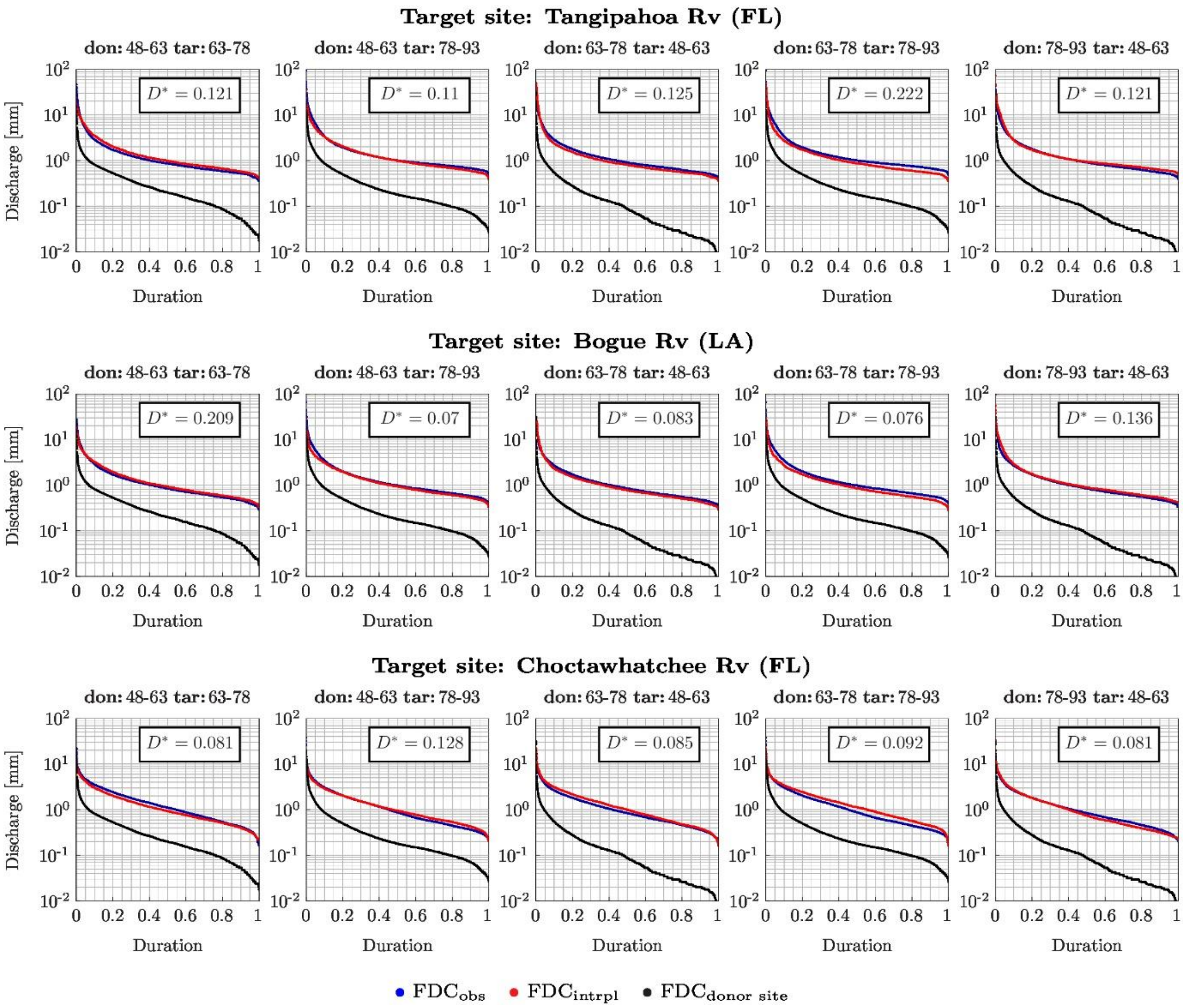


Figure ~~7~~8. Interpolated FDC at Tangipahoa River (FL), Bogue River (LA) and Choctawhatchee River (FL), upper, middle and lower panels, respectively. The donor ~~catchment~~basin is Blanco River (TX). The donor and target years are periods of 15 years. The blue and red dots are observed and interpolated FDC, respectively, at the target ~~catchment~~basin during the~~,~~ target period; ~~the dots are FDC at the target catchment and~~ the black dots are the observed FDC at the donor ~~catchment~~basin during the target period. In each box the KS distance, D*, is reported. The p-value of D* is always around zero.

