

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 work presents a method that aims to produce unbiased time series of daily discharges for ungauged basins. It consists in correcting the distributional bias in the time series of simulated discharges obtained for this ungauged basin from some simulation model (in the present case, the simulation model is a regression model where the daily discharge of each day is estimated via kriging from gauged stations in the neighborhood of the target ungauged station). The correction is a quantile –mapping correction where the reference statistical distribution is obtained*

C1

from the Flow-Duration-Curve of the target location, estimated via regionalization again from the observed FDC estimated for the neighboring stations.

The issue is a very relevant one and the approach a promising one. The novelty of the approach with respect to other similar studies (mentioned in the manuscript) has to be clarified.

Author Response 1: Thank you for your deep consideration of our manuscript. Your comments have helped us to identify weaknesses in presentation that substantially improve our presentation.

Reviewer Comment 2: *More specifically, the aim of the paper is not well explained and it has taken me more than 2 hours to understand what was done in the paper. One or possibly several graphical schemes are definitively required to figure out the methodology / objective of the paper and data used for the application of the method. A reformulation of the introduction and of the abstract is for me also necessary. Showing figures with hydrographs or FDC is also required (boxplots are not informative enough).*

Author Response 2: Based on your comments and those of another reviewer, we are proposing to reorganize the manuscript and present several additional graphics. One will be a two-panel figure where one panel overlays observed and simulated hydrographs, while the second panel shows observed and simulated flow duration curves. A second figure will show how the methodology proceeds from simulated hydrograph, through estimated FDC to bias-corrected hydrograph. We will also take another look at the clarity of the introduction and the abstract.

Reviewer Comment 3: *The formulation in the abstract for instance is somehow clumsy : “Based on an existing approach that separates the simulated streamflow into components of timing and magnitude, the timing component is converted into simulated nonexceedance probabilities and rescaled to new volumes using an independently estimated flow-duration curve (FDC).” I think that the authors do not rescale the timing*

C2

component. They just scale each daily discharge of the time series with a multiplicative correction factor that depends on the non exceedance probability of that discharge. . . The correction function is estimated from the FDC of the target location, estimated from the observed FDC of gauged stations in the region. I would have understood instantaneously the objective / approach of the work that way (I am perhaps wrong, but I guess this is roughly what you do). This has definitely to be clarified (here and everywhere else in the manuscript).

Author Response 3: The reviewer is correct, and we propose revising this sentence to “Based on an existing approach that separates the simulated streamflow into components of timing and magnitude, the timing component is converted into simulated nonexceedance probabilities and the volumes are rescaled using an independently estimated flow-duration curve (FDC) derived from regional information.” We will correct this impression throughout the manuscript.

Reviewer Comment 4: It is also not clear that the authors want to estimate unbiased time series of discharge at ungauged sites. This has to be clearly stated. One thus needs to have a regional simulation model and a regional model for the FDC estimation. The interest of the configuration where an observed FDC is used to bias correct the simulated time series has to be made clear in this context. This approach is presented as an alternative estimation approach. It is actually not and this is really confusing.

Author Response 4: The primary goal of this work is to estimate distributionally unbiased time series of discharge at ungauged sites. This statement will be reinforced. The use of observed FDCs, which, of course, are not available at ungauged locations are presented as an “idealized” case (P1, L10). It is presented merely as demonstration of the “theoretical potential of this [general] approach” (P6, L10), providing evidence to support the hypothesis that such an approach would work with a suitably-estimated FDC (P6, L30). We will strive to make this point more clearly.

C3

Reviewer Comment 5: The results of the study were rather expected. The authors find that the combined “time series simulation”+“FDC bias correction approach” is unable to correct the bias in the simulated hydrograph. The authors explain this is due to the bad quality of the regionalized FDC which present a rather significant distributional bias for all discharge percentiles. By construction, the “poor” result of the FDC bias correction was thus expected. The interest of the method would have been for a configuration where the distributional bias obtained from the “time series simulation” approach is larger than that of the regionalized FDC. Obviously, this is unfortunately here not the case. At least in average for the 1168 basins considered in the work. More convincing results could have perhaps been presented considering different configurations of basins with different relative performances of the two approaches considered here for the combination. We would have seen that the method works in configurations where regionalised FDC are good.

