Reply to Reviewer 1

We would like to thank William Farmer for the time dedicated to our paper and for his valuable comments that contribute to improve the paper.

In the following, reviewer's comments are in Italic (R2), Authors' comments are in normal text (AC). Some of the changes made in the manuscript by the Authors based on Referee's comments are reported below. Moreover, in the following, a summary of authors' changes in manuscript based on comments of the editor and all referees is given:

1. In the Introduction, we added a discussion on the works by Hughes and Smakhtin (1996) and Smakhtin and Masse (2000) clarifying the differences existing between our and their work.
2. The Methodology section was reorganized and improved to provide a clear description of the method. Moreover, for the sake of clarity, a figure was added.
3. We provided a clear statement of the hypothesis.
4. The sections are now organized on the base of the Introduction-Methods-Results-and-Discussion (IMRAD) format.
5. For both case studies we have used the same performance criteria and results are discussed in-depth.

R1: This work could benefit from a clear statement of hypotheses. In my opinion, the main hypothesis is that the daily flow duration curves at an ungauged location can be simulated with knowledge of the precipitation record at both the ungauged site and some index site. This hypothesis relies on a further assumption that the cumulative distributions of streamflow and precipitation correlate in some way. In the revisions I am proposing, I think the authors should clearly set out to quantifiably address these hypotheses. A figure may improve the understanding of the methodology.

AC: We thank the reviewer for the suggestion, we stated the hypothesis of the work (see comments below) and added a figure to give a clearer explanation of the procedure.

R1: As a previous commenter noted, the approach is difficult to understand and may be greatly simplified. The authors create cumulative distribution functions (CDFs) for streamflow and API at both an index site and a target site over some reference period. They then create a CDF of streamflow and API at the index site for some target period and a CDF for API at the target site for this same target period. The method then only uses the CDFs of (1) API at the index site in the target period, (2) API at the index site in the reference period, and (3) streamflow at the target site in the reference period.

AC: We thank the reviewer for the comment, we changed Sect. 3 to make it clearer and shorter.

R1: In addition to improving readability, revising the methods section might also address the concerns raised by the previous commenter. Namely, that it seems the approach could be greatly simplified through interpolation along the relevant relationships without the need for intermediate *exceedance probabilities. This would be accomplished by the following: (1) Create the CDF of API at the index site in the target period, (2) Plot the API of the index site in the reference period against the streamflow at the target site during the reference period, (3) interpolate each API from the target period, in order, along the curve created in (2) to produce the CDF of streamflow at the target site in the target period. While this approach is achievable algorithmically, and identical to the one proposed, it raises several concerns about the implicit assumptions. "Are API and streamflow ranked independently? (See step 3 of page 10.) If so, the implicit assumption is that the exceedance probability of the API on a given day is equivalent to the exceedance probability of the streamflow on that same day at the same site. This is a pretty sizeable assumption. As a start, it would be good to see if 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."*

AC: We thank the reviewer for the comment, the methodology section was largely modified and a clear statement of the hypothesis is provided. Specifically, it is also shown 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. For instance, in the manuscript the following table showing this correlation is reported for three catchments for four different periods of time.

	Correlation		
Period	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

R1: The second assumption arises when we move from the CDF of API at the index site in the reference period to the CDF of streamflow at the target site in the reference period. This movement introduces a second implicit assumption: namely, that the exceedance probability of the API on a given day at the index site is equivalent to the exceedance probability of streamflow on that same day at the target site. Put another way, if you accept the assumption in the previous paragraph, this step assumes that the temporal sequencing of API is identical at both sites. Again, this needs to be demonstrated: Is the temporal sequence of API exceedance probabilities at the index site highly correlated with the temporal sequence of streamflow exceedance probabilities at the target site in the reference period? It may be argued that the temporal sequencing is irrelevant. This is not the case. By assuming the same exceedance probabilities in step 7 of page 10, we are assuming a perfect correlation and, therefore, assuming a temporal correspondence.

AC: We thank the reviewer for the comment, in the revised manuscript we shown that the temporal sequence of API exceedance probabilities at the index site is highly correlated with the temporal sequence of streamflow exceedance probabilities at the target site in the reference period. As an example, the following table was added to the paper to report the correlation between Blanco (index site) and three target sites (specified in the table) for four different reference periods.

Sites	**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

R1: The third assumption, which was alluded to earlier, is that the CDF of API is identical across sites for both the index site and the target site in the same period. This is what allows the authors to use the CDF of the API of the index site in the target period for step 6 on page 10. It may be that this is what the authors meant by "the assumption of large scale precipitation" (line 16, page 9); if so, please clarify. Regardless, this assumption needs to be validated through correlation or a KS test.

AC: We thank the reviewer for the comment, in the revised manuscript we now show that the CDF of API is identical across sites for both the index site and the target site in the same period through correlation. In the table below an example of correlation between Blanco (USA) and three other sites is reported for four different time periods.

Sites	**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

Moreover, the distributions of API have the same type of distribution – Weibull is accepted for all of them. On the other hand, the distribution parameters may differ from site to site and from time period to time period. For instance, the Weibull is the best fitting distribution of the API at Blanco for the periods 1948-1968, 1968-1988, 1948-1963 and 1948-1958. The same applies to the API of the three sites above for the same periods. We now specify it in the paper as well.

R1: Without some quantifiable validation of these assumptions, the proposed method is tenuous at best and left vulnerable to criticism. With that in mind, and the comments of the previous commenter, I'd like to propose that exploring these assumptions might result in modifications of the methodology that might move away from the case of simple interpolation. Is the relationship between API and streamflow constant across periods or sites? Should API and streamflow be ranked independently or with some sort of dependence? Should the API of the index site in the target period be used to map to a different site in a different period (i.e., the target site in the reference period)? Exploring these questions, and validating the underlying assumptions, will produce a more robust approach."

AC: Thank you for this comment. We assume that the relationship between API and FDC is the same for the same site regardless the time period. For the reviewer we enclose the following table reporting the correlation at Blanco between API and q for different time periods. This correlation is stronger than the one obtained considering API and discharge from two different time periods.

Site	**1948-1968**	**1968-1988**	**1948-1963**	**1948-1958**
Blanco (USA)	0.977	0.995	0.997	0.970

R1: "In addition to their main hypotheses proposing this methodology, the authors assert that the FDC is a product of the basin and the weather. This is surely intuitive, but the evidence provided could be greatly strengthened. The authors use KS tests, but is unclear how they were applied. It would be informative to clearly communicate if the CDF of streamflow from one period and the CDF of streamflow from another period could be considered significantly different. The authors have done this, but the presentation is not clear. The extension would be to ask if the API can be correlated with any differences across time. (As an aside: Was there any discussion of selecting stationary sites? How would nonstationary behavior play a role here?)"

AC: The FDC seems to be significantly different from one time period to the other. The same applies to the API. In our opinion, this is not caused by non stationarity of the time series but more to some long memory effects. We show that this long scale variability is very visible in the discharge.

R1: This, in my opinion, raises another concern: The authors seem to be attempting to simultaneously address two very different problems. The first problem considers a target site that has a streamflow record overlapping with an index site, but the desired period has no overlap (the ungauged area is the same site, different period). In this case, the use of APIs within site, without an index site, would be most ideal. The second problem considers a site without any streamflow information; this situation necessitates the use of an index. Of course, when there are gaps in the API record as well, this transforms into four unique problems. Regardless, if we believe the underlying assumption that the CDF of streamflow is a product of basin and weather, then the solutions to these problems must be quite different. The first asks if knowledge of new weather can produce the CDF of streamflow, while the second alters both variables and asks if the CDF relationship can be transferred across weather and basin. Line 8 of page 3 implies that both problems are considered, but the remainder of the paper seems only to address the partially gauged site. I would advise addition of the second problem or, at least, a discussion of implication for the second problem (completely ungauged).

AC: We thank the reviewer for the comment, the methodology explained in the paper needs to be applied to partially gauged basins, we eliminated the reference to totally ungauged sites in the paper.

R1: In 1996, Hughes and Smakhtin (<https://doi.org/10.1080/02626669609491555>), among others, provided a technique for hydrograph simulation using flow duration curves. While their focus was on hydrographs, the extensions to ungauged FDCs can be made quite clearly (i.e., they could be derived from simulated hydrographs). Smakhtin and Masse (2000: <https://doi.org/10.1002/(SICI)1099-1085(20000430)14:6%3C1083::AID-HYP998%3E3.0.CO;2-2>) then extended this method to use a precipitation index. While I believe that the methods presented here are different, the novelty of this new method must be strongly articulated.

AC: In the new version of the manuscript we recalled the two papers in the Introduction section. We added a deep discussion on the assumptions underlying their work and highlighted which are the differences of our work.

R1: "I strongly encourage the authors to revisit the style of the manuscript. At times, it feels a bit disjointed and it may be improved by enforcing a strict Introduction-Methods-Results-and-Discussion (IMRAD) format. For example, section 3.1 is ostensibly a methods sections but presents a series of results that I think are pivotal to the paper (line 19, page 8). Similarly, the paragraph on page 13 and section 4.1 present new methods of analysis that have not been presented earlier in the paper. While

IMRAD is not a requirement, I do suggest thinking carefully about the best approach to presenting the narrative."

AC: We thank the reviewer for the suggestion, we performed a restyle of the paper. The results reported in Sect.3.1 are pivotal to the development of the paper, they are now moved under a paragraph named "Preliminary analysis". The presentation of the performance criteria was moved into the Methodology

R1: "In my opinion, this work needs more presentation and discussion of quantified results.

AC: In the updated version of the paper, we have further developed the discussion.

R1: "The results sections heavily rely on visualization. Even the presentation of metrics in section 4.1 is visual. While this is useful, we still need to see some discussion on the performance metrics. For example, the scale on NSE in Figure 11 makes all positive values appear as a single color. This presentation means we can't honestly see how the methods perform."

AC: We chose to present results in a visual style to better show them. Because of the large number of sites and the large number of time windows we have investigated, to a reader it would take too much time to go through a tabular presentation, while plots have an immediate impact. However, we agree with the reviewer regarding the scale of the plots that sometimes make difficult to understand the goodness of the results. We have now solved this issue.