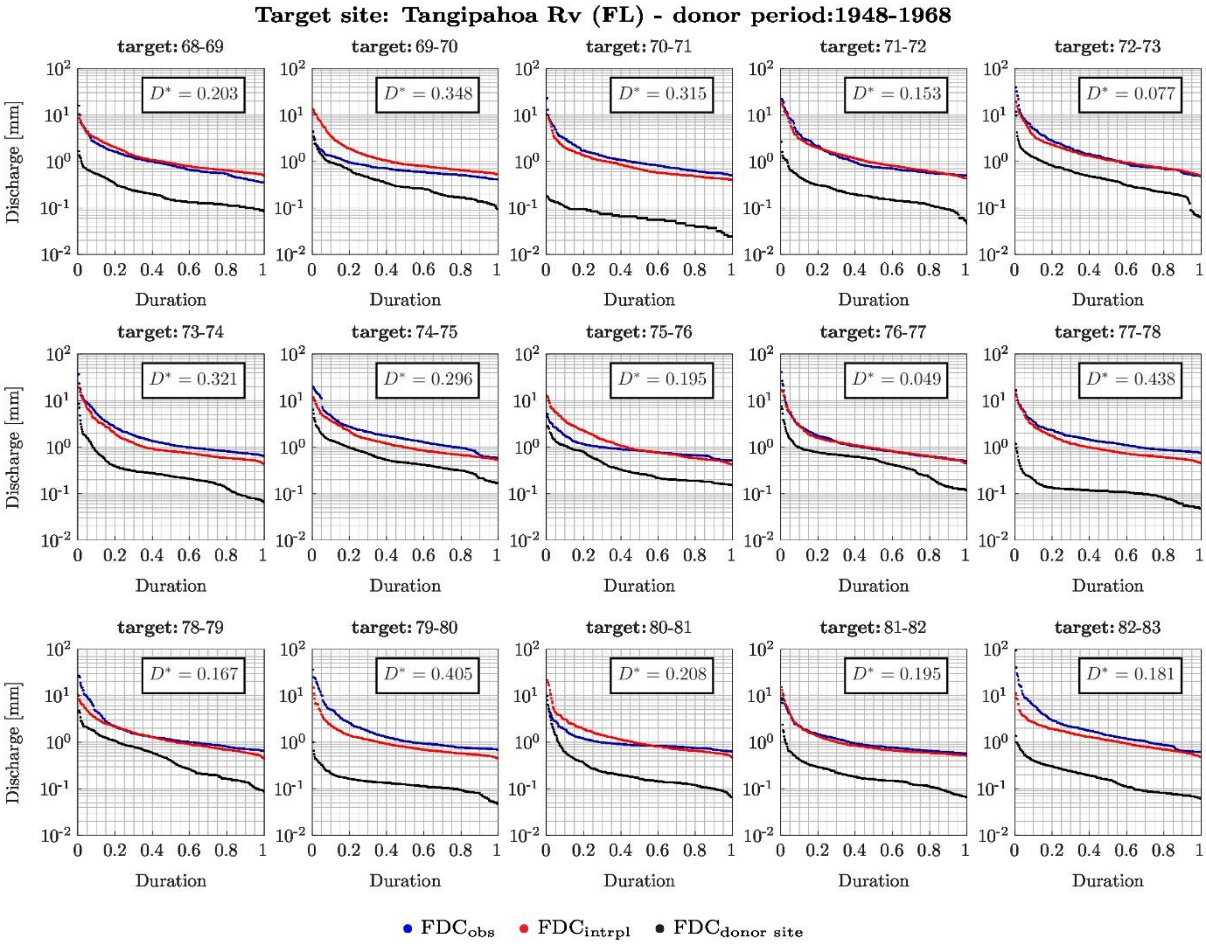


Figure ~~8~~9. Interpolated FDC at Tangipahoa River (FL). The donor ~~catchment~~basin is Blanco River (TX). The donor years are a 20 years time window from October 1948 to September 1968. Target years are each hydrological year from October 1968 to September 1983. The blue and red dots are observed and interpolated FDC, respectively, at the target ~~catchment~~basin during the~~,~~ target period; ~~the dots are FDC at the target catchment and~~ the black dots are the observed FDC at the donor ~~catchment~~basin during the target period. In each box the KS distance, D*, is reported. The p-value of D* is always around zero but for hydrologic years 1972-1973 and 1976-1977. ~~.~~

To estimate the goodness of the methodology, the NSE, BIAS and MAE are evaluated for the 1st, 3rd, 5th, 10th, 20th, 30th, 50th, 75th, 90th and 99th percentiles.

~~For U.S. catchments, w~~When a decade is used as both target and donor period, the performance measures show a good agreement between observed and interpolated values, Figures ~~9~~10. The NSE index shows accurate estimation, i.e. it is characterized by values close to 1, especially of intermediate percentiles. The BIAS provides information regarding the ~~over- and underestimation of the~~overall agreement between interpolated and observed values. Its magnitude is likely higher for high flows, while it attenuates for intermediate percentiles. ~~Also the~~The MAE as well shows a low performance for high streamflow values. This is due to the fact that the procedure is more able to reproduce the average streamflow values than extreme events such as high and low flows. However, low flows are more likely well estimated rather than high flows.

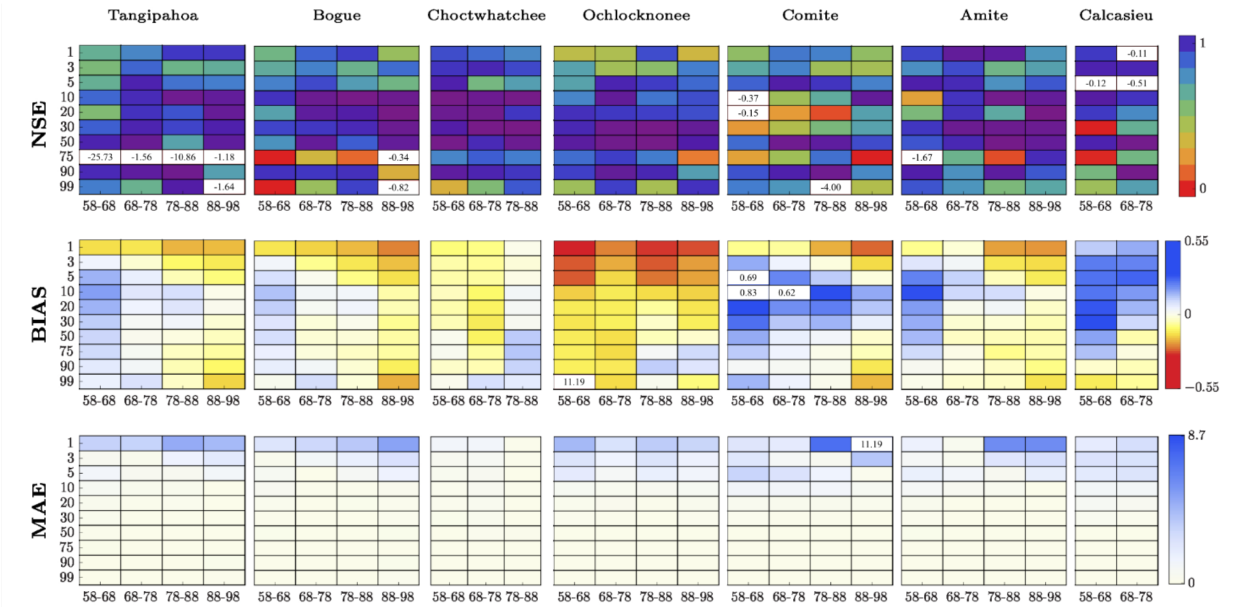


Figure ~~9~~10. Performance measures NSE, BIAS and MAE evaluated for specific percentiles (on the y-axis) and for specific target decades on the x-axis. The donor decade is 1948-1958, the donor ~~catchment~~basin is Blanco (TX). Each target ~~catchment~~basin is indicated ~~above~~ in the corresponding box. Negative values of the NSE as well as outliers of BIAS and MAE are reported ~~on~~ in the corresponding box.

