

## ***Interactive comment on “Bias correction of Simulated Historical Daily Streamflow at Ungauged Locations by Using Independently Estimated Flow-Duration Curves” by William H. Farmer et al.***

**William H. Farmer et al.**

wfarmer@usgs.gov

Received and published: 24 May 2018

**Reviewer Comment 1:** *The authors presented a bias-correction procedure useful for improving the accuracy of simulated daily streamflow series by using independently estimated flow-duration curves (FDCs). Although the procedure itself is not completely new (references of previous studies are included in the manuscript), the study is interesting as it considers an extended database in the US and focuses on the reproducibility of upper and lower tails, distinguishing between observation-dependent and observation-independent tails. This aspect is meaningful for highlighting the effect of*

C1

*timing on distributional bias. The study concludes that the significant potential of the bias-correction procedure is limited by the accuracy of the FDCs estimation method.*

*The paper is well structured and written, but some changes should be applied for making it more readable and for emphasizing some important aspects. I believe it is suitable for publication in HESS after the authors address some issues reported in the comments below.*

**Author Response 1:** Thank you for your deep consideration of our manuscript. Your kind comments will surely help us to improve the delivery of our results.

**Reviewer Comment 2:** *I would suggest to explain more in details the “Bias correction” procedure at Section 2.3. For instance, at Page 4 Lines 29-30, the sentence “[. . .] linearly interpolated along two types of independently estimated FDCs” is rather ambiguous: I found difficult to understand whether the authors refer to the resampling of the curves, or perhaps to the prediction of FDC quantiles, which is carried out with a linear regression. I would recommend to rephrase this sentence and add more information to it, in order to clarify better this fundamental aspect of the proposed procedure. Moreover, at least one figure could be useful for clarifying the procedure. I can suggest to show at least two plots, where the authors may report standard normal quantiles vs. logarithmically transformed streamflow percentiles for, in turn, (a) regionally regressed FDC and pooled ordinary kriging curve, and (b) observed FDC and pooled ordinary kriging curve. Finally, in my opinion, the bias-correction section should be extended: I recommend to add more detailed information about how the bias correction is applied to the simulated streamflows, maybe introducing a figure vignette, or, likewise, describing the procedure point by point.*

**Author Response 2:** In concert with the response of other reviewers, we will be re-formatting the section on methodology. It will be substantially enhanced by figures that show example hydrographs and example FDCs, as well as a figure demonstrating the steps of the procedure.

C2

**Reviewer Comment 3:** *I would stress more that in the majority of possible practical applications of the proposed method (i.e. predictions in ungauged sites), using observed FDCs for the bias correction would not be possible. Indeed, the only exception could be represented by those catchments in which we want to simulate streamflows for a given period, even though we have streamflow data for another period. I would add this reasonings to the revised version of the paper.*

**Author Response 3:** We will add clarification to point out that the use of observed FDCs is only provided as a theoretical demonstration of the upper-limit of performance in ungauged locations. We did not consider use of observed FDCs for record extension, but this approach might be useful in partially gauged locations. Though we did not explore that particular application, we will mention it in the discussion.

**Reviewer Comment 4:** *[3] Page 10, Lines 11-13 – The authors highlight that “initial exploration did not find a strong regional component to performance of the bias correction method”. In order to better support this sentence and to improve the effectiveness and completeness of the study, the authors could better discuss the spatial distribution of performance, especially given the high climatic variability among the conterminous United States. Therefore, I would suggest to add and discuss a new figure (or figures; e.g. a set of maps, similar to Figure 1), showing the spatial distribution of bias and root mean squared error in the study region for at least a couple of the cases considered in the study (Orig., BC-RR, BC-Obs.; upper tail, lower tail; observation-dependent tail, observation-independent tail; sequential, distributional, etc.).*

**Author Response 4:** We will develop such a figure.

**Reviewer Comment 5:** *In Figure 1, the differences between not-selected and selected streamgauges are not clear: in some areas, crosses overlap with points and the distinction is not simple. Maybe using different colors and symbol sizes might highlight better the differences between these two categories.*

C3

**Author Response 5:** We will work with the journal to improve the visibility of this figure.

**Reviewer Comment 6:** *Page 3, Line 24 – It is not clear what a 2-digit Hydrologic Unit is. Could you please explain?*

**Author Response 6:** The Hydrologic Unit system is a common method for delineating watersheds in the US. 2-digit hydrologic units (the polygons in Figure 1) roughly align with the major river basins of the United States. We will add this description with appropriate citations. In the figure, the units are the outlined polygons. (Seaber, Paul R., F. Paul Kapanos, and George L. Knapp (1987). "Hydrologic Unit Maps". United States Geological Survey Water-supply Papers. No. 2294: i-iii, 1-63.)