Author Response 5: It is surely expected that this two-step approach will depend on the accuracy of both steps. This project demonstrates the impact that FDC accuracy can have on the results. However, the results were not uniformly poor. As mentioned in the discussion, improvements were seen in certain parts of the distribution of daily streamflow. It is unfortunate that the results were not uniform improvements. In the revision we will identify a few sites that show improvements under the situation the reviewer describes (when bias obtained from the “time series simulation” approach is larger than that of the regionalized FDC). This will supplement our analysis of how performance of the estimated FDC corresponds to improvements in the time series.

Reviewer Comment 6: P2. “As attested to by many researchers focused on the reproduction of historical streamflow, this bias commonly appears as a general overestimation of low streamflow and underestimation of high streamflow”. This statement is not true. Simulations with any hydrological model will lead to the overestimation of some high stream flow and the underestimation of the others. The same for low flows : some are over-, some are underestimated. In rainfall-runoff models, it clearly depends

C4

on the amount of rainfall estimated from the small number of rain gauge available, and frequently not numerous enough to have the right estimate of rainfall input. It also depends on the quality of the model which is limited for the simulation of specific hydro processes. The fact that a given model systematically overestimate lowflows and underestimate highflows surely means that the model is not good enough and could likely be improved. As a matter of fact, some models present those limitations and this justify the present work. But a general statement can not be given there on such a case. (Note that this statement is invalidated in your work by the results and comments you mention (p6 – line 20/25). Please reformulate.

Author Response 6: The citations provided do note the tendency of the bias described. Furthermore, this statement is not made categorically. Instead, what is described is a *common* manifestation of bias, not a *categorical* or *general* statement of all bias. The results presented do not invalidate this statement, as the medians clearly demonstrate, but do show a range of performance. We will try to make this nuance more clear in the revision of this manuscript.

Reviewer Comment 7: P2. Ln 22. *The difference / novelty of the present methodology with that of Fennessey (1994) and Hughes and Smakhtin (1996) is not clear. Please clarify.*

Author Response 7: We will attempt to clarify this in the revision. Fennessey (1994) and Hughes and Smakhtin (1996) presented a method to simulate streamflows using a donor gauge for timing and a regionalized regression for FDCs. The approach presented in the current work is an extension where the timing is generated from a process-based model (rather than a donor).

Reviewer Comment 8: P3. *The description of the content of the paper is missing at the end of the introduction.*

Author Response 8: We will add a paragraph describing the headers of each section
C5

as a signpost to the reader.

Reviewer Comment 9: P3. *Simulated discharges. "An initial predictions of daily streamflow values for each stream gauge were obtained by applying the pooled ordinary kriging approach (Farmer, 2016) to each 2-digit Hydrologic Unit". The approach has to be better explained. It is not possible to understand how time series are obtained for any given gauge from what is said in the paragraph ("It builds a single, time-invariant semivariogram model of cross-correlation that is then used to estimate ungauged streamflow as a weighted summation of all contemporary observations"). A summary methodological scheme could be helpful. Is the variogram model estimated independently for each day ? for each target location from stations in its neighborhood (or do you use all stations of the United States to estimate the daily discharge of a given station) ? What is a 2-digit Hydrologic Unit ? Does the leave-one out procedure applies for the target location ?*

Author Response 9: Because the underlying hydrologic model is not the novelty of this method nor is it the only simulation method that can be bias-corrected in this way, we chose not to go into detail on the exact simulation methodology. The approach thoroughly described in Farmer (2016) is followed identically here. The semivariance was calculated for all contemporary pairs of daily streamflow observations, after transformation by dividing by drainage area and taking the logarithm, in a 2-digit HUC. These seminvariances were then summarized with a single semivariogram cloud, which was used to fit a semivariogram model that applies equally for all days. The 2-digit Hydrologic Units are the large regions used to classify basins in the United States and shown in Figure 1 (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.). These clarifications will be provided in the revision of this manuscript and we will explore the feasibility of a graphic to summarize the simulation routine.

Reviewer Comment 10: P4. Ln 1-8 : *Please clarify. "Daily FDCs were developed*
C6

independently of the streamflow simulation procedure by following a regionalization procedure similar to that of Farmer et al. (2014). You mention "These same percentiles were then estimated using a leave-one-out cross-validation of regional regression". The objective/process of this estimation was not clear. Please make clear that the FDC used to bias correct your target station is obtained from a regional FDC model, obtained from all (or a part of) the stations close to this target station.