When both target and donor periods equal 15 years, the agreement between interpolated and observed flow values is high, Figure ~~10~~11. The NSE shows values of efficiency around 1, thus there is a good match between interpolated and observed values, even though there are few exceptions. The errors are very low in value, as shown by the MAE, which also reveals a poor performance for high flows, while the performance improves for intermediate and low flows. The high flows are more likely ~~overestimated~~estimated with a higher error, ~~while~~ than intermediate and low flows ~~are more likely underestimated~~ as also shown by the BIAS.

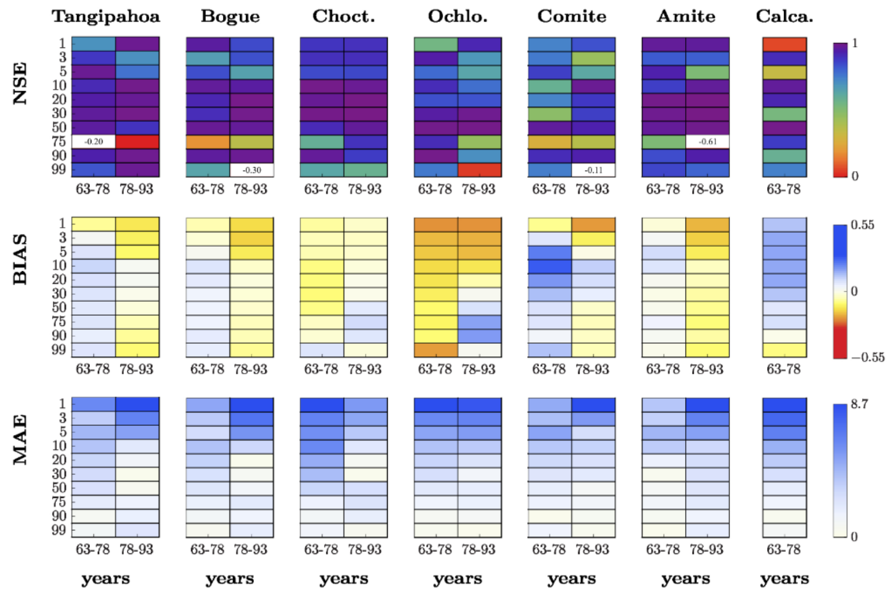
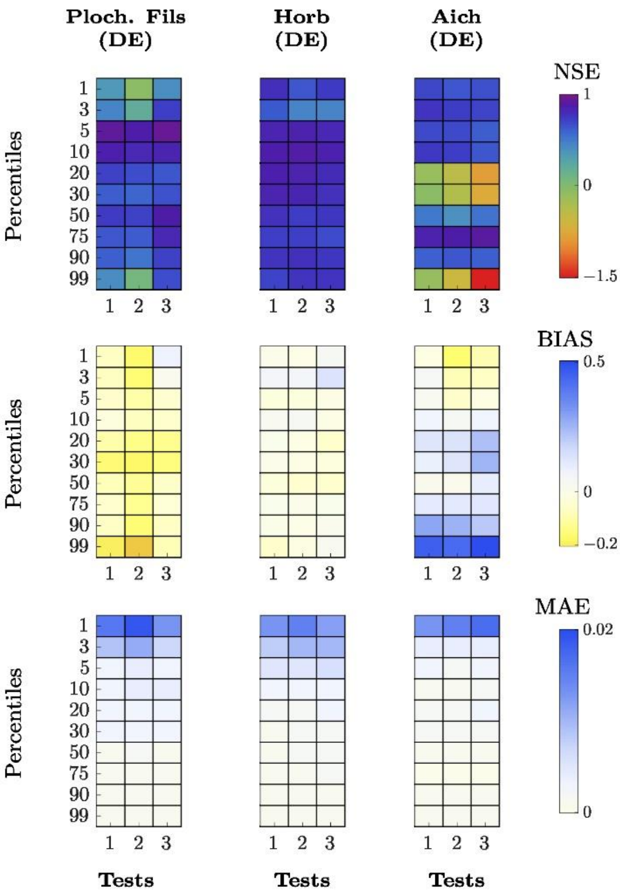


Figure ~~10~~11. Performance measures NSE, BIAS and MAE evaluated for specific percentiles (on the y-axis) and for~~ a~~ specific 15 target years (i.e., 1963-1978 and 1978-1993 on the x-axis). The donor decade is 1948-1963, the donor ~~catchment~~basin is Blanco (TX). Each target ~~catchment~~basin is indicated ~~above~~ in the corresponding box. Negative values of the NSE are reported in the corresponding box.