**Reviewer Comment 7:** *Page 3, Line 16 (and elsewhere) - “cubic feet per seconds (cfs)” is used. I believe that the International System of Units should be used in scientific papers. At the same time, if cfs is the standard adopted by USGS (and in the GAGES-II database), I think that at least the conversion factor to m<sup>3</sup> s<sup>-1</sup> should be indicated in parentheses the first time that cfs is mentioned.*

**Author Response 7:** We will add a conversion factor.

**Reviewer Comment 8:** *Page 4, Line 3 - It seems that you are referring to period-of-record FDCs; am I correct? Could you please state that explicitly?*

**Author Response 8:** Yes, you are correct. We will state this more clearly.

**Reviewer Comment 9:** *Page 4, Line 10 - It is not clear to me what “best-subsets regression” is. Could you please clarify and/or add at least one reference?*

**Author Response 9:** Best-subsets regression is a common regression technique that exhaustively searches the predictor space for the best model with a specified number of variables. The specified number of variables is then changed to explore a range

C4

of model sizes. As described by Farmer et al. (2014), of these models that then differ in size, the AIC (or some other metric) is used to select the "best" model in an unsupervised fashion.

**Reviewer Comment 10:** *Page 4, Lines 13-15 and 20-22; Page 8, Lines 19-21 - I would suggest to remove parentheses.*

**Author Response 10:** We will remove them.

**Reviewer Comment 11:** *Page 5, Lines 9-10 – I would recommend to add an equation showing the expression “ten to the power of the difference and subtracting one from this quantity”. I would also suggest to explain why you are referring to this equation for computing the percentage. Moreover, the authors could show equations for bias and root mean squared error, respectively, when introducing them.*

**Author Response 11:** We will add these equations and any relevant citations.

**Reviewer Comment 12:** *Page 5, Line 10 – The authors write “root-mean-squared error”, while use “root mean squared error” in the remainder of the text. I would suggest to use the same expression everywhere (“root mean squared error” should be fine).*

**Author Response 12:** We will make every effort to ensure consistency throughout the manuscript.

**Reviewer Comment 13:** *Page 5, Line 13 – Could you please add a reference for the Wilcoxon signed rank test?*

**Author Response 13:** We will add this citation: Wilcoxon, Frank (Dec 1945). "Individual comparisons by ranking methods". *Biometrics Bulletin*. 1 (6): 80–83. doi:10.2307/3001968.

C5

**Reviewer Comment 14:** *In Tables' captions, the acronyms “OD” and “OI” are used instead of “observation-dependent” and “observation-independent”, respectively. You could use these acronyms also in the body of the text, in order to improve the readability; e.g. “OD-tail” instead of “observation-dependent tail” and “OI-tail” instead of “observation-independent tail”.*

**Author Response 14:** A previous reviewer suggested also using BC-Obs, etc. We will attempt to use these abbreviations, and ensure that clarity is not harmed.

**Reviewer Comment 15:** *Page 10, Lines 6-7 – Could you please report some other methods (please provide references) for estimating FDCs?*

**Author Response 15:** Moving and expanding the last paragraph of the introduction will allow us to discuss other methods. Some might include TNDTK, kriging methodologies, index-flood methods, other hydrograph simulations, etc.

**Reviewer Comment 16:** *Page 11, Line 16 - Please delete “Summary and conclusions”.*

**Author Response 16:** Yes, we will fix this error.

**Reviewer Comment 17:** *Captions of Figures 2, 3, 4, 5 – Please remove comma in “pooled, ordinary kriging”.*

**Author Response 2:** Commas will be removed.

**Reviewer Comment 18:** *In the body of text, there are some references to a recent study by Pugliese et al. (2017), which is currently under review. I would suggest to update these references after its possible acceptance/publication.*

**Author Response 2:** We have been keeping an eye on this publication and will update the references when a decision on that manuscript is made.

C6