Author Response 10: We will provide clarification in the revision. Farmer et al. (2014) use unsupervised regional regression that relies on a best-subsets regression to estimate the complete FDC in a way that tries to capture the dependence between quantiles. The leave-one-out procedure is used to quantify performance as if the target site were completely ungauged. Regionalized regression relies on all stations within a pre-defined region (in this case, the 2-digit Hydrologic Unit).

Reviewer Comment 11: *P4. Explanatory variables. Please comment the possible correlations between those. Have you looked / used for uncorrelated sets of explanatory variables ?*

Author Response 11: Farmer et al. (2014) discuss how explanatory correlations were controlled. Most of this detail, being already published elsewhere, is beyond the scope of this current work. However, we also used a limitation of variance inflation factors of each model. This will be added to the revision of the manuscript.

Reviewer Comment 12: *P4 – In 10 : what are "best-subsets regression". Best subsets of what ?*

Author Response 12: 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 of model sizes. As described by Farmer et al. (2014), of these models that then

C7

differ in size, the AIC (or some other metric) is used to select the "best" model in an unsupervised fashion.

Reviewer Comment 13: *What is the difference with "three contiguous streamflow regimes you mention in line 17.*

Author Response 13: The three contiguous regions allow for different explanatory variables to be used to predict different streamflow quantiles. Each is contiguous in that it must span a contained range of quantiles, e.g. (0.02%, 0.05%, 0.1%, etc.) and not (0.02%, 75%, 99.98%). Because three are allowed, these can be thought of as "low", "medium" and "high" streamflows. We will provide this example in the revision.

Reviewer Comment 14: *P4; Please clarify. "the percentiles were grouped into a maximum of three contiguous streamflow regimes based on the behavior of the unit FDCs in the region". What is a unit FDC ? How are defined the regimes ? Are they relative to different parts of the FDCs or to different sets of different basins within the region ? Do you group 95th and 90th percentiles for instance ?*

Author Response 14: The unit FDC is the duration curve of streamflow divided by drainage area; this definition will be provided in the revision. Based on a national analysis of unit FDCs (Over et al., in press [expected to be published before revision]), it is possible to identify low, medium and high streamflow regimes. This explanation will be provided.

Reviewer Comment 15: *P4. "The percentiles in each regime were estimated by the same explanatory variables, allowing only the fitted coefficients to change.." Do you mean that a regression is estimated for each of the Twenty-seven percentiles you considered ? If yes, a very significant risk on a non monotonic behavior of the FDC is likely. Why not working with a analytical model fitted to the FDC (and then propose*

C8

a regionalization of the parameters of this model ?). This issue should be discussed somewhere.

Author Response 15: The answer to the reviewers question depends on how the reviewer is defining "regression". Within a given regime, as defined earlier, all of the quantiles have the same explanatory variables in the final equation. However, the coefficients on those variables are fit independently across quantiles. Because the same variables are used, monotonicity is made much more likely within the regime; the only nonmonotonicity therefore occurs at the separation of regimes. We will add a discussion of this point. Furthermore, identifying the optimal method for FDC prediction is beyond the scope of this work. For this reason, we did not explore analytical solutions for the FDC (an aside: <https://doi.org/10.5194/hess-21-3093-2017> found it is very difficult to find a suitable analytical solution).

Reviewer Comment 16: P4. "Further details on the approach of percentile grouping this methodology can be explored in the associated data and model archive Farmer et al. (2018)." The given reference is a technical report. The soundness of the "percentile grouping approach" is thus uncertain. It has thus to be fully explained in the present manuscript or it has to be described in another a peer reviewed journal.

Author Response 16: The report that the reviewer is looking for is in press and a citation will be added (Over et al., in press [expected to be published before revision]). The data release includes all the development code, allowing a user to explore the method in depth, if interested.

Reviewer Comment 17: P4. The section on the bias correction approach has to be made clear. Especially how the bias correction works and with which data in the case of ungauged catchments. Again a synthetic graphical scheme of the BC approach would be usefull. You say "The nonexceedance probabilities were then converted to standard normal quantiles and linearly interpolated along two types of independently estimated

C9

FDCs: the regionally regressed FDCs and the observed FDCs determined by applying the Weibull plotting position." It is something to do with quantile mapping? I can imagine that a correction function can be established for any given station where you have both observations and simulations. Is the idea here to transfer the correction function obtained for a gauged catchment to neighboring ungauged ones ? If yes, how do you define the neighborhood ? how do you consider the different corrections functions you can obtained for the different gauges stations you may have in the neighborhood of your target location ?