~~For the German catchments, three different pairs of donor and target years are considered. Test 1 is built using a target time periods of 20 years and a donor period of 10 years. Test 2 is built considering as both donor and target time periods 15 years, while for Test 3, the target time period equals 10 years and the donor one equals 20 years.~~

~~We recall that for the German case study the proxy variable is not the API but rather the discharge recorded during the donor period at the donor site. For the German case study, errors are reported for Plochingen Fils, Horb and Oberensingen Aich (henceforth named Aich) as target catchments, using as donor catchment Plochingen Neckar.~~

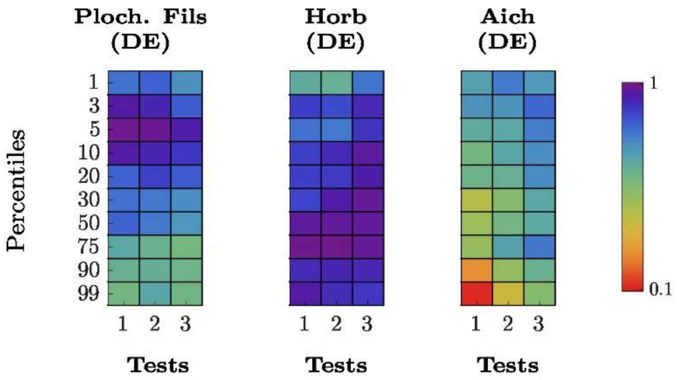
~~As for the U.S. catchments, the estimation metrics show a lower performance for extreme flows. For intermediate percentiles, the NSE shows values closer to 1 and the BIAS is generally close to zero. However, it is worth noticing that the overall agreement between observed and interpolated values is high as demonstrated by a low value of the MAE, Figure 11.~~



~~**Figure 11.** Performance measures NSE, BIAS and Mean Absolute Error (MAE) evaluated for specific percentiles (on the y-axis) and for three specific set of target and donor years (i.e., Test 1, 2 and 3). For Test 1 the target period is from 1961 to 1980 and the donor period is from 1981 to 1990. For Test 2 the target period is from 1961 to 1975 and the donor period is from 1976 to 1990. For Test~~

~~3 the target period is from 1961 to 1970 and the donor period is from 1971 to 1990. The donor catchment is Plochingen Neckar, while the target catchments are indicated on each corresponding box.~~

~~To better understand the relationship between a target and a donor catchment, the coefficient of correlation has been computed. Coefficient values are reported for all Test cases. Values are estimated between the donor catchment Plochingen Neckar and the three target catchments, Figure 12.~~



~~**Figure 12.** Correlation coefficient evaluated for each Test case at each percentile between Plochingen Neckar and each other catchment indicated on the boxes.~~

~~It is interesting to observe that the correlation coefficient shows the same trend of the NSE, as it shows a higher correlation where the NSE is closer to one, while generally they both decrease in correspondence of the same percentiles. The correlation coefficient shows how the proxy variable, in this case the discharge gauged at the donor site, co-moves with the target variable. As expected, where the correlation is high, there is a better estimation of the flow values. Therefore, this means that it is possible to know *a priori* whether a site is more suitable to be a donor site or not. If the correlation is low, also the performance of the method is expected to be low.~~

5. Discussion

As resulted from the KS test applied to pairs of FDCs obtained from recorded data at the same site in different periods, FDCs cannot be considered an invariant characteristic of a basin. The fact that FDCs are not invariant suggests that the weather is a driver of annual runoff variability. Indeed, the reason should be found in the weather conditions as others (e.g. the ~~catchment~~basin area, the land use) did not change. To better investigate these findings, we performed the KS test on pairs of observed and interpolated FDCs for two purposes. The first is to know if pairs of interpolated and observed FDCs at the same site have the same continuous distribution, the second is to know which is the distance between these pairs. The test performed on pairs of interpolated and observed FDCs revealed that the null hypothesis could not be rejected for nearly the half of the cases. For instance, for Tangipahoa River the test was not rejected in 48% of the cases, Figure ~~13~~12 right panel. On the contrary, the test rejected the null hypothesis that FDCs built at the same location in different periods had the same distribution. In the 73% of the cases, the distance between pairs of interpolated and observed FDCs of the same period is smaller than the distance between FDCs built at the same site from data recorded during different periods, Figure ~~13~~12 right and left panel, respectively. These results suggest that the methodology proposed here has a good performance and it is actually an interesting alternative to other methodologies, which assume that FDC of different ~~periods of time~~periods have the same distribution.

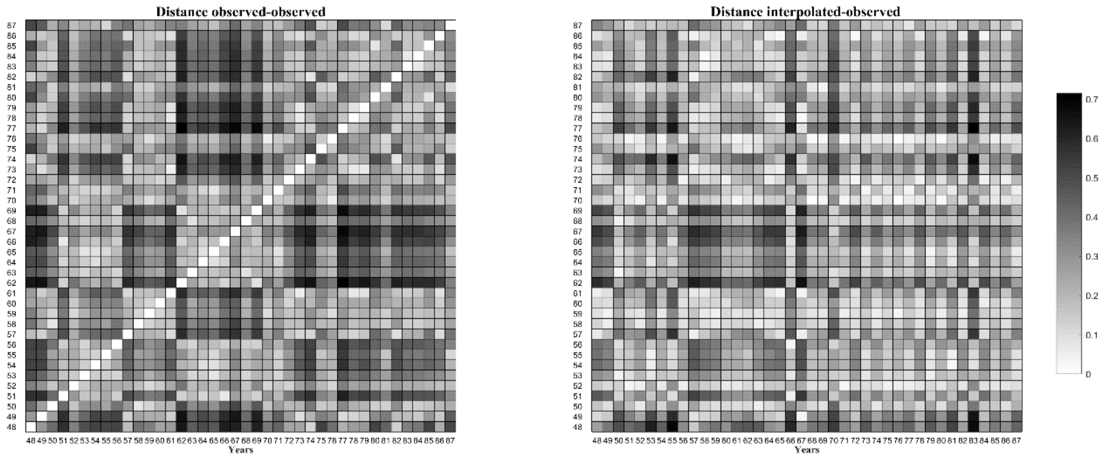


Figure ~~13~~12. Kolmogorov-Smirnov distance between couples of streamflow values observed (left panel) and between couples of streamflow values observed and interpolated (right panel) at Tangipahoa River (FL) from October 1948 to September 1987.

