Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-347-SC5, 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.

F. Serinaldi

francesco.serinaldi@ncl.ac.uk

Received and published: 21 July 2018

Dear Andras,

Thanks for your reply. Now my question is: why things you explained so clearly in few lines in your reply do not seem even mentioned in the paper? According to your response, the key point is that API shows a distribution close to stream flow (as it reflects the same generating mechanism). However, I cannot see this in the very short section 3.2, and figures, equations, etc. seem not to show this key fact, at least explicitly. On the other hand, there is a long discussion in section 3.3 and an appendix describing what is nothing but a quantile mapping, which should not be the main point of the paper, if I correctly interpret your response.

Printer-friendly version

Discussion paper

You are right when you say that my statements about independence are incorrect; in fact, in the second post that I sent, I showed diagrams (in the first figure) where I show the effect of correlation on the reproduction of 'FDCs'.

In my opinion, here the point is that the proxy variable you chose (whatever it is, apart some generic requirements) should closely follow the fluctuations of the target variable in order the quantile mapping to work, which is also what implies your first equation $Q_k(t) = A(t)$, or a less demanding $Q_k(t) \propto A(t)$. In this respect, I cannot see how your sentence 'The two series are not perfectly or weakly correlated, but their distributions are practically the same.' applies in this context.

In fact, what matters in a quantile mapping is, I think, that the variables co-move in order to guarantee quantile matching (Figure 2 of Smakthin and Masse (2000) provides a nice illustration), and this seems to me obvious. Quantile mapping in the appendix does not need that the distributions of Q and A are the same; instead, they can be arbitrary, apart from minimal general requirements (in the second figure of my previous comment I used normal and lognormal CDFs for X and Y, respectively). Quantile mapping works because low/middle/high quantiles of X correspond to (and tend to occur in your time sequence jointly with) low/middle/high quantiles of Y, irrespective of the shape of their marginal distributions (with usual general assumptions on continuity, etc.). Ideally, the matching is perfect if the variables are comonotonic: the weaker the co-movement (correlation) the worse the effectiveness of quantile mapping (figure below attempts to illustrate what I mean). If Q and A are not well correlated, even if the marginals of Q and A were identical, there would be no way to know which quantiles of Q occurred in a target period from the observed quantiles of A in the same period, because the latter can be e.g. systematically low/high for some reason, while the former can be whatever, and vice versa. In this respect, I think that the quality of your results depends on how closely A co-moves (is correlated) with Q, and nothing else. So, I repeat: every variable well correlated with Q does the job (of course, A can be one of them). P_k does not because the correlation with Q is poor, and P_k is discrete-continuous while Q is generally continuous (...leaving aside intermittent streams, etc., you know). Thus, in my comment, it was implicit that the proxy should share minimal properties with Q in order to closely follow flow fluctuations.

Thus, if the message is that a weighted moving average of antecedent rainfall (API) closely co-moves with streamflow and thus A is a very good proxy enabling reasonably reliable quantile matching, that's fine. However, in this case I think that the paper should be reworded to focus on this, by preliminary showing the evidence of this assumption (as also suggested by Dr. Farmer). In the present form, as you can see by yourself, most of the test is devoted to (i) description of KS tests adjusted in a rather rough manner and applied in a context where every underlying assumption is not fulfilled, (ii) quantile mapping, which is not even performed in a consistent way, and in turn can be summarized by a single equation, $q = \hat{F}_Q^{-1}(\hat{F}_A(a))$, strongly shortening and simplifying (if not removing) section 3.3, and (iii) a long list of well-known performance indices whose selection looks a bit random, at least to me. On the other hand, the only equation of interest and few explanatory sentences reported in your reply are not shown... I may have missed them.

To summarize, if you think that the co-movement of Q and A is new and important (along with the message that FDCs depends on both meteorological forcing and basin response... is this not known?... Anyway), and represents something to be communicated, that's fine. As you know, as I'm statistically biased, for me A is just a good co-variate/proxy, even though I recognize that it is chosen looking at the generating mechanism, and can be of practical interest for HESS audience. However, this should be stated much better in the text. Nonetheless, please forgive me if I cannot see what is interesting and new in the roughly performed and badly described quantile mapping, KS testing for ni/nid variables, apparently random selected measures of performance, and if I find the overall presentation very poor in terms of material organization, mathematical notation, clarity, and synthesis. I think that the empirical considerations/arguments (about FDC fluctuations, etc.) and statistical techniques can be compactly and clearly summarized in a few pages using simple and effective mathematical notation (as you are a mathematician, you know what I mean).

Honestly, I am perplexed. This is not your standard, and I am confident that you are conscious of this, even if you will legitimately defend this work. By my side, I am not a reviewer; so, my opinion does not matter very much. I am confident that you do not mind if expressed it with my usual rough tone.

As ever

F

Dear Andras,

Thanks for your reply. Now my question is: why things you explained so clearly in few lines in your reply do not seem even mentioned in the paper? According to your response, the key point is that API shows a distribution close to stream flow (as it reflects the same generating mechanism). However, I cannot see this in the very short section 3.2, and figures, equations, etc. seem not to show this key fact, at least explicitly. On the other hand, there is a long discussion in section 3.3 and an appendix describing what is nothing but a quantile mapping, which should not be the main point of the paper, if I correctly interpret your response.

You are right when you say that my statements about independence are incorrect; in fact, in the second post that I sent, I showed diagrams (in the first figure) where I show the effect of correlation on the reproduction of 'FDCs'.