Author Response 17: We will clarify in the revised manuscript that the idealized method that uses observed FDCs is only an example to demonstrate theoretical utility. In practice, only an estimated FDC can be used. There is no development or transferal of correction functions. We will provide a graphical demonstration of the methodology.

Reviewer Comment 18: P4. Ln – 29/30 : "The nonexceedance probabilities were then converted to standard normal quantiles and linearly interpolated along two types of independently estimated FDCs: the regionally regressed FDCs and the observed FDCs determined by applying the Weibull plotting position." Åž This is confusing. In the ungauged basin configuration : only one FDC is expected to be used : the regionalized one. The observed FDC is not expected to be available in a ungauged catchment. You use it here only to estimate the added value of the quality of the regionalized FDC on the resulting bias corrected discharge time series. The "independently estimated FDC" you mention in line 31 (and basically everywhere else in your manuscript) should first refer, I guess, to the regionalized FDC (and not to the FDC from observations).

Author Response 18: The reviewer is correct, as we have addressed in other comments. The use of an observed case is merely for the purpose of demonstration. The reason we differentiate between "independently estimated FDC" and "regionalized FDC" is that regionalization is not the only way to estimate an ungauged FDC. This will be clarified.

C10

Reviewer Comment 19: P5. Ln 3: reformulate : "by correcting the simulated volumes to an independently estimated FDC.

Author Response 19: We will revise to "by *rescaling* the simulated volumes to an independently estimated FDC."

Reviewer Comment 20: P5. Evaluation : Ln 10- the two different evaluations approaches were not clear for me at first. A reformulation would be worth Ln 20 & follow. Evaluation for the whole FDC, or for a given tail : what is the evaluation criterion : the mean of the bias estimated between each pair of percentiles ? the bias between the mean of the percentiles for the raw and corrected data respectively ?? Evaluation on observation-independent tails and observation dependant tails. A graphical scheme please to explain what is done in the second case, at least in a supplementary material !

Author Response 20: We will provide a graphical scheme of the two types of tails in the revision. Several criterion are provided for evaluation (e.g., average differences in common logarithms, RMSE, NSE). The median across all sites is taken as the average performance for our data set. There is, of course, some spread around the central tendency, and this is discussed, but using the central tendency is a traditional means of assessing bias.

Reviewer Comment 21: Results : all results are given in the form of boxplots. This is likely not enough to understand how the methodology work and how good it is. To give for a selection of stations the different curves (observed / regressed / regressed+bias corrected) would be helpful (with a good performance station and a bad one for instance)

Author Response 21: We will provide a figure of some example cases in the revision.

Reviewer Comment 22: P5 ln33 and p7 . Ln 25. I find the term "timing" and "error in C11

timing" not appropriate. You could perhaps say " an error in timing of the percentiles". This is however more an error in the temporal structure of your simulated time series. This results from an error in volume which is one day an over-estimation of the true volume and the day after an under-estimation.

Author Response 22: We will consider how to clarify this. "An error in timing of the percentiles" is misleading, as it is not the percentiles that provide timing. Timing is provided by the relative ranks and nonexceedance probabilities. By "timing error", we intend to refer to the "error in the temporal structure". We will use this formulation in the revision. It is true that a transposition of the volumes could appear as a timing error, but this is not explored here (a discussion point will be added).

Reviewer Comment 23: P7 ln34. You say "These timing errors also almost result in errors in a particular direction: low for high flow and high for low flows". You have perhaps such a mean behavior but as mentioned above, you have a number of low flows that are overestimated but also a number of low flows (less frequent) that are underestimated. . . Please reformulate to put this statement in perspective.

Author Response 23: We will clarify that we are not making a general statement, but observing a central tendency amid a range of values.

Reviewer Comment 24: P6. Ln 2 : "Figure 4 and Table 1 summarize the tail bias in all approaches to streamflow simulation considered here.". What are the 3 different approaches ? This has to be clearly explained previously. The BC-Obs is not an "approach" similar to the 2 others as you do not know in principle the observed discharge for the target ungauged basin. It just allows you to identify the influence of the quality of the regionalised FDC. A reformulation is required when relevant in all the manuscript. The presentation of the method has probably also to be restructured to make it clearer.

Author Response 24: This was addressed in a previous comment. A graphical presentation will be provided to distinguish between observation-dependent and

observation-independent tails. The BC-Obs is provided for exactly the reason the reviewer describes: not as a viable alternative for application.