As the weather conditions strongly influence the FDCs estimation, we analyzed the streamflow percentiles to assess the between-year variability. To this end, the moving average (MA) of 30th, 70th, 90th and 95th percentiles of streamflow is estimated. The MA values are estimated using three different fixed time windows (i.e., 10, 15 and 20 years), Figure ~~14~~13.

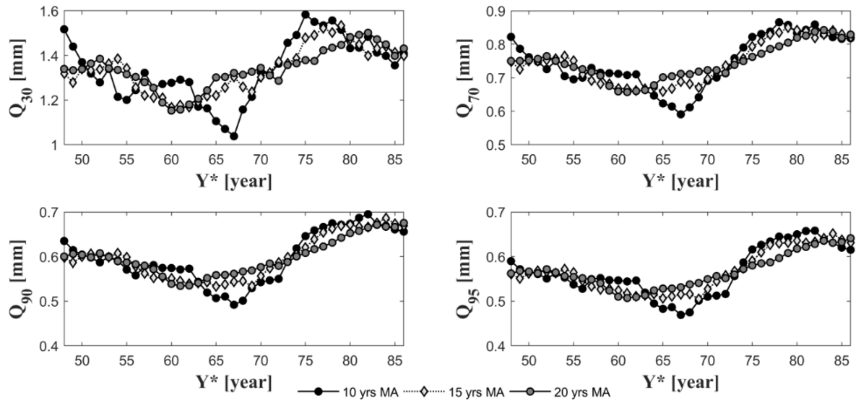


Figure ~~14~~13. Moving average (MA) of the 30th, 70th, 90th and 95th percentiles of daily streamflow values gauged at Tangipahoa. Three different fixed time windows are used to estimate the MA: 10, 15 and 20 years. On the x-axis the first year of each interval is plotted (Y*).

It is interesting to observe that the MA values are characterized by a strong variability throughout the time. The fluctuation of the flow percentiles suggests that the percentiles cannot be considered an invariant characteristic of the basin. Therefore, it is not possible to estimate the flow quantiles using regression methods that do not ~~take into account~~consider the weather characteristics. These methods, first, regionalize empirical runoff percentiles using multiple regression models. Then, regional evaluation of flow percentiles are interpolated across the percentiles (e.g., Franchini and Suppo, 1996; Smakhtin, 2001). If flow percentiles are estimated separately from weather characteristics, it may ~~results~~result in a misrepresentation of the percentiles themselves. Therefore, we suggest to add a weather factor to ~~take into~~ account for the influence of the weather in the percentiles estimates.

6 Conclusions

The paper presents a new, simple and model free methodology to estimate the streamflow behavior at partially gauged ~~catchments~~basins, given ~~the discharge and~~ the precipitation gauged at another ~~catchment~~basin. We show that two FDCs built for the same ~~catchment~~basin with data corresponding to two different time windows, cannot be regarded as the same continuous distribution. This means that the FDCs cannot be considered an invariant characteristic of a basin. As other conditions did not substantially change across time, such as the land use, the reason should be the weather. The influence of the weather is evident analyzing the between-year variability of flow percentiles. Indeed, the moving average of the 30th, 70th, 90th and 95th flow percentiles shows a strong variability throughout the time. This behavior has a strong consequence as it means that it is not possible to retrieve the streamflow percentiles without ~~taking into account~~considering the weather. Indeed, there exists several methodologies (i.e., regression models) that estimate flow quantiles separately from weather characteristics. FDCs and their selected properties cannot be considered as ~~catchment~~basin characteristics and should be used with caution for regionalization purposes. The FDC at a specific site is not a property of the corresponding basin, but ~~the FDC is a property~~rather of both the basin and the weather. Therefore, it is not possible to infer ~~a~~an FDC using parameters retrieved from the distribution of another FDC without considering the weather. The weather is indeed one of the main drivers of annual variability. The annual runoff variability depends on the different availability of energy and water in the basin. If more water than energy is available, the relationship between runoff and precipitation is almost linear, while if more energy is available, than the evaporation makes this relationship non-linear. Therefore, the runoff may vary largely depending on which element is prevalent. For this issue, we applied the methodology on basins with the same characteristics, i.e. energy limited ones.

Because of the dependence on the climate, discharge data are here retrieved using the precipitation data series. Since precipitation data series are characterized by a high number of zeros, here we used the Antecedent Precipitation Index (API)~~. The API is used~~ as it represents in a streamflow-like way the precipitation of the basin. It represents the memory of a basin providing the amount of precipitation released by the soil throughout the time.

The FDC at a target site is determined for a specific time window (i.e., target period) using API ~~and discharge~~ available for a so-called donor period at another ~~catchment~~basin (i.e., donor site).

~~To test the methodology, several donor and target periods are analyzed, such as 1 year, 10, 15 and 20 years and two case study areas are investigated, one located in USA and the other one in Germany.~~ Interpolated FDCs are compared with FDCs that were actually observed. ~~From the~~their ~~comparison of observed and interpolated FDCs, it r~~Results show that the methodology is able to correctly determine the missing streamflow data. The discharge values of the intermediate percentiles are better described than those of the extremes. Nevertheless, the error values between observed and interpolated FDCs are small. The difference between the interpolated and observed FDCs can be due to the different temperature values characterizing the donor and target ~~catchments~~basins. Indeed, a high difference in temperature can cause a different evapotranspiration, which in turn can influence the discharge. ~~To better analyze the relationship between donor and target catchments, the coefficient of correlation is computed between discharge data gauged at the two sites of interest during the donor period. As the performance criteria highlighted, the data series are more related at the intermediate~~

~~percentiles and less at the extremes. The correlation coefficient estimated for the donor period can help to determine in advance whether the discharge data of a donor and a target catchments are strongly correlated during that period of time. The FDCs interpolated at the target site will be more accurate if the correlation coefficient shows a strong correlation.~~

To test the methodology and to assess its performance depending on the extension of the period with missing data, several target periods are analyzed, such as 1 year, 10 and 15 years. ~~Results show that t~~The method performs better when the target period is longer, thus the lowest and the best performance correspond to target periods of one year and 15 years, respectively.