R1: Page 1, line 17: When talking about general duration curves, more commonly known as cumulative distribution functions, it is better to say "exceedance frequency" rather than "exceedance time".

AC: We replaced the word as suggested.

R1: Page 2, line 1: Please provide the citation for the Weibull plotting position.

AC: The citation was added (i.e., Weibull, W., 1939: A statistical theory of the strength of materials. Ing. Vetensk. Akad. Handl., 151, 1–45).

R1: Page 3, line 4: Please provide more discussion and literature of this important point.

AC: The sentence was referring to the results of the paper anticipating them for the readers. We rephrased the paragraph in the Introduction and we have introduced this concept in the Conclusions section together with an explanation.

R1: Page 3, line 7: It is not clear what the "distribution of the FDC" is. The FDC is a distribution, so it is confusing to talking about the distribution of a distribution.

AC: We rephrased the sentence as "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. [5]."

R1: Page 4, line 6: Florida, Louisiana and Texas are certainly not the East coast. I would suggest the Gulf Coast.

AC: We rephrased the sentence

R1: Page 5, line 7: misspelling of database

AC: We carefully checked the English spelling throughout the paper.

R1: Page 5, line 2 (?): This is an example of inconsistent citation style. Bloeschl should be in parenthesis.

AC: We carefully checked the citation style throughout the paper.

R1: Page 8, line 5: Please provide citation to KS test.

AC: We provided the citation (i.e. Massey, F. J. The Kolmogorov-Smirnov Test for Goodness of Fit. Journal of the American Statistical Association. Vol. 46, No. 253, 1951, pp. 68–78.).

R1: Page 9, line 10: What lead to this choice for alpha? (Also, note that the same symbol is used earlier in this section for significance: page 8, line 10.) Please provide citation or summary of initial exploration.

AC: Now α refers to the API only. We specified in the paper that 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, API represents the long memory of the basin as it includes the effect of precipitation occurred many days before. To capture this behavior, in this study α is chosen equal to 0.85. Moreover, this is in agreement with a previous study [6] which investigated the same case study area (i.e. Neckar catchment).

R1: Page 9, line 18: I strongly suggest referring to the "reference site" as an "index" or "donor". The reference connotation implies lack of human influence that might be confusing. The same could be said of the reference period

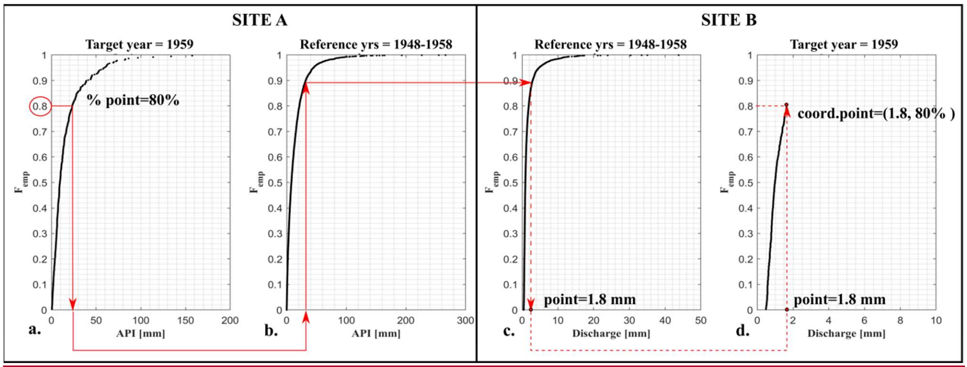
AC: We are now using "donor" site instead of "reference" one.

R1: Page 10, line 17: The series of Nr and Nt are both being indexed with i, which leads to confusion.

AC: Please notice that we consistently revised the Methodology section, however we changed the Appendix where both *Nt* and *Nr* exist.

R1: Page 10, line 19: So, API_Ati is equal to API_Arj?

AC: Yes, it is. We added the following Figure to better clarify the methodology and also we re-organized and strongly improved the methodology section.



R1: Page 11, line 4: What is a supporting variable? This is not described as such earlier, so it surprises the reader.

AC: The supporting variable is the one used to retrieve the streamflow values, in this example is the API. For the sake of clarity, we have now called it "proxy" variable.

R1: Page 11, line 17: "good agreement" is very subjective. Please provide thorough, quantifiable analysis. For example, a lot of the curves in figure 6 look rather poor for highs and lows (top row, second box from the left).

AC: The sentence is an introduction to the extensive explanation presented in the following lines of the manuscript. The paragraph reporting the results in terms of performance measures is more extended and further shows how the method has a higher performance per intermediate flows, while it is poorer for high and low flows. Nevertheless, the error in terms of Mean Absolute Error is small also for high and low flows, showing an overall good performance of the approach also for extreme flows.

R1: Figure 5: Why was the box for ref:68-88 and tar: 88-98 not included? The caption needs to do a better job of describing the different panels.

AC: The missing panel would not add more information to the paper, therefore because of the lack of space the panel was not included. We better described the panels in the caption.

R1: Page 12, line 9: Spelling of FDCs

AC: We carefully checked the spelling throughout the paper.

R1: Page 13: The methods for this paragraph were very unclear to me. Could a figure or a revision help?

AC: In this paragraph we compute the moving average to show that the between-year variability of the discharge of a specific percentile can be high. Therefore, this suggests that percentiles cannot be considered an invariant characteristic of the basin and thus they cannot be estimated using geographic and morphologic characteristic of the basin only. We decided to show the moving averages of these specific percentiles as they are the most used ones.

R1: Page 14: Please provide the citation for NSE. Even better, a metric like KGE might be more appropriate

AC: The citation was added. We used the NSE, despite the criticism it has received (e.g.,[7]) because of the familiarity most hydrologists and meteorologists have with it [8], facilitating the interpretation of the obtained values.

R1: Figure 10: What is the horizontal axis of this figure?

AC: Figure 10 was replaced with a revised figure

References

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

[2] V.Y. Smakhtin, D.A. Hughes, E. Creuse-Naudin, Regionalization of daily flow characteristics in part of the Eastern Cape, South Africa, Hydrol. Sci. J. 42 (1997).

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

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

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

[6] T. Sugimoto, Copula based Stochastic Analysis of Discharge Time Series, 2014.

[7] H. V. Gupta, H. Kling, K.K. Yilmaz, G.F. Martinez, Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, J. Hydrol. 370 (2009) 80–91.

[8] D.N. Moriasi, J.G. Arnold, M.W. Van Liew, R.L. Bingner, R.D. Harmel, T.L. Veith, Model evaluation guidelines for systematic quantification of accuracy in watershed, Simulations, Trans. Am. Soc. Agric. Biol. Eng. 50 (2007) 885–900.

Reply to Reviewer 2

We would like to thank Thomas Over for the time dedicated to our paper and for his valuable comments that contributed to improve the paper.

In the following, reviewer's comments are in Italic (R2), Authors' comments are in normal text (AC). Some of the changes made in the manuscript by the Authors based on Referee's comments are reported below. Moreover, in the following, a summary of authors' changes in manuscript based on comments of the editor and all referees is given:

1. In the Introduction, we added a discussion on the works by Hughes and Smakhtin (1996) and Smakhtin and Masse (2000) clarifying the differences existing between our and their work.
2. The Methodology section was reorganized and improved to provide a clear description of the method. Moreover, for the sake of clarity, a figure was added.
3. We provided a clear statement of the hypothesis.
4. The sections are now organized on the base of the Introduction-Methods-Results-and-Discussion (IMRAD) format.
5. For both case studies we have used the same performance criteria and results are discussed in-depth.

R2: This paper presents a method for estimating FDCs during an ungauged period at a "target" location that is gauged during another period. [...].As such it is similar to a record extension application of the approach of Smakhtin and Masse (2000, Hydrological Processes, Vol. 14, pp. 1083-1100), except they estimated daily flow in ungauged basins, and to the work of Hughes and Smakhtin [...] Ideally, however, the authors would investigate the distinction and show how in application one might make different choices regarding parameters or selection of the reference basin when estimating the FDC as opposed to daily streamflow.

AC: In the new version of the manuscript, we recalled the two papers in the Introduction section. We added a deep discussion on the assumptions underlying their work and highlighted which are the differences of our work.

R2: (1) Why select another gauged basin and use only its API, not its streamflow?

AC: For the German case study, we derived the streamflow and thus the FDC of each specific basin from the streamflow of another basin, this was now better specified in the text. We reported the evaluation metrics of the procedure to show the goodness of both methods.

p.C2: (2) Why not use the API at the target basin?

Because to retrieve the discharge at the target basin gauged during the target period, the API is not necessary at the target site.

R2: 2.a. The idea of using the reference gauge discharge is raised and in the first presentation of the methodology on page 10, the FDC at the reference basin is computed, but it is never used.

AC: We improved the presentation of the methodology and removed what is unnecessary to the development of the procedure.

R2: 2.b. The results for the US basins are presented quite differently than those of the German basins, including using different performance criteria.

AC: We decided to show results, in terms of flow duration curves, of the US catchments only to reduce the amount of possibly redundant data. In the revised version of the manuscript, we used the same estimation metrics for both sites.

R2: 2.c. How the performance criteria could be applied for individual predictions was not clear to me.

AC: We explained it the text (Sect. 3.3) with the following text: "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 percentile."

R2: 2.d. The choice of basins for the study seems rather arbitrary. For example, there are hundreds of basins in the MOPEX, including many others that do not have much snow.

AC: We decided to use these catchments also because the land use did not consistently change in the time window we used for analysis.

R2: 2.e. The consideration of energy versus water limitation as a measure of similarity is interesting but it is not clear that it is relevant when API is being used.

AC: Annual streamflow variability is driven by the availability of water (i.e., provided by the precipitation) and energy (i.e., the evapotranspiration), [5]. The API combines the precipitation to provide the amount of the water released by the soil. Therefore, it is relevant to distinguish between water and energy limited catchments, as their behavior is different.

R2: 1.a. Are there snow effects in the Upper Neckar basin? How addressed?

AC: 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 water which is then used for the calculation of the API. We specified it in the text.

R2: 1.b. Do you take the karstic effects on Upper Neckar flows into account?