In my opinion, here the point is that the proxy variable you chose (whatever it is, apart some generic requirements) should closely follow the fluctuations of the target variable in order the quantile mapping to work, which is also what implies your first equation $Q_k(t) = A(t)$, or a less demanding $Q_k(t) \propto A(t)$. In this respect, I cannot see how your sentence 'The two series are not perfectly or weakly correlated, but their distributions are practically the same.' applies in this context.

In fact, what matters in a quantile mapping is, I think, that the variables co-move in order to guarantee quantile matching (Figure 2 of Smakthin and Masse (2000) provides a nice illustration), and this seems to me obvious. Quantile mapping in the appendix does not need that the distributions of Q and A are the same; instead, they can be arbitrary, apart from minimal general requirements (in the second figure of my previous comment I used normal and lognormal CDFs for X and Y, respectively). Quantile mapping works because low/middle/high quantiles of X correspond to (and tend to occur in your time sequence jointly with) low/middle/high quantiles of Y, irrespective of the shape of their marginal distributions (with usual general assumptions on continuity, etc.). Ideally, the matching is perfect if the variables are comonotonic: the weaker the co-movement (correlation) the worse the effectiveness of quantile mapping (figure below attempts to illustrate what I mean). If Q and A are not well correlated, even if the marginals of Q and A were identical, there would be no way to know which quantiles of Q occurred in a target period from the observed quantiles of A in the same period, because the latter can be e.g. systematically low/high for some reason, while the former can be whatever, and vice versa. In this respect, I think that the quality of your results depends on how closely A co-moves (is correlated) with Q, and nothing else. So, I repeat: every variable well correlated with Q does the job (of course, A can be one of them). P_k does not because the correlation with Q is poor, and P_k is discrete-continuous while Q is generally continuous (...leaving aside intermittent streams, etc., you know). Thus, in my comment, it was implicit that the proxy should share minimal properties with Q in order to closely follow flow fluctuations.

Thus, if the message is that a weighted moving average of antecedent rainfall (API) closely co-moves with streamflow and thus A is a very good proxy enabling reasonably reliable quantile matching, that's fine. However, in this case I think that the paper should be reworded to focus on this, by preliminary showing the evidence of this assumption (as also suggested by Dr. Farmer). In the present form, as you can see by yourself, most of the test is devoted to (i) description of KS tests adjusted in a rather rough manner and applied in a context where every underlying assumption is not fulfilled,

(ii) quantile mapping, which is not even performed in a consistent way, and in turn can be summarized by a single equation, $q = \hat{F}_Q^{-1}(\hat{F}_A(a))$, strongly shortening and simplifying (if not removing) section 3.3, and (iii) a long list of well-known performance indices whose selection looks a bit random, at least to me. On the other hand, the only equation of interest and few explanatory sentences reported in your reply are not shown... I may have missed them.

To summarize, if you think that the co-movement of Q and A is new and important (along with the message that FDCs depends on both meteorological forcing and basin response... is this not known?... Anyway), and represents something to be communicated, that's fine. As you know, as I'm statistically biased, for me A is just a good co-variate/proxy, even though I recognize that it is chosen looking at the generating mechanism, and can be of practical interest for HESS audience. However, this should be stated much better in the text. Nonetheless, please forgive me if I cannot see what is interesting and new in the roughly performed and badly described quantile mapping, KS testing for ni/nid variables, apparently random selected measures of performance, and if I find the overall presentation very poor in terms of material organization, mathematical notation, clarity, and synthesis. I think that the empirical considerations/arguments (about FDC fluctuations, etc.) and statistical techniques can be compactly and clearly summarized in a few pages using simple and effective mathematical notation (as you are a mathematician, you know what I mean).

Honestly, I am perplexed. This is not your quality level, and I am confident that you are conscious of this, even if you will legitimately defend this work. By my side, I am not a reviewer; so, my opinion does not matter very much. I am confident that you do not mind if expressed it with my usual rough tone.

As ever

F

PS: API patterns look like simple and empirical counterparts of shot noise processes used to model stream flow data

F. Laio, A. Porporato, L. Ridolfi, and I. Rodriguez-Iturbe (2001) Mean first passage times of processes driven by white shot noise Phys. Rev. E 63, 036105

Claps, P., A. Giordano, and F. Laio (2005), Advances in shot noise modeling of daily streamflows, Adv. Water Resour., 28, 992–1000.

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

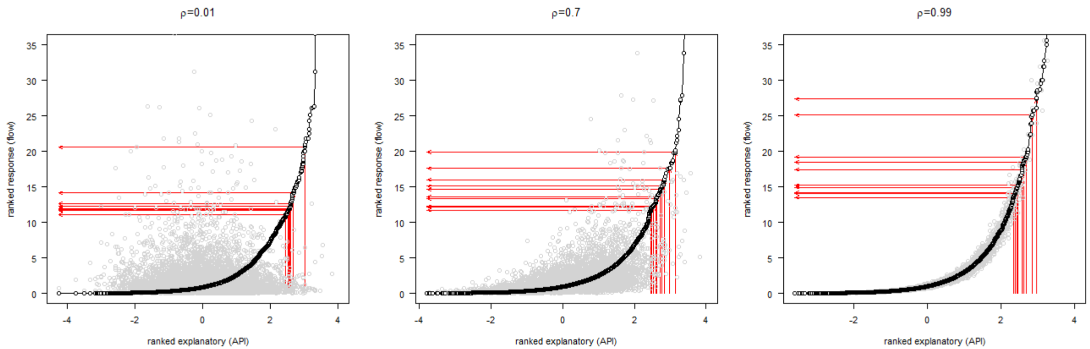

Fig. 1.