Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-347-RC2, 2018

Hydrology and Earth System Sciences
Open Access
Discussions

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

T. Over (Referee)

tmover@usgs.gov

Received and published: 4 September 2018

Overview:

This paper presents a method for estimating FDCs during an ungauged period at a "target" location that is gauged during another period. It is based on the antecedent precipitation index (API) at another ("reference") basin, where a correspondence is set up between the FDC at the target gauge during its gauged period and the FDC of the API at the target basin and between the FDCs of the API at the reference basin during the gauged and ungauged periods. 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 (1996, Hydrological Sciences Journal, vol. 41, pp. 851-871) who presented the use of time series of discharges at target and reference basins and their FDCs for daily record extension. Combining those two contributions yields the use of CPI or API for record extension, as was used by Straub and Over (2010) in an admittedly obscure report at: https://apps.ict.illinois.edu/projects/getfile.asp?id=3033, appendix A, section 3.2, and possibly elsewhere.

Arguably, while it seems the methodology in this paper could be considered a corollary of the prior work cited above, the focus of the present on FDCs rather than daily streamflow may distinguish it. 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.

Apart from that overlap with previous work, I have the following major concerns regarding this paper in its current form: 1. As mentioned, in the methodology presented, the API at another gauged basin, the "reference" basin, is used for the FDC extension at the target basin during its ungauged period. This is strange, for two reasons: (1) Why select another gauged basin and use only its API, not its streamflow, and (2) Why not use the API at the target basin? I am in fact rather surprised it seems to work as well as it does when the API used is not at the target basin. Perhaps this points to a difference between the application for daily streamflow as compared to FDCs (i.e., that the FDC application is comparatively forgiving).

2. A list of detailed comments is given below, but in general the presentation is rather uneven; some of the primary issues are:

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.

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

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

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.

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.

Specific technical comments:

1. Section 2.1:

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

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

2. Section 2.3

a. Figure 3:

i. Perhaps plot and check correlation with temperature of ET/P instead of just ET?

ii. 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.

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.

3. Section 3.1

a. 2nd paragraph:

i. It seems there should already be a well-established way of addressing autocorrelation effects on the K-S test.

ii. 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?

b. 3rd paragraph: This paragraph seems to include “Results”, not “Methodology”.

4. Section 3.2:

a. 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?

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

5. Section 3.3:

a. First paragraph, last sentence: Why do you assume “large scale precipitation”? What do you mean by that?

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.

6. 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.

7. 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.

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.

Comments on the presentation

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

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.

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…".

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.

5. Section 4, figures 5-8:

a. 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?

b. 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.

6. Section 4, figure 9 and discussion of it:

a. 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.

ii. 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.

iii. 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.

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.

8. Section 4.1, pp. 14-17, Definition of performance criteria:

a. Are all these computed for Q in mm units? Even though those are units used throughout, it would be worth re-emphasizing that here.

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?

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.

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?

11. Conclusions:

a. 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.

b. 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.)

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.

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