AC: For the karstic catchments, a direct transfer seams not to be plausible. However, the temporal stability of the API/Runoff can be considered as invariant.

R2: 2.a.i. Figure 3: Perhaps plot and check correlation with temperature of ET/P instead of just ET?

AC: The plot and the correlation of temperature and mean annual runoff (Q/P) already provide a similar information. It would be redundant.

R2: 2.a.ii. Figure 3: Need to consider uncertainty around correlation estimates: for the Peace R. it seems unlikely that rho = 0.027 is a significantly positive value; rather probably this basin is balanced between energy and water limitation by this criterion.

AC: We thank the reviewer for the suggestion, we added the following consideration to the paper: "For instance, measurements at Peace River (LA) suggest that the catchment is balanced between energy and water limitation by the correlation criterion, Figure 3 upper panel."

R2: 2.b. Last sentence on p. 6: It is not possible…" It sounds plausible, but has this assertion been tested? The statement itself is very categorical; in fact there are degrees of water and energy limitation. How different do they need to be to make this true (if it is)? In particular, for the present application, the methodology may account for the water versus energy limitation; it may be that the timing of the weather is the most important thing to have in common.

AC: We tested the assertion. As a result it was not possible to estimate streamflow values of a water-limited catchment from the data of an energy-limited ones. See also reply to comment 2e (above).

R2: 3.a.i. Sect. 3.1. 2nd paragraph: It seems there should already be a well-established way of addressing autocorrelation effects on the K-S test.

AC: Weiss [6] proposed a methodology to account for modifying the K-S test for autocorrelated data. Later, Xu [7] suggested a method that can be applied to two sample test. However, our way to take into account for the autocorrelation is easier to implement and has a nice interpretation of equivalent sample size adjustment. More importantly, our method can be easily generalized to two sample test. We introduced the following paragraph in the section: "… Since the streamflow data presents autocorrelation, the autocorrelation effects the KS test. Weiss (1978) proposed a methodology to account for modifying the K-S test for autocorrelated data. Later, Xu (2013) suggested a method that can be applied to two sample 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 the equivalent sample size. It is easier to implement and more importantly, it can be easily generalized to two sample test. We can assume that the autocorrelation effect attenuates after three days..."

R2: 3.a.ii. a. Sect. 3.1. 2nd paragraph: The last two sentences of this paragraph seem to be referring to a test on a particular basin, but they are stated as if these relations are generally (i.e., mathematically) true. Which is it?

AC: This example is given for streamflow values at daily resolution recorded during a year, thus for a time series of 365 values. We specified it with the following sentence in the paper: "For instance, let's take as an example the 1 year FDCs. If the samples were three times smaller and for instance their length would equal 122 …"

R2: 3.b. 3rd paragraph: This paragraph seems to include "Results", not "Methodology".

AC: In this paragraph, we provide details about the methodology, but we also anticipate results regarding the KS test. This is done because those results explain the reasons why we applied the methodology presented in the following part of the paper to estimate the FDCs. For the sake of clarity, we decided to move the Sect. 3.1 in Sect. 2 and name it "2.4 Preliminary analysis".

R2: 4.a. Section 3.2: First paragraph: It is not always true that the non-weather properties (land use) do not change. Did you check that your study basins satisfy this assumption?

AC: We considered basins where the land use did not deeply change in time.

R2: 4.b. Section 3.2: Last sentence: Did you test different values of alpha other than 0.85, or just select that value for the reasons given?

AC: We tested also other values of alpha. Then, we decided to proceed with alpha=0.85 because when α tends to 1, API represents the long memory of the basin as it includes the effect of precipitation occurred many days before. Moreover, this is in agreement with the value used by [8] for the Neckar Catchment. Therefore, we added a line: "To capture this behavior, in this study α is chosen equal to 0.85, this is in agreement with a previous study by Sugimoto (2014) who investigated the same case study area (i.e. Neckar catchment)."

R2: 5.a Section 3.3. First paragraph, last sentence: Why do you assume "large scale precipitation"? What do you mean by that?

AC: Small-scale variability of rainfall can be assumed to vary in a range lower than 10–20 km [9]. Therefore instead of APIs calculated from point precipitation areal precipitation is considered. The wording may be inappropriate and was changed.

R2: 5.b. Last complete paragraph on p. 10: It seems it would be better to interpolate between Pj and Pj+1 rather than taking the mean, but it may not make a lot of difference.

AC: Thank you for the suggestion, as you anticipated, the difference is not significant.

R2: 6.a. Section 4, p. 11, discussion of figures 5-8: a. Several statements regarding goodness of fit are made without being quantified. However the K-S technique has been presented and could be applied; indeed, it would be ideal to provide K-S test results to accompany the results in each panel of these plots.

AC: Thank you for the suggestion, the KS distance D* was added to each panel.

R2: 7.a. Section 4, p. 13, figure 10 and discussion of it: a. Why present 30, 70, 90, and 99th percentiles? As one can see, 90th and 99th (though the lower right panel of figure 10 is labeled as the 95th percentile), are almost the same. The complementary percentiles, i.e., 70, 30, 10, and 1st percentiles (exceedance probabilities) would be more interesting, in my opinion.

AC. We chose these percentiles as they are flow percentiles usually investigated in literature. For instance, the approach by Franchini and Suppo [10] regionalises these streamflow quantiles.

The title of the plots is now consistent with the caption.

R2: 7.b. You say (lines 5-6 of p. 13): "it is not possible to estimate the flow quantiles using regression methods that do not take into account the weather characteristics." This may be an over-statement. You have demonstrated that if you want to transfer across time, weather fluctuations need to be considered. But for prediction at ungauged basins for a fixed period of time, that may not be true

AC: The moving average is computed to show that the between-year variability of the discharge of a specific percentile can be high. Therefore, this suggests that percentiles cannot be considered an invariant characteristic of the basin and thus they cannot be estimated using geographic and morphologic characteristic of the basin only.

This is true also for prediction at ungauged basins for a fixed period of time as we demonstrated applying the KS test to streamflow values gauged during the same time window at two different sites (see Sect.2.4).

R2: 1. Section 3.1, 3rd paragraph: This paragraph seems to include "Results", not "Methodology.

AC: In this paragraph we provide details about the methodology, but we also anticipated results regarding the KS test. This is done because the results justify the methodology we present in the paper to estimate the FDCs. This paragraph is now in a new one named "Preliminary analysis".

R2: 2. Section 3.3, in steps 2&3 of the "procedure to predict" (p. 10), the FDCs of the reference catchment A is computed, but it does not seem to be used in the procedure.

AC: The methodology section is now revised.

R2: 3. Section 3.3, step 8 of the "procedure to predict" (p. 10): Suggest "qBrj is taken to be the value of discharge that occurred…" rather than simply "qBrj is the value of discharge that occurred…".

AC: The methodology section is now revised.

R2: 4. Section 3.3, last paragraph (p. 11): It is stated here that in the paper both discharge and precipitation will be used as the support variable. But everything before indicates that only precipitation will be used. And I don't see any results using discharge as the support variable.

AC: For the German case study, the support variable is the discharge. We better highlighted it, moreover the support variable is now recalled as "proxy" variable.

R2: 5.a. Section 4, figures 5-8: From what period is this FDCref_site that is plotted? As it does not seem to be used in the procedure, why is it plotted?

R2: 5.b. Section 4, figures 5-8: I think however you should add the FDC of the target site during the reference period to these plots so the reader can see how much the FDC has changed from reference period to target period.

AC: Since usually the FDC of a donor site is used to retrieve the FDC of a target site, the FDCref_site was plotted (now called FDC_donor site) to show the difference between the FDC at the donor site and the FDC at the target site recorded during the same period of time.

R2: 6.a.i. Section 4, figure 9 and discussion of it: Discussion of figure 9 on p. 12, lines 9&10: "Results shown that the distance between the former pairs is bigger than the distance between the latter, Figure 9.":i. I don't think you ever defined the K-S distance. That needs to be done.

AC: The K-S distance is now defined as:

"Moreover, the test allows us to estimate the distance between couple of FDC:

$$D^* = \max_x \left(\left| F_1(x) - F_2(x) \right| \right),$$

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."

R2: 6.a.ii. Section 4, figure 9 and discussion of it: a. Discussion of figure 9 on p. 12, lines 9&10. I am willing to believe this assertion is true, but it is hard to see just from the plot. Can you provide some summary results such as the mean and median difference between 9 (top) and 9 (bottom) to give evidence of the assertion.

AC: We now provide the following summary to evidence the findings: "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."

R2: 6.a.iii. Section 4, figure 9 and discussion of it: a. Discussion of figure 9 on p. 12, lines 9&10. This assertion should be restated without the shorthand of "former" and "latter". It is hard to understand the way it is currently phrased, and it is a very important point.

AC: We rephrased the sentence "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. 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 have the same distribution."

R2: 7. Section 4, pp. 11-14: It is not clear why the Results section starts by giving a lot of results for the U.S. catchments and none for the German ones.

AC: We decided to show results of U.S. in a comprehensive way (both FDCs and performance criteria are shown) to keep compact the manuscript avoiding redundant plots. For the German case study, we shown the performance criteria which are much more representative than the FDCs. The performance criteria are shown in an extensive way as they are reported for both case studies.

R2: 8.a. Section 4.1, pp. 14-17, Definition of performance criteria: Are all these computed for Q in mm units? Even though those are units used throughout, it would be worth re-emphasizing that here.

AC: Yes, Q is in mm. We specified it.

R2: 8.b. BIAS: This is not a simple bias as it is normalized by Qsim; it is more like a relative bias or "relative mean error"; however usually one divides by Qobs. Actually, ME (defined later) is more like a simple bias.
c. Why apply different criteria for the German catchments?
d. "Ratio":
i. Can you give it a more meaningful name?
ii. This formula looks odd. If the square root were only on the numerator, it would be the standard error divided by the mean error (and the quantity would be non-dimensional). But why apply the square root to the mean error in the denominator?

