Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-145-RC3, 2016

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Disentangling timing and amplitude errors in streamflow simulations" *by* S. P. Seibert et al.

Anonymous Referee #3

Received and published: 16 May 2016

Summary

This study focuses on the Series Distance criterion (SD) introduced in Ehret and Zehe (2011) to consider both magnitude and timing errors when evaluating streamflow simulations. The paper first details the principle of the method, its technical aspects, as well as the improvements brought to SD since the work of Ehret and Zehe (2011). It then proposes to apply SD to represent the uncertainties of deterministic simulations. The authors use a case study to illustrate and evaluate the ensemble dressing approach based on SD and based on a benchmark approach.

General comments

I really enjoy this article. I think it is well-structured and I found it pleasant to read because the authors set the context for each step and well describe the concepts. The

illustrations well support the explanations of SD.

Overall, I would advise shortening Sections 2.2 and 2.3 and adding some text to better explain and discuss the new aspects of SD that are introduced in this paper. In Section 3, I found very interesting the use of the proposed error to build uncertainty ranges. However, in Section 5.3, I don't understand why the uncertainty ranges are so different even though they were both built to achieve an 80 % coverage. My questions on this section are listed in the detailed comments.

Detailed comments

Lines 151-162 : transform this list into a table so that the reader better visualizes the new aspects proposed. Here, the "New" sections are lost at the end of the sentences

Line 176 : I think "low-flow" should be replaced with "low flows" when it is not an adjective.

Line 200 : "this approach has been shown to work well" : could you specify ?

Lines 262-263 : I recommend removing "(or vice versa, if a falling segment is dissolved)". The sentence line 261 follows the example on line 249, and I don't think the information in parenthesis necessary to understand as it is already clear. More generally, I would recommend removing most parenthesis in the article and either include the content in the text or remove it. This could further improve the reading of the article.

Lines 268-271: "Since the segments can differ in length. . .between the time series edge nodes." Here, you say that you interpolate between edge nodes to obtain the same number of points to compare between the obs and sim, and, in the case of the obs time series, each edge node corresponds to an actual observation. Is that right? If so, my question is: if you linearly interpolate a segment between edge nodes in the obs time series, won't you create "fictive" observations by filling a segment? And wouldn't this "falsify" the evaluation of your simulation time series? Could you clarify this?

Line 290: "In this case," a space is missing after the comma.

Line 317: “obs an sim”, “an” misses the final “d”.

Line 367: Shouldn’t the uncertainty ranges capture a significant portion of “observed” values rather than “simulated” values?

Sections 3.1 and 3.2: I am not sure I understand how you use the relative contribution to the total error to select your sample. For instance, in the one-dimensional case, do you assume that your uncertainty range is symmetrical around your simulated value? Since the relative contribution is positive, do you pick the value v that corresponds to 80% of the error and use it to build both the lower (-v) and upper range (+v) ? If not, the upper left graph of Figure 4 may have misled me. Or do you directly use the error values behind the relative errors within 0 and 80% and select the largest negative one to build the lower range and the largest positive one to build the upper range?

Line 398: Don’t all ds^2 sum up to 100?

Line 449: Remove one of the “period”.

Lines 552 and 773: The reference for Ewen is not formatted properly.

Line 661: I am not sure about the use of “relevant” here.

Lines 670-672: Just a question, in your opinion, could this be addressed by applying the dressing to adapted time steps? E.g. so that the distance between two dressed time steps is equal to the error in timing? Would there be an optimal time step to apply the dressing (one for which the percentage of sampled errors and the overall coverage would match)?

Section 5.3.2: (1) From Figure 6, it is not intuitive that the SD and BM approaches have the same coverage. Is it specific to this example? Do you have cases when there are more observations outside the SD envelope than outside the BM envelope that would compensate for the case you show? (2) Based on how the SD envelope is constructed, you would expect a third of the 20 % outside the envelope occur in low flows, a third in rising limbs and a third in falling limbs. On the opposite, the 20 % falling outside the BM

envelope can occur whenever in the time series, and can, for instance, always occur when it is harder to model streamflow, i.e. during events. Does this have an impact here? How does that affect the results?

Lines 700-702: From this sentence, it may seem that low flows also have a 2-d error distribution, but errors in low flows are 1-d. Could you clarify this?

Lines 719-720: What would be the difference between “uniting” and “intersecting” in this case?

Line 723: “as proposed by” I believe a reference is missing here.

Figure 5: The dot of the sampled subset in the upper left graph is black whereas the sampled subset is orange.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-145, 2016.