Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-83-RC1, 2020

Hydrology and Earth System Sciences
Open Access
Discussions

HESSD

Interactive comment

Interactive comment on "Comparison of occurrence-bias-adjusting methods for hydrological impact modelling" *by* Jorn Van de Velde et al.

Stefan Lange (Referee)

slange@pik-potsdam.de

Received and published: 22 July 2020

General comments

The authors of "Comparison of occurrence-bias-adjusting methods for hydrological impact modelling" test different methods for the bias adjustment (BA) of daily precipitation time series. In particular, they combine different methods for dry-day frequency BA with one method for the BA of wet-day precipitation intensities (Quantile Delta Mapping or QDM), and they test how an additional temporal shuffling of the time series for the purpose of multivariate bias adjustment (MBA) influences the results. The different method combinations are evaluated using precipitation occurrence statistics, precip-

Printer-friendly version

Discussion paper

itation quantiles, and discharge quantiles, with discharge computed using a lumped rainfall-runoff model. The authors find that the stochasticity added by the more complicated methods worsens results and conclude that simpler methods should be used for BA until the stochasticity-related problems identified here are better understood.

While I see value in investigating the performance of combined BA methods whose components have only been separately evaluated before, I am afraid that the present study is not a good example of such an investigation, for the following three main reasons. First, QDM alone is found to already perfectly adjust the dry-day frequency, hence no added value can be expected from combining QDM with any designated dry-day frequency BA method. Secondly, most of those combinations are done such that they are bound to fail by construction. Thirdly, where bad performance is found it is not explained mechanistically.

The first point is so important because the biggest differences in performance between method combinations were found for precipitation occurrence statistics. The authors write that QDM is only applied to wet-day precipitation (more than 0.1 mm/day), yet for some mysterious reason that application also perfectly adjusts the dry-day frequency (Fig. 4, Table A1). This result, which surprises me, is neither discussed nor explained. It could be related to a QDM application that was not done as described. Or it happened by chance and would not happen again if different climate data were used. In any case, it is an unfortunate result because it has the effect that no other method combination was able to better adjust the dry-day frequency and hence the simplest approach was bound to perform best.

The second point makes that even worse. You combine Stochastic Singularity Removal (SSR) and Triangular Distribution Adjustment (TDA) with QDM such that they are unable to adjust the dry-day frequency by construction. Your version of SSR+QDM first turns all dry days into wet days; then QDM is applied to all days (because they are all wet now), which changes the number of days with precipitation less than P_{min} in an uncontrolled way; then all days with precipitation less than P_{min} are turned into dry days, so the dry-day frequency is adjusted in an uncontrolled way. The problem is less pronounced but similar for TDA+QDM, where TDA is first applied to properly adjust the dry-day frequency, but then QDM is applied, which moves some formerly wet days below the wet-day threshold and hence turns them into dry days; therefore TDA+QDM always generates too many dry days (Tables A1 and A2).

The third point is that as you can read above it is relatively easy to explain most of your results. Yet you do not even try to do that. Just presenting the results without explaining them will not help others to better combine bias adjustment methods in the future.

Specific comments

l 157: I think you should mention that your Thresholding method is not able to account for a climate change signal in the dry-day frequency.

l 165-167: Vrac et al. (2016) set $P_{\text{min}} = 8.64\,10^{-4}$ mm/day. That is much lower than your P_{min} value and I am sure that this influences your results.

l 173: Why is SSR not also applied month by month like Thresholding and TDA? Say so if that would not make a difference.

l 175: What is the motivation behind TDA? You only explain how the method works but not why (with which intentions) it was developed. What can it do better than other methods?

l 177f: What if that is not possible, for example if $f^{\text{ho}} = 0.9$, $f^{\text{hs}} = 0.3$, $f^{\text{fs}} = 0.5$?

l 194: Not clear to me: Do you loop through 1–3 one day at a time or how? I am confused because l 189f suggests something else.

Fig. 2: This figure is currently not helping to understand the TDA method. Please add text to the caption that explains what you intend to illustrate here. What are the red and black arrows? Is this supposed to explain the dry-to-wet case or the wet-to-dry case? Is precipitation decreasing on the x axis in panel (a) and if yes then why? Why

is $x_{\text{th}} < 1mm$, should it not be bigger because $b = 0.9$? Is that supposed to be the CDF of the triangular distribution in panel (b) and if yes then why is the curve not smoother?

l 218-220: There are indices i missing in the innermost brackets of both equations.

l 221f: Which version is used for evaporation?

l 223: You should also specify how many empirical quantiles you used in your QDM configuration.

l 227-229: Which threshold was used for evaporation and why? What do you mean with "calculated"? Do you mean that the empirical CDFs in Eq. (6) where based on wet-day values only? Please be more specific.

l 248-250: Does that mean that MBCn was only applied to days with precipitation > 0.1 mm and evaporation > 0.1 mm? Or do you mean that QDM in the last step of MBCn was only applied to days with precipitation > 0.1 mm and evaporation > 0.1 mm?

l 255f: Cannon (2016) uses the additive version of QDM for that. Did you do the same?

l 270-285: This description is quite confusing. I think you do not need to mention the Schaake shuffle and you do not need Eqs. (11-16) because all you do in the last step is shuffle $\mathbf{X}^{\text{fa}}$ along the time axis such that the rank time series become identical to those of $\mathbf{X}^{\text{fs}}_{[k]}$, where k is the number of iterations of steps one to three.