AC: We thank the reviewer for the suggestions, we used the same metrics for both case study areas. We reviewed the estimation metrics, there was a typo in the BIAS formula, the correct form is now reported in the paper and below. Results were estimated with the following formula in agreement with Castellarin et al. [11]

$$\mathrm{BIAS} = \frac{1}{N} - \sum_{i=1}^{N}\left(\frac{Q_{sim,i} - Q_{obs,i}}{Q_{obs,i}}\right).$$

R2: 9. Section 4.1, figures 11-13: Many of the colors these figures are shifted so each box has more than one color, making them hard to interpret. This effect needs to be fixed.

AC: We are sorry for this issue, we fixed it.

R2: 10. Section 4.1: I don't see how the Performance criteria were applied to create the results shown in figures 11-13. As I understand, there is only one prediction of each quantile, for a fixed reference catchment and decade. Then how does one do the summations indicated in the performance criteria formulae? Following the definition of NSE on p. 15 it says: "N is the number of discharge values related to a specific percentile". How many of those are there? Is there ever more than one? If so, how? The situation with the correlation coefficient values presented in figure 14 seems to be the same: How does one compute correlation coefficients without multiple values? If there are multiple values, where are they coming from?

AC: We better explained it the text with the following sentence: "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 percentile."

R2: 11.a. Conclusions: p. 21, lines 2-5: "Here it is shown that two FDCs built for the same catchment, but with data corresponding to two different time windows, cannot be regarded as the same continuous distribution. The same results when two FDCs of two different catchments built for the same time window are analysed. Thus, it is not possible to infer a FDC using parameters retrieved from the distribution of another FDC without considering the weather." The first sentence supports the assertion in the third, but the second does not. If two different catchments experience possibly similar weather but produce a different streamflow, the cause is not the weather.

AC: We thank the reviewer for the comment. The sentences is now: "We show that two FDCs built for the same catchment 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."

R2: 11.b. Conclusions: p. 21, lines 13-14: "Since precipitation data series are characterized by a high number of zeros, here we used the Antecedent Precipitation Index (API)." This statement misses the more important fact that the API combines in a streamflow-like way the history of the precipitation. (A similar statement is made at the beginning of section 3.2 near the bottom of p. 8.)

AC: We thank the reviewer for the suggestion. We have rephrased as:

"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 as it provides the amount of precipitation released by the soil throughout the time."

R2: 11.c. p. 22, lines 26-27: Qualitative statement about similarity in shape from beginning of Section 4 is repeated. This assertion needs to be quantified somehow.

AC: A quantitative assessment of the goodness of the methodology is performed through the performance criteria. To describe why there can be a difference between the two FDCs, we rephrased the sentence as:

"The difference between the interpolated and observed FDCs can be due to the different temperature values characterizing the donor and target catchments."

References

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

[2] V.Y. Smakhtin, D.A. Hughes, E. Creuse-Naudin, Regionalization of daily flow characteristics in part of the Eastern Cape, South Africa, Hydrol. Sci. J. 42 (1997).

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

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

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

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

[7] X. Xu, Methods in Hypothesis Testing, Markov Chain Monte Carlo and Neuroimaging Data Analysis, 2014. https://dash.harvard.edu/handle/1/11108711.

[8] T. Sugimoto, Copula based Stochastic Analysis of Discharge Time Series, 2014.

[9] D.K. Nykanen, E. Foufoula-Georgiou, W.M. Lapenta, Impact of Small-Scale Rainfall Variability on Larger-Scale Spatial Organization of Land–Atmosphere Fluxes, J. Hydrometeorol. 2 (2001) 105–121.

[10] M. Franchini, M.Suppo, Regional analysis of flow duration curves for a limestone region, Water Resour. Manag. 10 (1996) 199–218.

[11] A. Castellarin, D. Burn, A. Brath, Assessing the effectiveness of hydrological similarity measures for flood frequency analysis, J. Hydrol. 241 (2001) 270–285. doi:10.1016/S0022-1694(00)00383-8.

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

Formatiert: Links, Einzug: Links: 0 cm, Hängend: 1.13 cm, Abstand Vor: 12 Pt., Nach: 0 Pt., Zeilenabstand: einfach, Keine Absatzkontrolle, Leerraum zwischen asiatischem und westlichem Text nicht anpassen, Leerraum zwischen asiatischem Text und Zahlen nicht anpassen

Formatiert: Schriftart: Nicht Fett, Kursiv, Schriftfarbe: Automatisch

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 at a specific site has a key role to the knowledge of the 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 of its importance, because of the lack of streamflow gauging stations, the FDC construction can be a not straightforward task. In partially gauged catchments, FDCs are usually built using regionalization methods among the others. 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 (i.e., donor catchment). 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. 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

Formatted: Top: 2.5 cm, Header distance from edge: 1.25 cm, Footer distance from edge: 1.25 cm

Field Code Changed

Deleted: important

Deleted: ungauged or

Deleted: of time

Deleted: paper

Deleted: reference

Deleted: reference

Deleted: reference

Deleted: reference

Deleted: analysed

Deleted:

Deleted: time

Deleted: time

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

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 request 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 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 vegetation change (Brown et al., 2005).

Spite of FDC importance, 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. 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

Deleted:

Deleted: O

Deleted: FDCs have a

Deleted: in hydrology for

Deleted: analysis

Deleted: as

Deleted: testified by the

Deleted: For instance,

Deleted: s

Deleted: are

Deleted: Through FDCs it is possible to evaluate the available flows for intake purposes. Indeed, the FDC provides information about the reliability of the water resource for water abstraction activities (Dingman, 1981).

Deleted: are

Deleted: they

Deleted: estimated on the basis of the FDC

Deleted: s

Deleted: ¶

Deleted: s

Deleted: have

Deleted: The

Deleted: methodology

Deleted: s

Deleted: s

Deleted: are

Moved (insertion) [4]

Deleted: The determination of streamflow characteristics at ungauged sites was one of the main objective of the IAHS's initiative in the Predictions in Ungauged Basins (PUB) decade framework (Sivapalan et al., 2003).

Deleted: techniques

Deleted: s

Deleted:

Deleted: Moreover,t

Moved up [4]: The determination of streamflow characteristics at ungauged sites was one of the main objective of the IAHS's initiative in the Predictions in Ungauged Basins (PUB) decade framework (Sivapalan et al., 2003).

Formatted: Font: (Default) Times New Roman, 12 pt 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 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. 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 studies are presented and catchments 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 applied to a set of catchments located in Germany and in U.S. 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),

Deleted: A

Deleted: s

Deleted: ¶
In this work we present a methodology to predict FDCs in poorly gauged and ungauged basins using a transformation based on the corresponding weather conditions. Smakhtin and Masse (2000) used the weather at a donor site to extend the daily hydrograph at a "destination" site through the monthly FDC of the destination site itself. The monthly FDC at the "destination" site is found 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). The procedure presented by Smakhtin and Masse (2000) is an extension of a previous work proposed by Hughes and Smakhtin (1996) to extend and/or filling in daily flow time series. The drawback of the procedures proposed by Hughes and Smakhtin (1996) and Smakhtin and Masse (2000) is the necessity of retrieving the monthly FDC of the target site with well-known literature methods before applying the methodology to extend the hydrograph. On the other hand, the novelty of the approach we propose is the possibility to retrieve the FDCs at partially ungauged sites from weather recorded at a donor catchment. Thus, without the need of applying procedures such as the regionalization. Our Aanalysis show that

Deleted: periods of time

Deleted: This implies that it is not possible to use the same FDC when dealing with different time windows. Moreover, i

Deleted: Results show This also implies that the FDC is not a property of a specific basin, because different time windows (such as decades) relate to different FDCs. ¶

Deleted: parameters

Deleted: FDC

Deleted: use the distribution of a FDC built at a location to determine the distribution of the FDC at another location based on catchment properties only.

Deleted: partially and completely

Deleted: Indeed, there still remains the question of how the FDC can be determined in case of missing streamflow data.

Deleted: Here we want to define a methodology to build FDCs facing the aforementioned issues. T

Deleted: main

Deleted: hat it is possible to

Deleted: a

Deleted: for a given target period using

Deleted: has been

Deleted: East

Deleted: U.S.A.

Deleted: Upper Neckar

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).

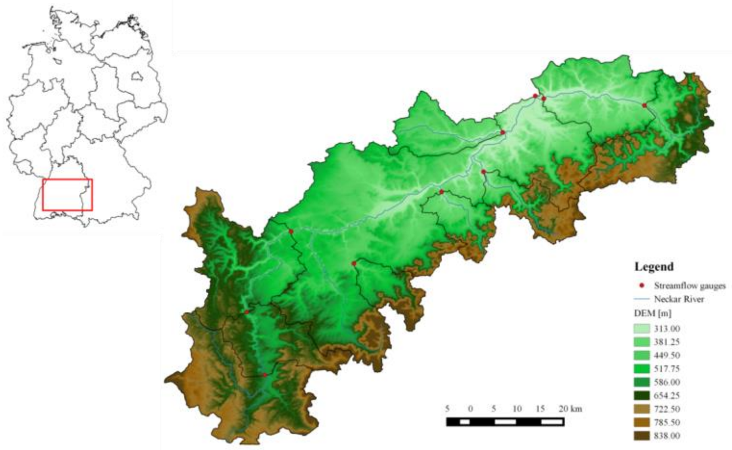


Figure 1. Streamflow gauges (red circles) used to test the methodology in the corresponding catchments located on the Upper Neckar River, Germany.

Deleted: green

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

2.2 USA catchments

The catchments on the Gulf coast of the USA are located in three different States: Florida, Louisiana and Texas, Figure 2. 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 for different time windows, Table 2.

Deleted: .
Deleted: .
Deleted: .
Deleted: East
Deleted: .
Deleted: .
Deleted: .
Deleted: characterised

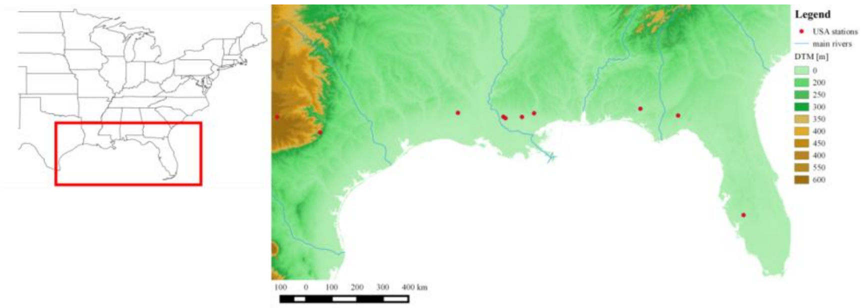


Figure 2. Streamflow gauges (red circles) used to test the methodology in the corresponding USA catchments.