The method is tested on basins with a mild climate, however it ~~is possible t~~can be ~~o use~~applied ~~it~~ also ~~in~~to basins characterized by the presence of snow, converting the snow into the corresponding liquid amount.

Appendix A

In this Appendix we want to provide an easy example to better understand the method that we applied ~~for~~ to U.S. ~~catchments~~basins. This method is based on the use of the API of a donor site to retrieve the FDC at a poorly gauged site. We recall that a "donor period" is a period of time for which streamflow values are available at ~~both donor and~~ target ~~catchments~~basin, while a "target period" is a period of time during which streamflow values are not available at the target ~~catchment~~basin. ~~As t~~The rainfall is available at ~~both~~ the donor site~~s~~ for both periods~~, also API values are~~.

Let suppose that we want to know the discharge value at ~~catchment~~basin B (i.e., Bogue Rv, LA) corresponding to the 10.11th percentile (i.e., 10.11%) for the year ranging from October 1968 to September 1969. Let suppose that the donor period has a length of 15 years. Every hydrological year ranges from October to September of the following year. We present the method step by step in the following.

1. Select the mean daily precipitation occurred at the donor ~~catchment~~basin (i.e., Blanco Rv) during the target period and estimate the API as in Eq.6 assuming α equal to 0.85;

2. sort in descending order the API values evaluated for the target period at the donor ~~catchment~~basin (i.e., Blanco Rv, TX);

3. assign to each sorted value the corresponding rank *i*, with i = 1,..., N_t where N_t is the length of the target API series and thus equals 365, and then estimate the exceedance probability P(~~$Q < q_i$~~$API < API_i$) of each value using a Weibull plotting position $i/(N_t + 1)$, Table A1;

4. in the sorted API series, identify the value with frequency equal to 10.11%. This value equals 37.72 mm (bold line in Table A1);

5. estimate the API from the mean daily precipitation occurred during the donor period at the donor ~~catchment~~basin (i.e., Blanco Rv, TX) and sort in descending order the API values, estimate the rank and the associated exceedance probability P(~~Q~~ API < ~~q~~API_j) of each value as j/~~=~~$(N_r + 1)$ where N_r equals 5475;

6. find the exceedance probability $P(API < API_j)$~~$P(Q < q_j)$~~ associated to the value 37.72 mm in the sorted API sample. From Table A2 it is possible to observe that there is not such an API value. Therefore, look for the two most similar values: one should be bigger and the other smaller than the searched value. Then, take their empirical frequency values (i.e., 7.52 % and 7.54%; in bold, Table A2);

7. sort in descending order the streamflow values gauged during the donor period at the target ~~catchment~~basin (i.e., Bogue Rv, LA), estimate the rank and the associated exceedance probability $P(Q < q_j)$ of each value as j/~~=~~$(N_r + 1)$;

8. find the two streamflow values which have an empirical frequency equal to 7.52 % and 7.54%. These values are in bold, Table A3;

9. estimate the mean value of these two streamflow values. The resulting value is the streamflow value with empirical frequency equal to 10.11% evaluated for the target ~~catchment~~basin and the target period that we were looking for, Table A4.

Table A1. API values sorted in descending order and the corresponding percentiles estimated for the target year (i.e., 1968-1969) at the donor ~~catchment~~basin (i.e., Blanco RV, TX).

Rank	$P(API<API_i)$ %	$API_{Blanco, tar}$ mm
1	0.27	76.78
2	0.55	73.39
…	…	…
30	8.20	39.65
31	8.47	39.35
32	8.74	38.71
33	9.02	38.31
34	9.29	38.18
35	9.56	38.10
36	9.84	37.97
37	**10.11**	**37.72**
38	10.38	36.99
…	…	…
365	99.73	0.61

Table A2. API values corresponding to specific percentiles estimated for the donor years (i.e., 1948-1963) at the donor ~~catchment~~basin (i.e., Blanco RV, TX).

Rank	$P(API<API_j)$ %	$API_{Blanco,\ ref}$ mm
1	0.02	266.17
…	…	…
410	7.49	37.81
411	7.51	37.78
412	**7.52**	**37.74**
413	**7.54**	**37.61**
414	7.56	37.61
415	7.58	37.55
…	…	…
5475	99.98	0.01

Table A3. Streamflow values corresponding to specific percentiles gauged during the donor years (i.e., 1948-1963) at the target ~~catchment~~basin (i.e., Bogue RV, LA).

Rank	$P(Q<q_j)$ %	$q_{Bogue,ref}$ mm
1	0.02	38.81
…	…	…
410	7.49	3.28
411	7.51	3.28
412	**7.52**	**3.21**
413	**7.54**	**3.21**
414	7.56	3.20
415	7.58	3.19
…	…	…
5475	99.98	0.31

Table A4. Streamflow value corresponding to the 10.11th percentile estimated for the target year (i.e., 1968-1969) at the target ~~catchment~~basin (i.e., Bogue RV, LA).

$P(Q<q_i)$ %	$q_{Bogue,tar}$ mm
10.11	3.21

Competing interests. No competing interests are present.

Acknowledgements: The Authors would like to express their gratitude to William Farmer, Thomas Over and Francesco Serinaldi for the fruitful discussion that consistently improved the paper.

References

Alaouze, C.: Transferable water entitlements which satisfy heterogeneous risk preferences, Australian Journal of Agricultural Economics, 35, 197–208, 1991.

Archfield, S. A., & Vogel, R. M.: Map correlation method: Selection of a reference streamgage to estimate daily streamflow at ungaged catchments. Water Resources Research, 46, W10513, https://doi.org/10.1029/2009WR008481, 2010.