l 286f: Does that imply that MBCn was also applied using 91-day running windows etc.? You better be more explicit here.

l 295f: Cannon (2016) also suggests early stopping. Have you tested that? What did you do to detect and prevent overfitting?

l 297: It would help the readers a lot (see below) if you inserted an extra section here that spells out how exactly you combine the dry-day frequency BA methods with QDM/MBCn, and how QDM/MBCn are applied when they are not combined with any dedicated dry-day frequency BA method. For example, in TDA+QDM, is TDA done before or after QDM? The same for Thresholding+QDM etc.

l 310: You also need to specify whether you account for seasonality or just pool data from all days of the year.

l 312: Why can that be misleading?

l 321-327: Which variables are used to drive PDM? Precipitation only or precipitation, evaporation, and temperature? I assume all three variables because otherwise it would make no sense to apply MBCn.

l 327: What did you use as your true discharge? Did you use observed discharge or did you generate it running PDM based on the observed climate data?

l 332: MBCn is also stochastic, or did you use fixed rotation matrices?

l 368: Which kind of extrapolation are you talking about here? There is no extrapolation necessary to make Eq. (6) work for all values of x^{fs}.

l 380-396: It is impossible to understand these results if you do not spell out exactly how you combine the dry-day frequency BA methods with QDM and how QDM is applied when it is not combined with any dedicated dry-day frequency BA method.

l 389f: Who assumes that? Do you have a reference or can you motivate that assumption? I would not have expected the stochasticity added by SSR and TDA to improve the time series.

l 403-407: Your discussion of those small differences raises the question of statistical significance. There are no uncertainty estimates provided anywhere in your results section. Either you change that (using bootstrapping or the uncertainty originating from the stochastic nature of SSR and TDA, for example) or you are more careful with such statements.

l 421-423: I do not understand. Please rewrite.

l 423-428: This part is quite difficult to read. Please rewrite. You are just describing what can be seen in Fig. 7. This can be done more intelligibly.

l 432: Do those ranges include results for SSR or not?

l 432f: Here you start discussing the influence of QDM on precipitation occurrence but then you just stop and do not follow up in the discussion section.

l 435: “might be responsible for the performance” and “plausible cause”: Why so vague? The shuffling changes autocorrelation and transition probabilities but not the distribution. There is no doubt about that.

l 436-438: Aha. You should mention this earlier, ideally in a dedicated section, see above.

Table 2: It is unfortunate that you do not just plot all index values in your figures. I suggest you just extend the axis limits and get rid of this table. Then all the information can be found in one place and in one format.

l 445-448: Can you explain that?

l 450-452: Is that conjecture supported by Q_5 being too high?

l 452 You keep writing “5th quantile” where I think you should write “5th percentile”.

l 457f: Again, why would you assume that?

Table 3: What is the point of this table? All of this can be seen in Fig. 8, so I suggest to remove it.

l 461-465: I think you are right but this can be spelled out more: SSR and TDA change precipitation ranks randomly and the partly randomized ranks are then shuffled by MBCn. Then PDM calculates discharge, the result of hydrological processes over many days. One could say that discharge integrates precipitation (an other fluxes) over space and time. Therefore, changing the temporal order of events has an effect on discharge. The impact on the high discharge quantiles is particularly strong for TDA (Fig. 8 vs. Fig. 5) because by setting $b = 0.9$, you have configured TDA such that it may turn pretty heavy precipitation days into dry days.

l 466: Why do you cite Switanek et al. (2017) here? What in that paper is related to your statement?

l 467: I think you should show those figures.

l 469-472: You are talking about quite an extreme quantile here. Such a quantile can only be estimated with limited precision. You really have to quantify the statistical uncertainty of your results and measure the significance of differences if you want to make statements like these.

l 473-476: I am unable to see in Table A2 what you write here.

l 482f: Also about this I would have liked to read a deeper discussion.

l 489-490: That is not correct for all indices.

l 495: You write that SSR+QDM and TDA+QDM performed "unexpectedly" worse. I think they did not, see my general comments.

l 495-501: What you do not mention here is that Vrac et al. (2016) used a much lower P_{min} value (see above) and that they replaced zeros with random numbers from $]0, P_{\text{min}}[$. Then they apply CDFt, which by design adjusts the dry-day frequency, in contrast to QDM. The SSR method was designed to work well in combination with CDFt. You are not using SSR as intended by Vrac et al. (2016).

l 506f: This has already been done, see for instance François et al. (2020, doi:10.5194/esd-11-537-2020).

l 524f: That is easily explained by the drizzle effect which is of course perfectly adjusted by thresholding but not by SSR+QDM and TDA+QDM, see my general comments.

l 525f: That needs an explanation!

l 526f: That statement is just wrong. MBCn has no influence on the performance of the dry-day frequency BA methods. All it does is shuffle the time series.

l 530-535: To also say something positive, I support those statements.

l 535-538: I support your call to always use the simplest possible method. Yet as outlined in the general comments, I am afraid your study is ill-designed to support it.

Table A1: QDM alone does not only perfectly adjust N_{dry} but also P_{P00} and P_{P10}. How is that possible?

Tables A1 and A2: I suggest you use better column headings: QDM, SSR+QDM, TDA+QDM, Thresholding+QDM, and similarly for MBCn.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-83, 2020.