Deleted: .
Deleted: .
Deleted: .

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).

Deleted: ¶

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

Deleted: .
Deleted: .A.

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

Formatted: Swedish (Sweden)

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 Energy and water limited catchments

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 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 since the catchment infiltrates 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 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 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:

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

Deleted: , Blöschl et al. (2013)
Deleted: As a consequence
Deleted: ¶
Deleted: a
Deleted: impact
Deleted: behaviour
Deleted: analysed

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

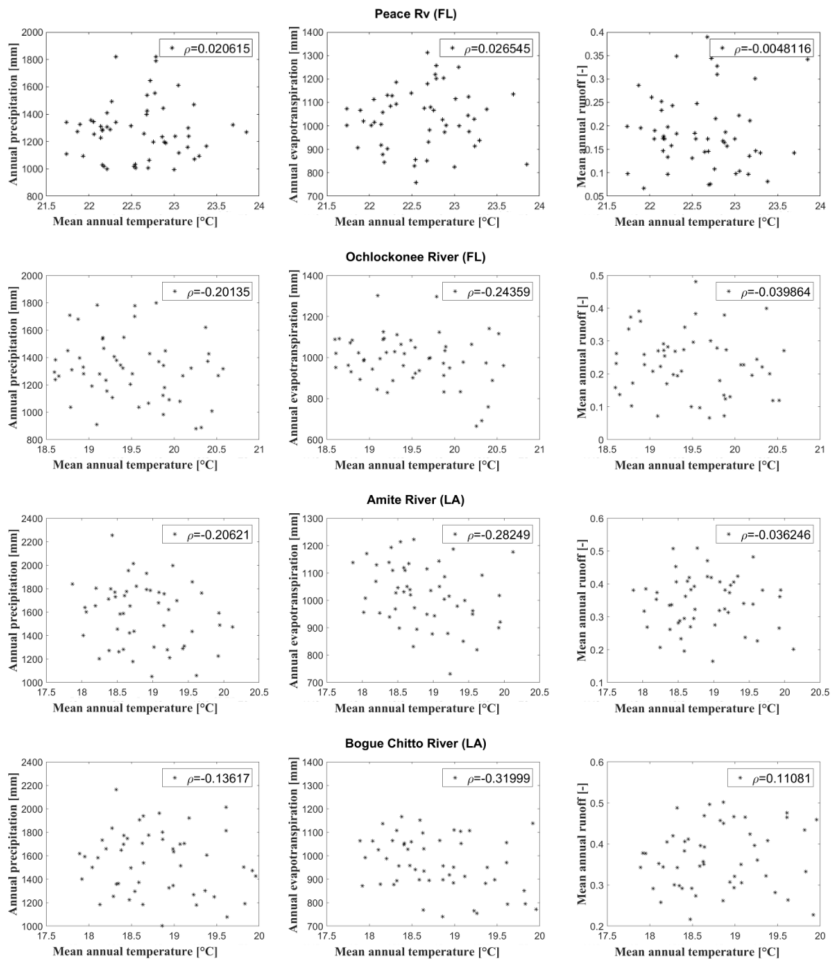


Figure 3. Annual precipitation against mean annual temperature (left panel), annual evapotranspiration against mean annual temperature (middle panel) and annual runoff coefficient against mean annual temperature (right panel) for four different catchments: Peace River (FL), Ochlockonee River (FL), Amite River near Denham Springs (LA), Bogue Chitto River (LA). In each plot, the Pearson correlation coefficient ρ is reported in box.

Results show that catchments 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; in the latter the evapotranspiration is limited by the available energy which happens in energy-limited catchments. For instance, measurements at Peace River (LA) suggest that the catchment is

Deleted: behaviours

balanced between energy and water limitation by the correlation criterion, Figure 3 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 from the data set of an energy-limited one, analysis have been carried out on climatically homogeneous sets of basins.

2.4 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, 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:

$$D^* = \max_x(|F_1(x) - F_2(x)|), \qquad (3)$$

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 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 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 FDC sampled at another location. Therefore, it is necessary to develop a methodology that accounts for the weather as it is main driver of FDCs variability as shown in the following. Figure 4 shows how

Deleted: water-limited

Deleted: , Figure 3

Deleted: Data p

Deleted: couples

Deleted: -

Deleted: following

Deleted: 's

Deleted: 3

Deleted: R

Deleted: period of time

different can be FDCs built at the same location using streamflow data gauged during different time windows.

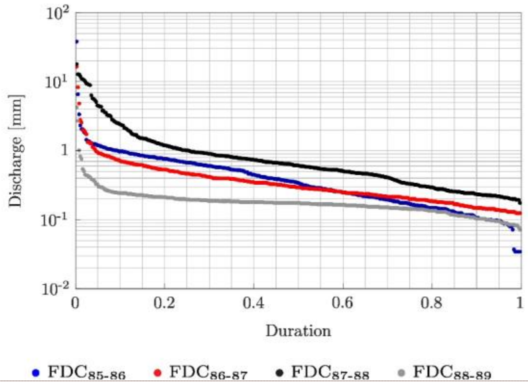


Figure 4. FDCs built for Tangipahoa River (FL) for four different hydrological years. Every hydrological year starts in October and ends the following September.

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:

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

where k stands for the location, 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 additional errors and it may be biased for long subperiods. Thus, the main idea is to get rid of a complicated non-linear processes 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):

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

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

Deleted: ¶ ¶

Moved (insertion) [1]

Deleted: ¶

Deleted: ¶

Deleted: are

Deleted: main idea

Deleted: have an approximation

Deleted: of

Deleted: basic

Deleted: wa

Deleted: ¶
The FDC can be interpreted as distribution function of discharge over a given time period. To determine if samples are drawn from the same distribution, the Kolmogorov-Smirnov test is carried out on each pair of samples. Then, in the followings, the methodology proposed in this paper to build a FDC at a location using precipitation and discharge data gauged at another location is described step by step. Since the methodology involves the use of the API index, it will be introduced in this Section.¶
3.1 Application of the Kolmogorov-Smirnov test to FDCs ¶
To determine if two FDCs built either at a location in different periods or at two different locations in the same period of time could be regarded as the same distribution, the Kolomogorov-Smirnov (KS) statistic is performed. 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 alpha. Here, alpha is set equal to 5%. ¶
The KS statistic is applied on daily streamflow data sampled in several periods of record (e.g. 1 year, 10 years, 15 years). 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 couples sampled at two different sites. Since the streamflow data presents autocorrelation, the autocorrelation effects the KS test. Thus, we need to be more strict on the rejection of the null hypothesis. We can assume that the autocorrelation effect attenuates after three days. However, if the samples were 3 times smaller and for instance their 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. ¶
Results show that streamflow data gauged in different period of time (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 FDC sampled at another location. Figure 4 shows how different can be FDCs built at the sa ...

Deleted: ¶

Deleted: 3.2 The Antecedent Precipitation Index ¶
As resulted from the KS test, FDCs cannot be considered an invariant characteristic of a basin. The fact that FDCs are not invariant ...

Deleted: rainfall data series are characterized by a high amount of zeros which would affect the methodology presented here.

Deleted:

Deleted: Therefore, we need a weather descriptor which is not affected by ties. The methodology proposed in this paper involves the use of the Antecedent Precipitation Index (The API). This index ...

Moved up [1]: ¶

Deleted: specific

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, API at the *i-th* day is given by:

$$\text{API}_\text{i} = \sum_{j=0}^{29} \alpha^{j} P_{i-j}, \qquad (6)$$

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, API represents the long memory of the basin as it includes the effect of precipitation occurred many days before. To capture this behavior, in this study α is chosen equal to 0.85, this is in agreement with a previous study by Sugimoto (2014) who investigated one of the two case study areas (i.e. Neckar catchment); nevertheless this value was found to be suitable also for the US catchments. Here the API is calculated from areal precipitation instead of point precipitation.

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

3.2 How to determine the FDC at a partially gauged basin

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

Deleted: The API is used to investigate precipitation data in a similar way to discharge data .

Deleted: .¶

Deleted: 3

Deleted: behaviour

Deleted: good

Deleted: following

Deleted:

Deleted: and

Deleted: .

Deleted: 1

Deleted: 3

Deleted: Since it is found that two FDCs built for the same period of time at two different sites cannot be regarded as the same continuous distribution, we present here a methodology to overcome this issue and to estimate the FDC at an ungauged site. The FDC at one site is determined using climatological data gauged at another site. The FDC is expressed in millimeter, thus the area of the catchment is not an issue while retrieving the FDC from data of another catchment. based ona procedure $\widehat{F_Z}(z) = \widehat{F_Q}(q)\widehat{F_Z}(z)$and $\widehat{F_Q}(q)$ proxy variable at the donor site at the target site¶

Deleted: The methodology is carried out under the assumption of large scale precipitation.¶
Let consider two catchments, A and B. We want to determine the Flow Duration Curve at catchment B for each time period from data available at catchment A. Therefore, A will be named referencedonor catchment, while B will be called target catchment. ¶
Let suppose that in a given number of years, discharge is available at both sites A and B, these will be named reference years. On the other hand, in another number of years catchment B is characterized by a lack of streamflow data, while data are available for A. Specifically, t The procedure involves daily Antecedent Precipitation Index (API) of A and the daily streamflow values gauged of at both A and B during the reference years. The API is considered as a support variable to obtain the FDC at the ungauged site from the FDC of the gauged one.