Bárdossy, A., Pegram, G. G. S., and Samaniego, L.: Modeling data relationships with a local variance reducing technique: Applications in hydrology, Water Resources Research, 41, https://doi.org/10.1029/2004WR003851, 2005.

Bárdossy, A., Huang, Y., and Wagener, T.: Simultaneous calibration of hydrological models in geographical space, Hydrology and Earth System Sciences, 20, 2913–2928, https://doi.org/10.5194/hess-20-2913-2016, 2016.

Blöschl, G., Sivapalan, M., Thorsten, W., Viglione, A., and Savenije, H.: Runoff prediction in ungauged basins: Synthesis across Processes, Places and Scales, Cambridge University Press, 2013.

Bonta, J. V. and Cleland, B.: Incorporating natural variability, uncertainty, and risk into water quality evaluations using duration curves, Journal of the American Water Resources Association, 39, 1481–1496, https://doi.org/10.1111/j.1752-1688.2003.tb04433.x, 2003.

Brown, A. E., Zhang, L., McMahon, T. A., Western, A. W., and Vertessy, R. A.: A review of paired catchment studies for determining changes in water yield resulting from alterations in vegetation, Journal of Hydrology, 310, 28–61, https://doi.org/10.1016/j.jhydrol.2004.12.010, 2005.

Budyko, M. I.: Climate and Life, Academic Press, San Diego, California, USA, 1974.

Carrillo, G., Troch, P. A., Sivapalan, M., Wagener, T., Harman, C., and Sawicz, K.: Catchment classification: Hydrological analysis of catchment behavior through process-based modeling along a climate gradient, Hydrology and Earth System Sciences, 15, 3411–3430, https://doi.org/10.5194/hess-15-3411-2011, 2011.

Castellarin, A., Camorani, G., and Brath, A.: A stochastic model of flow duration curves, Advances in Water Resources, 30, 937–953, https://doi.org/10.1029/93WR01409, 2007.

Castellarin, A., Botter, G., Hughes, D., Liu, S., Ouarda, T., Parajka, J., Post, M., Sivapalan, M., Spence, C., Viglione, A., and Vogel, R.: Prediction of flow duration curves in ungauged basins, in: Runoff prediction in ungauged basins: Synthesis across Processes, Places and Scales, edited by Blöschl, G., Sivapalan, M., Wagener, T., Viglione, A., and Savenije, H., 2013.

Castellarin, A., Burn, D., and Brath, A.: Assessing the effectiveness of hydrological similarity measures for flood frequency analysis, Journal of Hydrology, 241(3-4), 270-285, 2001.

Chow, V. T.: Handbook of applied hydrology, Mc-Graw Hill BookCo., New York, N.Y., 1964.

Dingman, S. L.: Planning level estimates of the value of reservoirs for water supply and flow augmentation in New Hampshire, Journal of the American Water Resources Association, 17, 684–690, https://doi.org/10.1111/j.1752-1688.1981.tb01277.x, 1981.

Duan, Q., Schaake, J., Andreassian, V., Franks, S., Goteti, G., Gupta, H., Gusev, Y., Habets, F., Hall, A., Hay, L., Hogue, T., Huang, M., Leavesley, G., Liang, X., Nasonova, O., Noilhan, J., Oudin, L., Sorooshian, S., Wagener, T., and Wood, E.: Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops, Journal of Hydrology, 320, 3–17, 2006.

Farmer, W. H., Archfield, S. A., Over, T. M., Hay, L. E., LaFontaine, J. H., & Kiang, J. E.: A comparison of methods to predict historical daily streamflow time series in the southeastern United States: U.S. Geological Survey Scientific Investigations Report. https://doi.org/10.3133/sir20145231, 2015.

Fennessey, N. and Vogel, R. M.: Regional Flow-Duration Curves for Ungauged Sites in Massachusetts, Journal of Water Resources Planning and Management, 116, 530–549, https://doi.org/10.1061/(ASCE)0733-9496(1990)116:4(530),1990.

Franchini, M. and Suppo, M.: Regional analysis of flow duration curves for a limestone region, Water Resources Management, 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 Resources Research, 45, 1–10, https://doi.org/10.1029/2008WR007472, 2009.

Hänggi, P. and Weingartner, R.: Variations in Discharge Volumes for Hydropower Generation in Switzerland, Water Resources Management, 26, 1231–1252, https://doi.org/10.1007/s11269-011-9956-1, 2012.

Hughes, D.A., Smakhtin, V.Y.: Daily flow time series patching or extension: a spatial interpolation approach based on flow duration curves, Hydrol. Sci. J., 41 851–871, 1996, doi:10.1080/02626669609491555.

Jothityangkoonad, C. and Sivapalan, M.: Framework for exploration of climatic and landscape controls on catchment water balance, with emphasis on inter-annual variability, Journal of Hydrology, 371, 154–168, 2009.

Kohler, M. and Linsley, R.: Predicting the runoff from storm rainfall, U.S. Weather Bureau Research Paper No. 34, p. 10, 1951.

Linsley, R., Kohler, M., and Paulhus, J.: Applied Hydrology, McGraw-Hill. New York, first edit edn., 1949.

Massey, F.J.: The Kolmogorov-Smirnov Test for Goodness of Fit. Journal of the American Statistical Association. 46 (253), 68–78, 1951.

McMahon, T. A., Laaha, G., Parajka, J., Peel, M., Savenije, H., Sivapalan, M., Szolgay, J., Thompson, S. E., Viglione, A., Wood, E. A., and Yang, D.: Prediction of annual runoff in ungauged basins, in:

Runoff Prediction In Ungauged Basins: Synthesis Across Across Processes, Places and Scales, edited by Bloschl, G., Sivapalan, M., Wagener, T., Viglione, A., and Savenjie, H., Cambridge University Press, 2013.

