Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-347-RC1, 2018

Hydrology and Earth System Sciences
Discussions
Open Access

Interactive comment on "A methodology to estimate flow duration curves at partially ungauged basins" *by* Elena Ridolfi et al.

W.H. Farmer (Referee)

wfarmer@usgs.gov

Received and published: 20 July 2018

The authors have provided an interesting hypothesis that a daily flow duration curve, as function of both basin characteristics and climate, can be reproduced with the use of a precipitation index. While this is an intriguing assertion, the experiment and manuscript need some revision before being published. A previous commenter has noted some concerns, and I will try to expand on those concerns here with an aim towards supporting revisions of this work.

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

Printer-friendly version

Discussion paper

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.

The methods section needs a great amount of revision. A figure may improve the understanding of the methodology. 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.

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.

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.

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.

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.

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?)

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).

Before moving to some more specific comments, I would like to direct the authors to a couple previous works that should probably be discussed. 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.

SOME MORE SPECIFIC COMMENTS:

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.

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

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

Printer-friendly version

Discussion paper

than “exceedance time”.

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

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

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.

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

Page 5, line 7: misspelling of database

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

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

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.

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.

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

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

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

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).

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.

Page 12, line 9: Spelling of FDCs.

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

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

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

CLOSING REMARKS:

I want to express my deepest appreciation for this work and convey my strongest encouragement. I think the topic is quite relevant and the hypotheses are very intriguing. With some additional redevelopment, I think this work could address some fundamental questions of hydrology. It is for this reason that I invested so much time in the review. While some of my conclusions might be inaccurate, I hope I have provided some food for thought as to improving this work. If you have any concerns, please do not hesitate to reach out to me again. This is a topic I'd be more than happy to discuss at length. Thank you.

William Farmer

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-347, 2018.