Moved (insertion) [3]

Formatted: Font: (Default) Times New Roman, 12 pt

Formatted: Font: (Default) Times New Roman, 12 pt

Formatted: List Paragraph, Numbered + Level: 1 + Numbering Style: I, II, III, … + Start at: 1 + Alignment: Right + Aligned at: 0.63 cm + Indent at: 1.27 cm

Formatted: List Paragraph, Space Before: 0 pt, Numbered + Level: 1 + Numbering Style: I, II, III, … + Start at: 1 + Alignment: Right + Aligned at: 0.63 cm + Indent at: 1.27

Formatted: Font: (Default) Times New Roman, 12 pt

Formatted: List Paragraph, Numbered + Level: 1 + Numbering Style: I, II, III, … + Start at: 1 + Alignment: Right + Aligned at: 0.63 cm + Indent at: 1.27 cm

Formatted: Font: (Default) Times New Roman, 12 pt

Formatted: Font:

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, A and B. We want to determine the Flow Duration Curve at catchment B from data available at A. Therefore, A is the donor catchment, while B is the target catchment. 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 A and B, respectively. These will be named donor years (e.g. 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

Deleted: The supporting variable may be either the API or the streamflow values gauged at site A. ¶

Moved up [3]: Let suppose that in a given number of years, catchment B is characterized by a lack of streamflow data, while data are available for A.

Deleted: Let suppose that in a given number of years, catchment B is characterized by a lack of streamflow data, while data are available for A. ¶

Deleted: The proceduremethod to predict estimate streamflow values in at catchment B is explained step by step in the followings using the API as a supporting variable. Similarly it is possible to use the streamflow values recorded at the reference site as supporting variable.

Deleted: *Reference*

Formatted: Space Before: 12 pt

Deleted: 1.

Deleted: discharge and

Deleted: both

Deleted: s

Moved down [2]: (e.g. duration of 1 year, 10, 15, 20 years)

Deleted: reference

Deleted: reference

Moved (insertion) [2]

Deleted: reference

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.

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

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

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

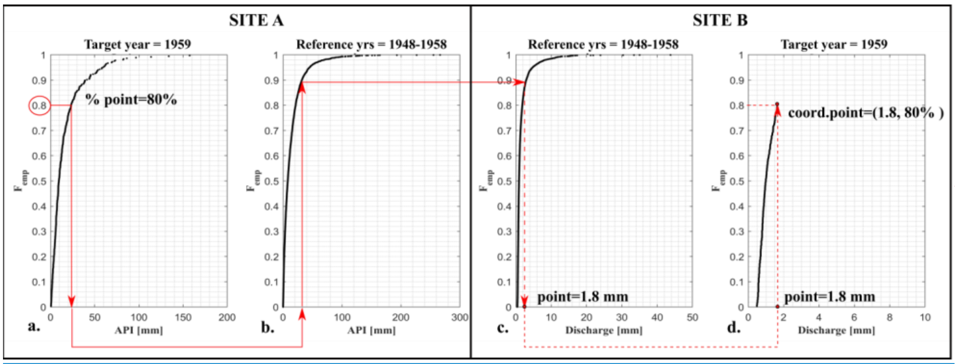


Figure 5. Illustration of FDC generation using the interpolation with the API of the donor site as a proxy.

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

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

3.3 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:

Formatted: Font: Not Italic

Deleted: reference

Deleted: one

Formatted: Font: Italic

Formatted: Font: Italic

Deleted: reference

Deleted: reference

Formatted: Font: Italic

Deleted: a

Formatted: Centered

Deleted: 2. From the precipitation values, estimate the API. Sort API and discharge values of each basin in descending order and assign to each sorted value the corresponding rank i, with i = 1,…,Nr where Nr is the length of each reference data series. ¶
3. For both basins, from the rank values, determine the exceedance probability of each API value as Nr_{i+1}. In the same way, for both catchments, find the exceedance probability of each discharge value. ¶
4. Select a period of time (e.g. 1 year, 10, 15, 20 years) in which streamflow values are missing at catchment B, while they are available at A. This period of time is then called target period and has length equal to Nt. ¶
5. For basin A, sort the API data series of the target period in descending order. Assign to each sorted value the corresponding rank and evaluate the associate exceedance probability. ¶
6. In order to determine the streamflow associated to a given probability of exceedance Pi (with i = 1,…,Nt) occurred in the target period at site B, take the API value with probability equal to Pi occurred at site A during the target year. This will be named API_{Ati}. ¶
7. From the sorted API data series occurred during the reference period at site A, find the API value equal to API_{Ati}. This API value has a probability of exceedance equal to Pj and is then called AP IArj (with j = 1,…,Nr). ¶
8. From the sorted streamflow data series gauged at site B during the reference period, find the discharge value qBrj having exceedance probability equal to Pj in the reference period. qBrj is the value of discharge occurred during the target year at site B with a probability of exceedance equal to Pi. ¶
If in the reference period at site A, it does not exist a value of API equal to API_{Ati}, then, from the sorted API data series occurred during the reference period at site A, take the ranks of the two most similar values to API_{Ati}, one should be bigger and the other smaller than API_{Ati} (i.e., API_{Arj} and API_{Arj+1} having probability of exceedance Pj and Pj+1, respectively). In the sorted streamflow data series gauged at site B during the reference period, look for the two discharge values of B having exceedance probability equal to Pj and Pj+1, respectively. Evaluate the mean value of these discharge values to obtain the discharge occurred during the target year at site B having probability of exceedance equal to Pi. The procedure is repeated for all exceeding probabilities Pi with i = 1,…,Nt. Therefore, the streamflow value at site B is¶
determined at each duration for the chosen target period. From the streamflow values obtained for the specific target period, the corresponding FDC is plotted. This will be named simulated FDC. The simulated streamflow values at each percentile (i.e., streamflow value associated to a specific duration) are compared with the observed values, that were actually available at site B.¶
Similarly, the procedure is

Deleted: In this paper, the procedure is carried out using either API values or discharge values as support variable.

Deleted: For the German and U.S. catchments, results are shown using discharge and API as support variable, respectively.

Deleted: used for the U.S. catchments

Deleted: 2

Deleted: (1970)

Deleted: simulated

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

where Q_{obs} is the observed discharge value at the target catchment during the target period; $\bar{Q}$ is the mean value of the observed discharge during the target period in the target catchment; 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 percentile.

The BIAS represents the mean difference between observed and interpolated values:

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

If the BIAS equals zero there is a perfect fit between 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:

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

Discharge values are in mm and so 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.

4. Results

The procedure explained above was tested on several target catchments 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, Figure 6. 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 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. On each panel, the two-sample Kolmogorov-Smirnov test distance between observed and interpolated value, 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 same period, the FDC of the donor catchment recorded during the target period was also plotted. It is noteworthy to

Deleted: $NSE = 1 - \frac{\sum_{i=1}^{N}\left(Q_{obs,i}-Q_{sim,i}\right)^2}{\sum_{i=1}^{N}\left(Q_{obs,i}-\bar{Q}\right)^2}$,

Formatted: Font: Times New Roman, Not Italic

Deleted: sim

Deleted: simulated

Deleted: Discharge values are in mm.

Deleted: simulated

Deleted: ¶
$BIAS = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{Q_{sim,i}-Q_{obs,i}}{Q_{obs,i}}\right)$.

Formatted: Swedish (Sweden)

Formatted: Swedish (Sweden)

Formatted: Swedish (Sweden)

Formatted: Swedish (Sweden)

Formatted: Swedish (Sweden)

Formatted: Swedish (Sweden)

Formatted: Swedish (Sweden)

Formatted: Swedish (Sweden)

Formatted: Swedish (Sweden)

Deleted: ¶

Deleted: $MAE = \frac{\sum_{i=1}^{N}\left|Q_{obs,i}-Q_{sim,i}\right|}{N}$.

Deleted: ¶

Deleted: simulated

Deleted: ¶

Deleted: TangipahoaChoctawthachee0.99610.99710.9930.9980.9950.998The procedure explained above has been tested on several target catchments varying both referencedonor and target time periods. For the German catchments, three different pairs of refere ...

Deleted: simulated

Deleted: , especially when a larger data set is used as both target ...

Deleted: ¶

Deleted: T

Deleted: simulated

Deleted: reference

Deleted: show a good agreement with observed FDCs

Deleted: 5

Deleted: It is interesting to note that observed and simulated FD ...

Deleted: agreement

Deleted: in some cases

Deleted: smaller

Deleted: 5

Deleted: as it can be observed at both Tangipahoa and Bogue si ...

Deleted: and

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.

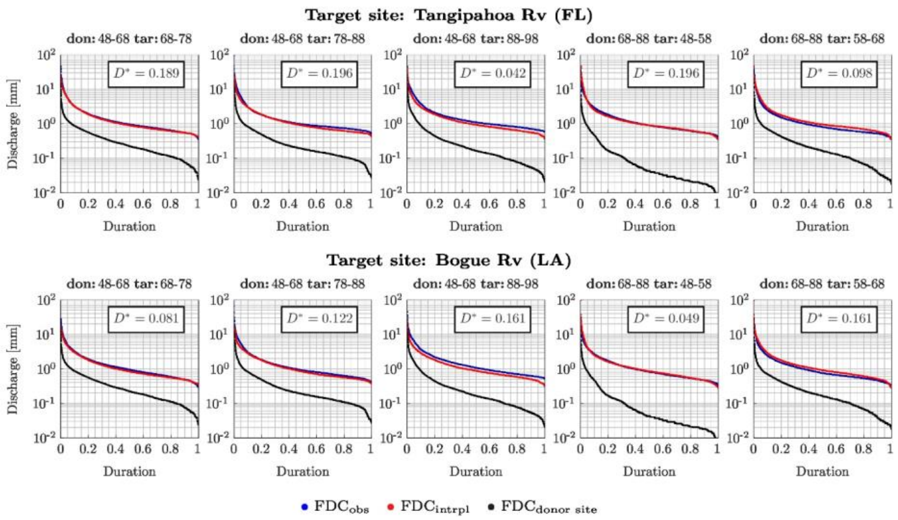


Figure 6. Interpolated FDC at Tangipahoa River (FL) and Bogue River (LA), upper and lower panel, respectively. The donor catchment is Blanco River (TX). The donor years are a 20 years time window from October 1948 to September 1968 and from October 1968 to September 1988. Target years are the decades shown above each panel. Blue dots and red dots are the observed and interpolated FDC at the target catchment, respectively; the black dots are the observed FDC at the donor catchment during the target period.

Interpolated and observed FDCs almost perfectly match when obtained using long donor and target periods, Figures 6 and 7. 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 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). 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 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.

