Geosci. Model Dev. Discuss., 5, C621–C625, 2012
www.geosci-model-dev-discuss.net/5/C621/2012/

Geoscientific
Model Development
Discussions

Interactive comment on "A web-based software tool to estimate unregulated daily streamflow at ungauged rivers" *by* S. A. Archfield et al.

Anonymous Referee #1

Received and published: 3 September 2012

The paper, "A web-based software tool to estimate unregulated daily streamflow at ungauged rivers" seeks to present "the first publically-available, map-based regional software tool to interactively estimate daily streamflow time series at any user-selected ungauged river location." The paper is well-written and relevant to the readership base of the journal. I generally support publication of this manuscript subject to my general and specific comments listed below.

General Comments:

1. The manuscript title states that the tool presents a way to estimate "unregulated" daily streamflow. This could be interpreted in two ways: (1) in a given record that has been subject to some regulation (such as downstream of a dam) the tool can remove some aspects of this regulation (2) the tool is only valid in unregulated basins, since

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

it cannot remove regulations. I believe #2 is what was intended. In general this point should be clarified, though, to avoid confusion. I would even go so far as to suggest removing "unregulated" from the title to avoid confusion.

2. In methods such as this, does one need to consider routing? Let's assume that we are working in a basin that is ungauged to a large degree, and the only index basin is far upstream of the ungauged point. Ostensibly there would be a timing lag between the index point and the ungauged point. If the authors' proposed method is invalid in such situations, a comment should be made to this effect.

3. From a stylistic perspective, this manuscript assumes that the reader has very good knowledge of statistical hydrology and some of the sub-methods employed in the framework. When discussing the modeling components, the authors use statements such as, "Method A was chosen, it was shown to perform better than B and C." Much of these methods have been published previously, though, so the justifications (comparisons to B and C) may not be necessary. Instead, I'd like to see some more description of the conceptual basis behind the techniques, especially the local regression to estimate the middle of the quartiles, for example. This is important for readers to understand, given that the framework will do these calculations for them and they will not have to do the calculations on their own. Additionally, more explanation would allow the manuscript to have a greater appeal to those readers who may not be intimately familiar with statistical hydrology but need to use the results, such as water utilities or managers. In general there are non-hydrologists in the readership base of this journal so it's a relevant issue. There are some additional specific comments relating to this point below.

4. As far as I can tell, the main problem to be solved is to provide ungauged streamflows for a specific period in the historical record at a specific site. This explicitly assumes stationarity in the record. I could think of three other relevant problems that hydrologists or water managers may be interested in: (1) calculation of streamflows in which there is a trend or changing moments, (2) hydrologic forecasting (of future periods) at an

GMDD

5, C621–C625, 2012

ungauged site, or (3) generation of synthetic streamflows that match the statistics of the historical record. I realize that these three applications are out of the scope of this paper. However, I think the use of StreamStats for delineating a catchment, as shown in Figure 1, the resampling of historical low-flows, the area-weighted method, and other components here could help solve problems (1)-(3). To increase the impact of the manuscript, I really think there should be some discussion of this (i.e., the applicability of some of these concepts to broader water management problems).

5. There is a lot of great information in the figures, but I have two comments regarding their presentation. (1) they are probably way too small, especially figures 3-4. It was very hard to read the text in the figures, and I'm not sure how big they will be in the final manuscript. (2) When discussing the methods, the authors refer to the figures without walking through and explaining each panel. Especially regarding figure 1, it seems like the treatment of text around the figure is incomplete and should be revised. Further specific comments along these lines will be provided below.

6. I'm not sure what the policy of this journal is, but there should be a more formal software disclosure section that tells readers where they can download a copy for themselves. The material is already in the manuscript, but it would be nice to reiterate it at the end. Also, is there a general version of the code or was the model only developed for Connecticut? This seems like a disparity from the title/abstract and the case study treatment later in the manuscript. In fact the authors mention it is a basin-specific tool [page 2512, line 11].

Specific Comments:

1. [page 2504, line 12] I don't think "efficiency" should be discussed in the abstract without any specific mention of what metric is used. I suggest removing this result from the abstract; I don't think it adds much to the discussion.

2. [page 2505, lines 8-11] I was confused between method (1) about rainfall-runoff modeling and (2) empirical approaches. I'm assuming #2 was intended in this

GMDD

5, C621–C625, 2012

manuscript, but it was unclear. This should be rewritten for clarity.

3. [page 2506, lines 18-25] This paragraph was unclear. I suggest the authors introduce the steps of figure 1 explicitly. Then, they should link to the specific sections by number (i.e., "Section 2.1 discusses...").

4. [page 2507, lines 7-10] I would like to see more discussion of how the quantiles from 0.02 to 0.85 were derived. This relates to my general comment #3 above.

5. [page 2509, lines 8-9] The authors argue that the map-correlation method is better than a method that includes drainage-area ratio, because the drainage-area ratio is already used in another part of the method. Why is this an advantage?

6. [page 2510, lines 9-10] "the software tool can be considered a general framework to provide daily streamflow time series at ungauged locations in other regions". I don't understand how it logically follows that because data is available in the US, the authors' method is general. Furthermore, it looks like the software is specifically derived for Connecticut, so the point here is unclear. On the other hand, I do believe the authors have presented a general approach – it's just that the discussion here in the text is confusing.

7. [page 2511, line 10] "The Microsoft Excel spreadsheet..." In this section the authors are talking about some components of the software as if we already know what they are, even though we are hearing about them for the first time. I recommend some explanatory text in the beginning of this section such as "Software package X consists of the StreamStats web applet and an Excel spreadsheet." Then continue with these explanations. I had to parse through this several times to understand what was really happening. Also it's really difficult to understand what's going on with Figure 3 in my copy of the manuscript, so it should probably be enlarged or broken across multiple figures (see general comment #5).

8. [page 2513, line 1] "dam management": Here, as per general comment #1, it is

GMDD

5, C621–C625, 2012

important to distinguish between regulated vs. unregulated flows.

9. [page 2513, line 23] "leave-one-out cross validation" Is there a citation for this method? In general, the case study is so quick here that the authors don't have a lot of time to discuss the specifics behind their method. I would be careful to explain as much as possible in the space allotted. Unfortunately part of the case study reads as if we already understand what the authors did, even though we are obviously reading it for the first time here.

10. [Conclusion] It may be helpful to discuss some issues such as the ones presented in general comment #4 here. Right now the conclusion is a summary of the paper, but the paper was short and straightforward enough where I don't think a straight summary is warranted. I think the readers would benefit from a broader discussion of some of the key issues here, beyond what is written already.

11. [Table 2] What is the ** in the caption? I don't see this in the table at all.

12. [Figure 1] There should be some discussion of the StreamStats tool in the caption as well. There's a gray box in the legend but it's not clear what StreamStats is, if you're only looking at the figure in isolation.

13. [Figure 2] What do the multiple points represent in each log-streamflow plot? Multiple sites? Also there are some typos in the text; "interpolated" and "quantiles" are spelled wrong. Finally, I noticed there's a point in the FDC that does not lie on the interpolation. What type of interpolation is used, and are there allowed to be outliers here? (In other words is the interpolation like a least squares approach where the interpolation line doesn't have to go through all the points, or a piecewise method where every point is preserved?

Interactive comment on Geosci. Model Dev. Discuss., 5, 2503, 2012.

GMDD

5, C621–C625, 2012