Reviewer Comment 25: *P7. 1. The analysis of the second paragraph in this section 3.1 is clumsy. The Observed FDC is in principle perfect and thus the bias in simulation for the observation independent evaluation case should fully vanish after correction. You should have nothing or roughly nothing to comment here. The results should be perfect. Why is there some remaining bias with the BC-obs approach ?? Please comment. (could it be a difference in the time period used for the simulation and the time period used to identify the observed FDC? Is it something related to the epsilon value you add to discharge data for the logarithmic transformation issue ? to the reduced number of percentile used to describe the FDC ? something else ?) The comment on this in the conclusion section has also to be modified accordingly (p12 – ln 1/10). P7. 24: "To understand the effect of errors in timing further, consider Figure 6, which shows the mean error in the nonexceedance probabilities of the observation-dependent upper and lower tails." I can not understand (just guess) what is referred to here ? Please clarify.*

Author Response 25: The use of an observed FDC will still rely on the timing of the simulated hydrograph. It is for this reason that you are likely to see some residual bias in the observation-dependent tails (Fig. 4, boxplots C and F). With the observation-independent tail (Fig. 4, boxplots I and L), the much smaller residual bias is likely a result of the interpolation along an FDC defined by a finite set of quantiles. There may be some effect from the epsilon value and censored regression used for zero-valued streamflows, but we did not find a major impact. We will add a note on this discussion.

Reviewer Comment 26: *Fig 2 and following : how many data are used for each boxplot (no. stations x no. percentiles ???). Is there one point for each station/percentile ?*

Author Response 26: There are 1168 points in each boxplot (one for each station);
C13

we will make this clearer in the caption. The metric is the mean difference in common logarithms of the complete FDC at a given site.

Reviewer Comment 27: *Discussion : what about the likely seasonal dependence of the correction function ? Is there some potential for improvement here ? The estimated FDCs are composed of 27 quantiles, of which the upper and lower tails contain only the eight values with nonexceedance probabilities 95% and larger and 5% and smaller, respectively.). A comment on the number of quantiles used to describe the FDC would be worth (a sensitivity analysis of the results to this number could be also included in a discussion section)*

Author Response 27: No correction function, as such, was used. (It could be conceived as a function, but that is not explored or proposed here.) There is, of course, always room for improvement; we provide a discussion of the sensitivity to the representation of the FDC. We will strengthen that discussion by noting the limitations of an FDC defined by a finite number of quantiles.

Reviewer Comment 28: *What is the influence of the duration of the observation time period used to estimate the observed FDC > on the quality of the FDC estimation and then on the quality of the bias correction ?*

Author Response 28: As the period of analysis is fixed here, it is not possible to explore this effect. We will make a recommendation for future work, as it will certainly have an impact on something like future projections.

Reviewer Comment 29: *Is there any dependence of the results to the hydroclimatic context of the basin ? How is it structured is space in US ?*

Author Response 29: We could not see any obvious pattern: "Initial exploration did not find a strong regional component to performance of the bias correction method. For

some regions, like New England, where FDCs are well estimated by regional regression, there is a general improvement in accuracy under bias correction with regionally regressed FDCs, but the improvement is highly variable. Instead, the strongest link is with the reproduction of the FDC. " It may be that the hydroclimate is driving the ability to reproduce the FDC through regression, but that is left for future research.

Reviewer Comment 30: *Minor remark : P2. Ln10. The interest of the "long term forecast term" here is not clear. It seems to be out of the scope of the work. To be better explained. What is long term ?*

Author Response 30: Exploration of long-term forecast is beyond the scope of this work and no effort is made to explore it here. Long-term (decadal and beyond) forecasts are mentioned as an example of a hazard of underlying bias. We will clarify this statement by defining long-term as decades and beyond.

Reviewer Comment 31: *P2. 15/20 : " this paragraph is not clear " to be clarified / double-formulate. The "interpolation of non-exceedance probabilities along the FDC" is a rather clumsy formulation. What does it mean ?*

Author Response 31: Described further in the methodology, the interpolation must occur because the FDC is being represented as a finite set of quantiles (27). If a non exceedance probability does not fall exactly on one of those percentiles (P4, L10), it must be interpolated. We will reference the methodology in the revision.