Deleted: this implies

Deleted: 5

Deleted: Simulated

Deleted: reference

Deleted: reference

Deleted: (left panels)

Deleted: (right panel)

Deleted: the decades

Deleted: from October 1968 to September 1998

Deleted: reference

Deleted: donor

Deleted: ,

Deleted: and 8

Deleted: (i.e.

Deleted:)

Deleted: the KS distance is much higher when the target period is 1 year, Figure 67.

Deleted: , Figure 67

Deleted: there is a

Deleted: shift

Deleted: ; the difference

Deleted: that

Deleted: reference

Deleted: On the other hand, it is note worthy to observe that when larger time periods are considered both as target and as referencedonor, the simulated and observed FDCs almost perfectly match, Figure 78, for 15 years and Figure 8 9 for decades. ¶

Deleted: ¶

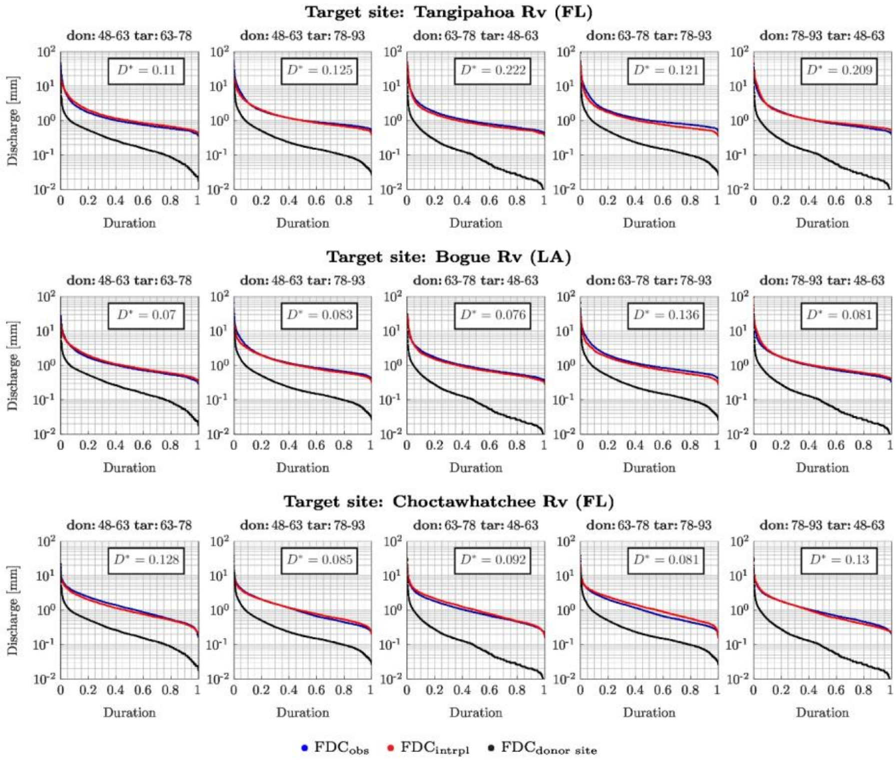


Figure 7. Interpolated FDC at Tangipahoa River (FL), Bogue River (LA) and Choctawhatchee River (FL), upper, middle and lower panels, respectively. The donor catchment is Blanco River (TX). The donor and target years are periods of 15 years. The blue and red dots are observed and interpolated FDC, respectively, at the target catchment, target period; the dots are FDC at the target catchment and the black dots are the observed FDC at the donor catchment during the target period.

Deleted: 7

Deleted: Simulated

Deleted: reference

Deleted: reference

Deleted: simulated

Deleted: reference

Deleted: donor

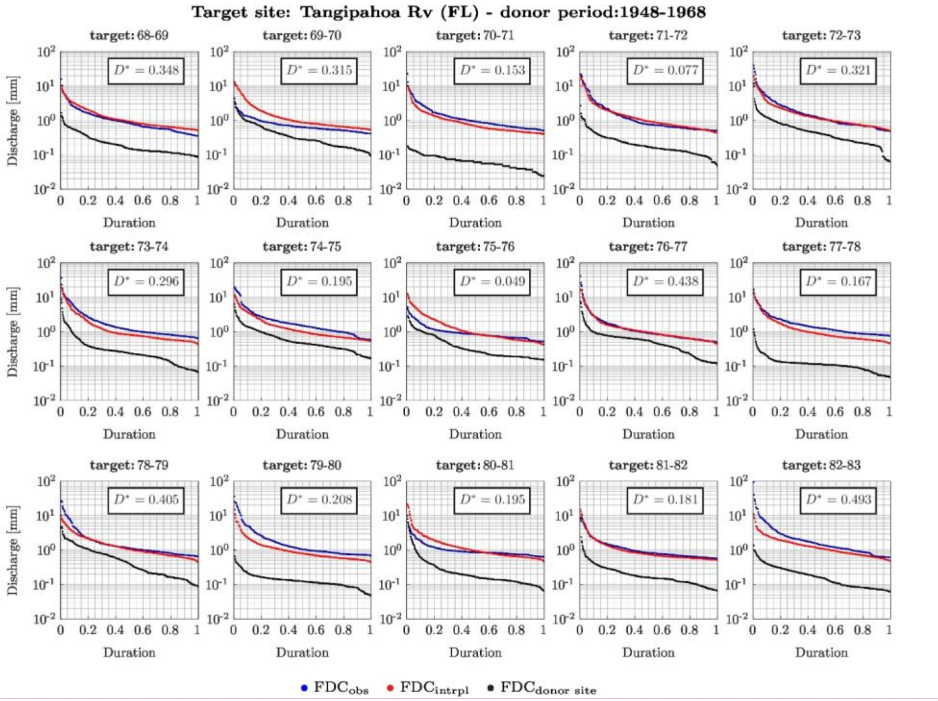


Figure 8. Interpolated FDC at Tangipahoa River (FL).The donor catchment is Blanco River (TX). The donor years are a 20 years time window from October 1948 to September 1968. Target years are each hydrological year from October 1968 to September 1982. The blue and red dots are observed and interpolated FDC, respectively, at the target catchment, target period; the dots are FDC at the target catchment and the black dots are the observed FDC at the donor catchment during the target period.

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, 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. 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 observed value. Its magnitude is likely higher for high flows, while it attenuates for intermediate percentiles. Also the MAE 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.

Deleted: ¶
4.1 Performance criteria¶
To determine the performance of the procedure proposed in this paper, different criteria are selected: the Nash-Sutcliff efficiency index, the BIAS and the mean absolute error (MAE).¶
¶

Deleted: 6

Deleted: 9

Deleted: Simulated

Deleted: reference

Deleted: reference

Deleted: simulated

Deleted: reference

Deleted: donor

Deleted: Moreover, tThese results show that the FDC is not a characteristic of the catchment properties only, but of both the catchment and the weather. To better investigate these findings, w

Formatted: Font: (Default) Times New Roman, 12 pt

Deleted: It measures the overall agreement between observed and simulated values. It is a non-negative metric without upper or lower bounds. A perfect model would result in a MAE equals to zero. Th

Deleted: For a deep understanding of

Deleted: estimation metrics presented in Sect. 3.2

Deleted: reference period

Deleted: simulated

Deleted: 112

Deleted: for

Deleted: the BIAS and

Deleted: (i.e., low percentiles)

Deleted: This is in agreement with what was already observed in Figure 89.

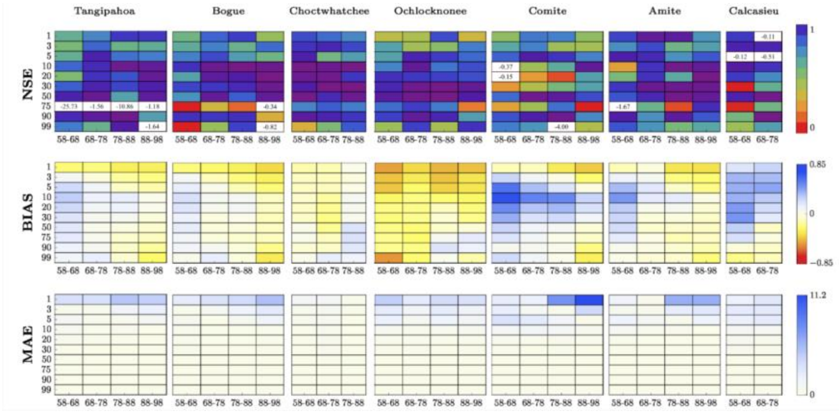


Figure 9. Performance measures NSE, BIAS and MAE evaluated for specific percentiles (on the y-axis) and for specific target decades on the x-axis. The donor decade is 1948-1958, the donor catchment is Blanco (TX). Each target catchment is indicated above the corresponding box. Negative values of the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

Deleted: 112

Deleted: (i.e., 1958-1968, 1968-1978, 1978-1988 and 1988-1998

Deleted:)

Deleted: reference

Deleted: reference

When both target and donor periods equal 15 years, the agreement between interpolated and observed flow values is high, Figure 10. 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, while intermediate and flows are more likely underestimated as shown by the BIAS.

Deleted: The same behaviourbehavior can be observed when both target and reference periods equal 15 years, Figure 123.¶

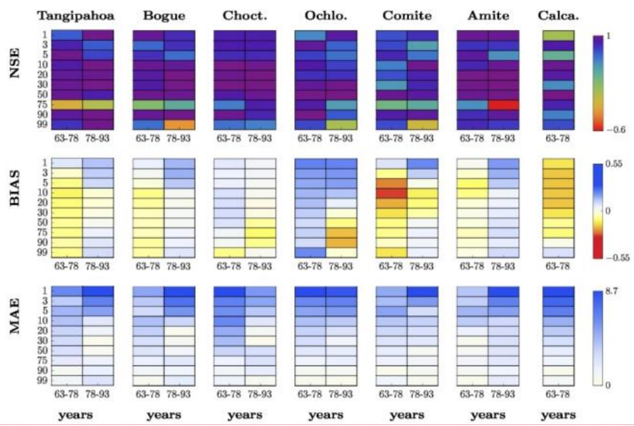


Figure 10. Performance measures NSE, BIAS and MAE evaluated for specific percentiles (on the y-axis) and for a specific 15 target years (i.e., 1963-1978 and 1978-1993 on the x-axis). The donor decade is 1948-1963, the donor catchment is Blanco (TX). Each target catchment is indicated above the corresponding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

Deleted: 2

Deleted: 3

Deleted: years

Deleted: reference

Deleted: reference

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.

