Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-216-AC2, 2020

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Interactive comment on "Implications of Model Selection: A Comparison of Publicly Available, CONUS-Extent Hydrologic Component Estimates" *by* Samuel Saxe et al.

Samuel Saxe et al.

ssaxe@usgs.gov

Received and published: 14 September 2020

Note: The author responses were originally going to be provided as PDFs but upload failed due to technical errors on the Copernicus website. If any reviewers or editors would prefer author responses in PDF or Word format, please contact the lead author via email.

The authors are grateful for the invaluable input from the referee on our paper "Implications of Model Selection: A Comparison of Publicly Available, CONUS-Extent Hydrologic Component Estimates". The reviewers comments are insightful and identify to the authors areas where clarification would increase the effectiveness of the manuscript.

Printer-friendly version

Discussion paper

Major Comments

Regarding the readability of section 3.3:

We agree that the use of subsections would increase the readability, as would explaining to the reader how each method contributes to the analysis and discussion of the manuscript.

Regarding the use of time periods for analysis:

Time periods were used primarily to limit the effects of variable models counts within each water year. Datasets were typically available in either the (a) 1985-1999, (b) 2000-2014, or (c) 1985-2014 time periods, with counts of models available per year changing between water balance components. By dividing summary statistics into two time periods (Early vs. Late), we attempted to reduce biases that would be introduced into uncertainty values by having more models (i.e. greater uncertainty) or less models (i.e. less uncertainty). However, the other reviewer suggested using a bootstrapping methodology to calculate uncertainty that will further assist in reducing the biases. Using that method may allow us to remove the time periods.

Regarding the exclusion of two ecoregions from Figure 8:

The two smallest ecoregions were excluded because, as the reviewer notes later in their comments, many of the figures in this manuscript are extremely informationally dense. In this case, we simply wanted to provide a more concise figure because it is used more to visually relate water balance uncertainty than it is to provide the reader with raw values.

Regarding more clearly highlighting the originality of the study:

We agree that this can be better discussed within the paper. This will be amended during revisions.

Regarding line 294:

Line 294 states that “Disagreement in the presence of significant trend and trend direction is quantified using the unalikeability coefficient (u) which measures how often categorical variables differ on a $0 \leq u \leq 1$ scale, with 0 and 1 being complete agreement and disagreement, respectively (Kader and Perry, 2007). Thus, the value of 0 occurs if all datasets agree on trend direction. By stating on Line 464 that “Runoff datasets show the most consistent spatial distribution of $u > 0$ across the study ecoregions”, we are explaining (in an unclear way) that runoff datasets are most commonly in disagreement across the CONUS. This will be edited for clarity in the revisions

Regarding the description of Figure 3c (line 355):

We agree that it would be useful to label each hydrologic model by its category (LSM vs. CM vs. WBM). This had actually been done earlier but discarded because the figures became too dense for the reader, but perhaps we can fit labels along the right-hand y-axis, or potentially shade the figure background by model type (e.g. white = LSM, light grey = CM, dark grey = WBM).

Regarding missing subplot in Figure 7:

The empty space is left in place because we only have one remote sensing dataset for SWE and wanted each water balance component to have it’s own row. We can’t fill the empty space with a subplot of runoff (R) because there is not a satellite dataset that measures runoff, at least until SWOT is launched in 2022 and even then will not have retrospective estimates during our study period. This was not mentioned in the methods section discussing correlation statistics, so the revision will add it in for clarity.

Regarding Figure 8:

The 10th and 90th percentile boxplot lines are not shown when they exceed the boundaries of the subplots. We were on the fence about whether to even include the boxplots as overlays on the histograms because they tend to “busy” the figure, so to speak, without providing much additional information. If parts of the boxplots have to be excluded

in some subplots, perhaps it is best just to drop the boxplots entirely and leave the subplots as just histograms.

Regarding Appendix Figures 2.5-2.9:

Yes, several of the dataset names overlap each other. This is something we will correct in the revisions. It was difficult to squeeze so much text into individual figures. We placed these figures into the appendix rather than the main manuscript body because they are quite large but difficult to shrink down considering all the included text. With Figure 6, our goal was to provide most readers with a simple visual representation of the general disagreement in model trends. Appendix Figures 2.5-2.9 were attached to provide more detailed results for readers that may be interested in specific models used in the study.

Regarding Appendix Figure 2.10:

I believe you are misinterpreting the grouping labels in this figure. The first three groups, labeled as “Cor > 0.90”, “Cor 0.50-0.90”, and “Cor < 0.50”, are grouping values of correlation measured with Spearman’s rho, denoted in the text with the Greek letter . Models are only assigned to these groups when their correlation is statistically significant, calculated using a binary p-value significance test. If the p-value of a significance test is less than our assumed alpha value (α = 0.05), then correlation is assumed to be significant. So the first group, “Cor > 0.90”, identifies models with very strong significant correlation. The second group, “Cor 0.50-0.90”, identifies models with moderate to good significant correlation. The third group, “Cor < 0.50”, identifies models with poor to negative significant correlation, including anything with rho values of -1 to +0.50. The fourth group includes any model with statistically insignificant correlation but does not report the actual correlation value since it is deemed irrelevant by the significance test.

Regarding the general remark:

We agree that there are much more figures than are typically found in a journal paper, as well as being more informationally dense than usual. However, our goal with this study was to provide readers access to as much information as possible while still maintain a decent “readability” so that those readers interested in specific models can find the relevant information without having to delve into the actual datasets being released in tandem with this study. For example, we want the reader who is utilizing the NLDAS2-Noah land surface model in California to be able to compare their monthly and annual estimates to a range of other models without having to acquire, process, and interpret all the other models themselves.

Our hope is that these information-dense figures will allow the scientific community to more easily include uncertainty constraints in their results and analyses. We see this manuscript essentially as a review of the current state of knowledge within the various modeling communities measured in terms of uncertainty.

Minor Comments

Regarding abstract length:

This will be shortened during revisions.

Regarding dataset types:

Unfortunately, remote sensing datasets are much more limited than hydrologic model datasets. We try to discuss differences between dataset types in more general terms to soften potential biases resulting from different numbers of available datasets by water balance component. However, different remote sensing datasets estimating the same water balance component will likely use the same underlying observational measurements (e.g. MOD16-A2 and SSEBop both use MODIS sensor data). Because of this, we believe that comparing the magnitudes of just one or two remote sensing datasets against many hydrologically modeled datasets is effective in at least representing general differences.

Regarding the use of RZSM:

This was also noted by the other reviewer. We will switch to using “SM” as an abbreviation during revisions.

Regarding numbering errors:

This was also noted by the other reviewer. This will be corrected during revisions.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-216, 2020.

Printer-friendly version

Discussion paper