Reviewer Comment 32: *p2. Ln 33 : please clarify what is meant there*

Author Response 32: As before, if the nonexceedance probability falls within the range of the quantiles (P4, L10), it can be interpolated. However, it falls outside of the range of quantiles (P4, L10), it must be extrapolated. The two nearest points were used for linear extrapolation. We will consider how to clarify.

C15

Reviewer Comment 33: *P2. Ln 35 and following. This does not belong to the introduction but to a discussion section. The discussion should probably give the evaluation of the other method suggested here (Additional research to explore if estimating nonexceedance probabilities directly, as opposed to the conversion of simulated streamflow to nonexceedance probabilities used here, might further improve nonlinear spatial interpolation using FDCs or simulation more generally*

Author Response 33: This will be moved to the discussion. Further evaluation is outside the scope of the present work.

Reviewer Comment 34: *p3. Ln 3 : give the structure of the paper*

Author Response 34: This will be added, recapitulating the headers of the document.

Reviewer Comment 35: *p5. "The root-mean-squared error of the common logarithms of streamflow and the differences therein were used to quantify accuracy." Do you mix streamflow and differences between streamflow in the computation of a single RMSE criterion ? If yes, I fear it is not relevant or please clarify / justify.*

Author Response 35: We do not mix streamflow and differences between streamflow in the computation of a single RMSE. The RMSE is calculated for two different approaches; we then observed the differences across approaches. We agree that this sentence is confusing, so will explore other options. One possible revision could be "The differences in the root-mean-squared error of the common logarithms of the predicted streamflow for the two approaches were used to quantify accuracy."

Reviewer Comment 36: *P6. Ln 9 – add a subsection title "simulated hydrographs without correction"*

Author Response 36: This section will be added.

C16

Reviewer Comment 37: P6 – 25. “These results show upward bias in lower tails and downward bias in upper tails.” No, this is not the case in general. See your paragraph above.

Author Response 37: We do not claim that the result is general, but the result do show a tendency in the direction described. We also discuss the variability of this performance, but this will be further highlighted in the revision.

Reviewer Comment 38: P9 “For the observation-independent case, the errors are removed almost completely, and the remaining errors in the observation-dependent case mimic the timing (nonexceedance probability) errors.” This is not true (only if the observed FDC is used)

Author Response 38: The first half of this paragraph only discusses the case where observed FDCs are used. We will clarify this transition.

Reviewer Comment 39: P10. 2 : “The change in the absolute bias of the observation-independent lower tail has a 0.72 Pearson correlation with the absolute bias of the lowest eight percentiles of the FDC estimated with regional “regression.” I do not understand what is meant here.

Author Response 39: We will add, “That is, the residual bias in the bias-corrected FDC is strongly correlated with the bias in the independently-estimated FDC.”

Reviewer Comment 40: Figures P10- 6 “as regional regression is not the only tool for estimating FDCs, improved methods for DC estimation would further increase the impact of this bias correction procedure.” Mention other such methods.

Author Response 40: 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.

C17

Reviewer Comment 41: P10.15 “It may not always be possible to determine the accuracy with which a given FDC estimation technique might perform, making it difficult to determine if these results can be generalized.” There is no reason why the accuracy of any FDC regionalization approach could not be assessed (this is done in all FDC regionalization study). At least with a leave-one out procedure. Before applied for the bias correction of any time series simulation model, this quality should be checked and estimated to be better than that of the simulation model.

Author Response 41: We will revise to, “At a particular ungaged location, it may not always be possible to determine the accuracy with which a given FDC estimation technique might perform (beyond a regional cross-validation assessment), making it difficult to determine if these results can be generalized.”

Reviewer Comment 42: Fig. 1 – What is meant here : “The outlines of 2-digit Hydrologic Units are provided for further context.

Author Response 42: The polygons represent the 2-digit Hydrologic Units, which are the regionalization areas. The meaning of these units is described in an earlier comment and will be incorporated into the revised manuscript.

Reviewer Comment 43: Fig. 2 : BC-RR and BCObs have to be defined in the main text.

Author Response 43: We will consider adding these into the report. They are not currently used in the main body, as the repetition of opaque acronyms might detract from clarity.

Reviewer Comment 44: Figures and tables : simplify the captions : a number of repetitions could be removed (and a reference to the caption details of one reference figure added in the caption of all other figures)

C18

Author Response 44: We prefer that the captions allow each figure to stand alone as much as possible.

Reviewer Comment 45: *Are the 3 tables usefull ?*

Author Response 45: We feel they are useful, as they provide a summary of the numbers discussed in the report.