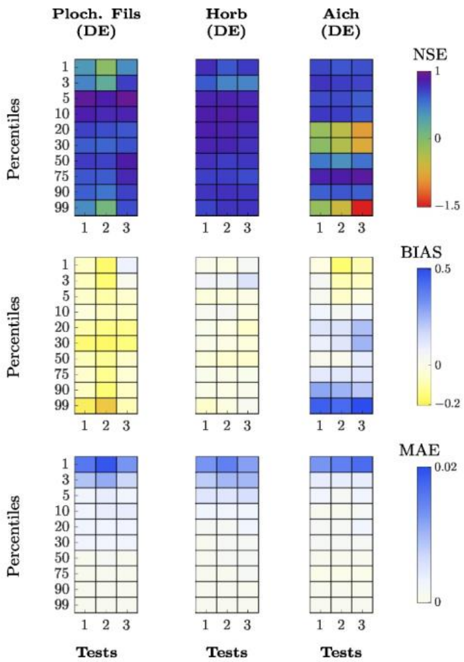


Figure 11. Performance measures NSE, BIAS and Mean Absolute Error (MAE) evaluated for specific percentiles (on the y-axis) and for three specific set of target and donor years (i.e., Test 1, 2 and 3). For Test 1 the target period is from 1961 to 1980 and the donor period is from 1981 to 1990. For Test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

Deleted: Wannweil

Deleted: reference

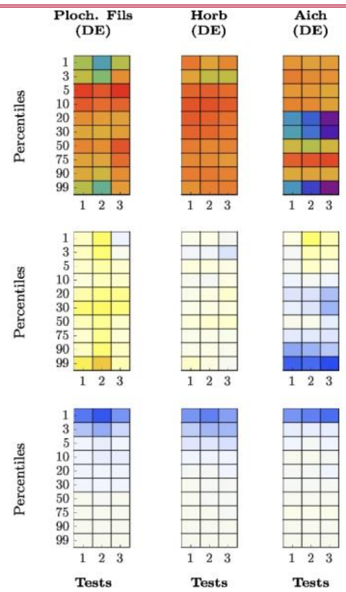
Deleted: ,

Deleted: ¶

Deleted: 3

Deleted: 4

Deleted:



Formatted: English (United States)

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.

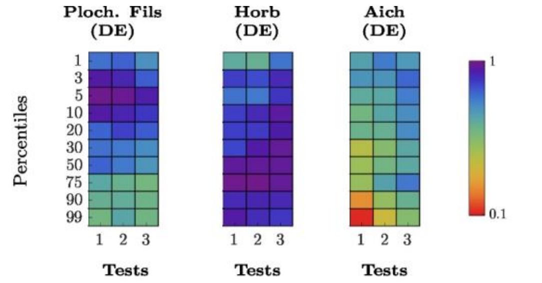


Figure 12. Correlation coefficient evaluated for each Test case at each percentile between Plochingen Neckar and each other catchment indicated on the boxes.

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 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. 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. 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 have the same distribution.

Deleted: for Wannweil and Horb,

Deleted: period of time

Deleted: must

Deleted: be

Deleted: T

Formatted: Font: (Default) Times New Roman, 12 pt

Deleted: ,in turn,

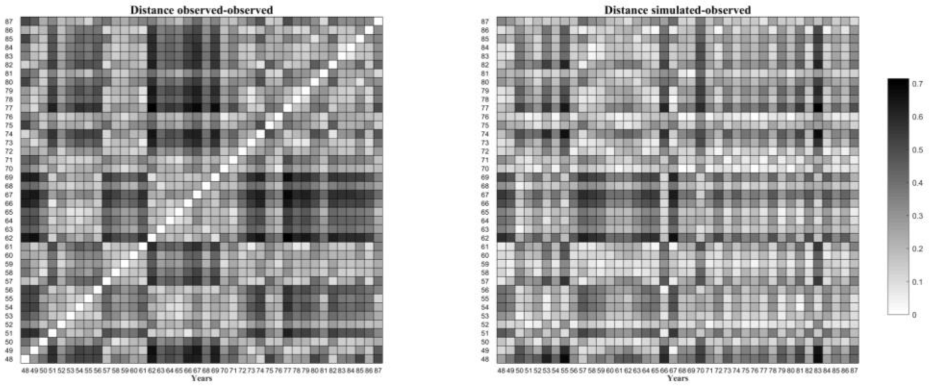


Figure 13. Kolmogorov-Smirnov distance between couples of streamflow values observed (left panel) and between couples of streamflow values observed and interpolated (right panel) at Tangipahoa River (FL) from October 1948 to September 1987.

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.

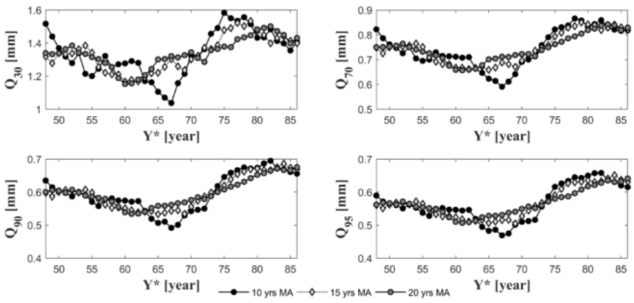


Figure 14. Moving average (MA) of the 30th, 70th, 90th and 95th percentiles of daily streamflow values gauged at Tangipahoa. Three different fixed time windows are used to estimate the MA: 10, 15 and 20 years. On the x-axis the first year of each interval is plotted (Y*).

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 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 in a misrepresentation of the percentiles themselves. Therefore, we suggest to add a weather factor to take into account 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, given the discharge and the precipitation gauged at another catchment. We show that two FDCs built for the same catchment 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 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 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 of both the basin and the weather. Therefore, it is not possible to infer a FDC using parameters retrieved from the distribution of another FDC without considering the weather.

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 (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 comparison of observed and interpolated FDCs, it results 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. 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

Deleted: the NSE shows values close to one for intermediate percentiles, while the performance decreases for the extreme percentiles. While for Oberensingen Aich, the intermediate percentiles shows a poor performance. These performances are confirmed also by the Ratio index. However, the error values given by the ME and the VE are close to zero in the majority of the cases. The errors increase for the percentiles corresponding to floodings for Horb and Oberensingen Aich.¶
To better understand the relationship between a target and a referencedonor catchment, the coefficient of correlation has been computed. Coefficient values are reported for Test 1. Values are estimated between the referencedonor catchment Plochingen Neck ...

Deleted: 5

Deleted: behaviour

Deleted: ungauged and

Deleted: n

Deleted: , but

Moved down [5]: The FDC at a specific site is not a property of

Deleted: The same results when two FDCs of two different ...

Deleted: analysing

Deleted: 99th

Moved (insertion) [5]

Deleted: as it

Deleted: es

Deleted: The API is determined from the precipitation and allow ...

Deleted: streamflow data and thus, the

Deleted: are

Deleted: for

Deleted: reference

Deleted: time period

Deleted: reference

Deleted: ¶

Deleted: Moreover, we tested a method to retrieve the FDC of a ...

Deleted: FDCs simulated at the target site are compared with FD ...

Deleted: reference period

Deleted: s (i.e., for which streamflow data are available)

Deleted: ones (i.e., for which streamflow data are to be determined)

Deleted: analysed

Deleted: simulated

Deleted: To exploit the methodology goodness several estimatio ...

Deleted: simulated

Deleted: Moreover, it is interesting to observe that

Deleted: distance

Deleted: simulated

Deleted: are similar in shape, even when the two curves are not ...

Deleted: t

Deleted: wo

Deleted: analyse

Deleted: reference

Deleted: reference period

Deleted: reference period

Deleted: reference

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.

Deleted: estimated

Deleted: This could be useful especially if no data are actually available at the target site for the target period and it is not possible to estimate the goodness of the simulated FDC.¶
The catchments used for analysis here are characterized by a mild climate, thus the snow melt effect was not taken into account.¶
However, if the methodology would be applied to catchments characterized by snowfalls, the snow melt effect should be taken into account adding the snow melt to the API.¶
¶
¶
The catchments used for analysis here are characterized by a mild climate, thus the snow melt effect was not taken into account.¶
However, if the methodology would be applied to catchments characterized by snowfalls, the snow melt effect should be taken into account adding the snow melt to the API.¶

Appendix A

In this Appendix we want to provide an easy example to better understand the method that we applied for U.S. catchments. 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, while a “target period” is a period of time during which streamflow values are not available at the target catchment. As the rainfall is available at both sites for both periods, also API values are.

Let suppose that we want to know the discharge value at catchment 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 (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 (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 $P(Q < q_i)$ 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 (i.e., Blanco Rv, TX) and sort in descending order the API values, estimate the rank and the associated exceedance probability $P(Q < q_j)$ of each value as $j=(N_r + 1)$ where N_r equals 5475;

6. find the exceedance probability $P(Q < q_j)$ 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 (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 and the target period that we were looking for, Table A4.

Deleted: reference period
Deleted: reference
Deleted: available at both catchments for both periods of time
Deleted: reference period
Deleted: and
Deleted: 1948
Deleted:
Deleted: 1963
Deleted: reference
Deleted: 3
Deleted: reference
Deleted: reference period
Deleted: reference
Deleted: q_i
Deleted: i
Deleted: q_i
Deleted: reference period
Deleted: q_i
Deleted: i

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 (i.e., Blanco RV, TX).

Deleted: reference

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 (i.e., Blanco RV, TX).

Rank	$P(API<API_i)$ %	$API_{Blanco,\ ref}$ 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

Deleted: reference

Deleted: reference

Deleted: API_i

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

Deleted: reference

Rank	$P(Q<q_i)$ %	$q_{Bogue,ref}$ 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

Deleted: q_i

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

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

Competing interests. No competing interests are present.

References

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

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.

Field Code Changed

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.

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.

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.

Field Code Changed

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.

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.

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

Field Code Changed