Mohamoud, Y. M.: Prediction of daily flow duration curves and streamflow for ungauged catchments using regional flow duration curves, Hydrological Sciences Journal, 53, 706–724, https://doi.org/10.1623/hysj.53.4.706, 2008.

Montanari, L., Sivapalan, M., and Montanari, A.: Investigation of dominant hydrological processes in a tropical catchment in a monsoonal climate via the downward approach, Hydrology and Earth System Sciences, 10, 769–782, https://doi.org/10.5194/hess-10-769-2006, 2006.

Nash, J.E. and Sutcliffe, J.V.: River flow forecasting through conceptual models part I – a discussion of principles, J. Hydrol., 10, 282–290, 1970.

Pugliese, A., Castellarin, A., and Brath, A.: Geostatistical prediction of flow-duration curves in an index-flow framework, Hydrology and Earth System Sciences, 18, 3801–3816, 2014.

Pugliese, A., Farmer, W. H., Castellarin, A., Archfield, S. A., and Vogel, R. M.: Regional flow duration curves: Geostatistical techniques versus multivariate regression, Advances in Water Resources, 96, 11–22, https://doi.org/10.1016/j.advwatres.2016.06.008, 2016.

Quimpo, R.G., Alejandrino, A.A., McNally, T.A.: Region- alized flow duration for Philippines, J. Water Resour. Plan. Manag., 109, 320–330, 1983.

Rianna, M., Russo, F., and Napolitano, F.: Stochastic index model for intermittent regimes: from preliminary analysis to regionalisation, Natural Hazards and Earth System Science, 11, 1189–1203, https://doi.org/10.5194/nhess-11-1189-2011, 2011.

Rianna, M., Efstratiadis, A., Russo, F., Napolitano, F., and Koutsoyiannis, D.: A stochastic index method for calculating annual flow duration curves in intermittent rivers, Irrigation and Drainage, 62, 41–49, https://doi.org/10.1002/ird.1803, 2013.

Ridolfi, E., Rianna, M., Trani, G., Alfonso, L., Di Baldassarre, G., Napolitano, F., Russo, F.: Advances in Water Resources A new methodology to define homogeneous regions through an entropy based clustering method. Advances in Water Resources, 96, 237–250, 2016.

Samaniego, L.: Heft 118 Hydrological Consequences of Land Use / Land Cover and Climatic Changes in Mesoscale Catchments von, Ph.D. thesis, 2003.

Samaniego, L., Bárdossy, A., and Kumar, R.: Streamflow prediction in ungauged catchments using copula-based dissimilarity measures, Water Resources Research, 46, https://doi.org/10.1029/2008WR007695, 2010.

Scott, D., Prinsloo, F., Moses, G., Mehlomakulu, M., and Simmers, A.: A re-analysis of the South African catchment afforestation experimental data. Water Research Commission, Pretoria, Report 810/1/00, Tech. rep., 2000.

Singh, S. K., Bárdossy, A., Götzinger, J., and Sudheer, K. P.: Effect of spatial resolution on regionalization of hydrological model parameters, Hydrological Processes, 26, 3499–3509, https://doi.org/10.1002/hyp.8424, 2012.

Smakhtin, V.Y.: Low flow hydrology: A review, Journal of Hydrology, 240, 147–186, 2001.

Smakhtin, V.Y.: Generation of natural daily flow time-series in regulated rivers using a non-linear spatial interpolation technique, Regul. Rivers Res. Manag., 15, 311–323, 1999.

Smakhtin, V.Y., Hughes, D.A., and Creuse-Naudin, E.: Regionalization of daily flow characteristics in part of the Eastern Cape, South Africa, Hydrological Sciences Journal, 42, 1997.

Smakhtin, V.Y. Masse, B.: Continuous daily hydrograph simulation using duration curves of a precipitation index, Hydrol. Process. 14, 1083–1100, 2000, doi:10.1002/(SICI)1099-1085(20000430)14:6<1083::AID-HYP998>3.0.CO;2-2.

Vogel, R.: Flow Duration Curves. I: New interpretation and confidence intervals, Journal of Water Resources Planning and Management, 120, 485–504, https://doi.org/10.1017/CBO9781107415324.004, 1994.

Vogel, R. M., Sieber, J., Archfield, S. A., Smith, M. P., Apse, C. D., and Huber-Lee, A.: Relations among storage, yield, and instream flow, Water Resources Research, 43, https://doi.org/10.1029/2006WR005226, 2007.

Weibull, W.: A statistical theory of the strength of materials. Ing. Vetensk. Akad. Handl., 151, 1–45, 1939.

Weiss, M.S.: Modification of the Kolmogorov-Smirnov statistic for use with correlated data, J. Am. Stat. Assoc., 73, 872–875, 1978. doi:10.1080/01621459.1978.10480116.

Westerberg, I. K., Guerrero, J.-L., Younger, P. M., Beven, K. J., Seibert, J., Halldin, S., Freer, J. E., and Xu, C.-Y.: Calibration of hydrological models using flow-duration curves, Hydrology and Earth System Sciences, 15, 2205–2227, https://doi.org/10.5194/hess-15-2205-2011, 2011.

Worland, S. C., Steinschneider, S., Farmer, W., Asquith, W., Knight, R.: Copula Theory as a Generalized Framework for Flow-Duration Curve Based Streamflow Estimates in Ungaged and Partially Gaged Catchments, 55(11), 9378-9397, 2019.

Xu, X.: Methods in Hypothesis Testing, Markov Chain Monte Carlo and Neuroimaging Data Analysis, Ph.D. Thesis, Harvard, 2014. https://dash.harvard.edu/handle/1/11108711.

Yu, P.-S. and Yang, T.-C.: Using synthetic flow duration curves for rainfall-runoff model calibration at ungauged sites, Hydrological Processes, 14, 117–133, https://doi.org/10.1002/(SICI)1099-1085(200001)14:1<117::AID-HYP914>3.0.CO;2-Q, 2000