Geosci. Model Dev. Discuss., 5, C824–C828, 2012
www.geosci-model-dev-discuss.net/5/C824/2012/

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 #2

Received and published: 26 October 2012

The paper presents a first attempt to develop a software tool to generate daily stream flow time series at an arbitrary location on a stream network. In my opinion, such a tool will be an extremely useful addition to the "hydrology toolbox" currently used by scientists and water practitioners. The paper is generally well written and structured. However, I do see the need for a number of adjustments to improve its clarity and comprehensibility (not only to scientists but also to general water planners), before it is published, as set out below.

General Comments

1. As far as I understand, the tool, as it is currently applied in the Connecticut River Basin consists of two parts: (a) The StreamStats tool to delineate watershed boundary and basin characteristics, and (b) The spreadsheet tool which performs the rest of the

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

estimation procedure. Currently only (a) works online (web-based) while (b) works offline (non- web-based), whereas the title of the paper suggests the existence of a fully "web-based" operational tool. It is not clear whether the intention of the authors is to develop the current tool into a fully web-based tool in the future. If so, this is not stated in the text. The tool in its present form appears to be a "work in progress" towards a fully web-based tool. Hence I suggest that the title of the paper is amended to read as "Developing a web-based software tool. . . . . ." or "Towards a web-based software tool. . . . . ."

2. The StreamStats tool contains information limited to US watersheds and only those in some states at that. The authors do not sufficiently explain how the tool may be practically implemented in any other part of the world, including the underlying data of the StreamStats tool, the components, applications and functionality expected in such a tool, and what data is output by it (See specific comment on "basin Characteristics" given below). The CRUISE worksheet tool too is specific to the Connecticut River Basin as far as I understand. What features should be included in this tool if it is to be implemented in any other region of the world? It would also be useful to know how practical it is to build the two separate components into one compact standalone software (whether web-based or not).

3. The paper is written assuming that the reader is familiar with all the methods mentioned in it. For example the "map-correlation" method is referred to in several places, but nowhere is it explained. The authors also state that the FDC at the ungauged site is estimated using regional regression equations based on basin characteristics, but do not elaborate further on what specific characteristics are considered or what the form and type of the regression equations are (Also see specific comments below). 4. The text does not sufficiently explain the information presented in figures and tables leaving it to the reader to figure them out, which makes the reader's life extremely difficult. All figures are too small and it is next to impossible to read some of them (especially Figures 3 and 4). Also see some specific comments below on figures.

GMDD

5, C824–C828, 2012

Specific Comments

1. The one before the last sentence in the abstract reads as “For the demonstration region. . . . . with efficiency values computed from observed and estimated streamflows ranging from 0.69 to 0.92”. I suggest that the term “efficiency” here is qualified as “Nash-Sutcliff Efficiency” or the sentence is reworded in another way to indicate that the values presented are goodness of fit statistics, in order not to confuse readers by using the term “efficiency” which could mean any number of things. Also it is noted that the 0.69 to 0.92 values have been obtained using natural logarithms of generated streamflows while the same for untransformed streamflows is 0.04 to 0.92. Generally, goodness of fit statistics are evaluated against untransformed streamflows, and presenting the statistics for the transformed streamflows in the abstract might mislead readers about how good the suggested methods are. I suggest that the actual (0.4 t0 0.92) values are reported in the abstract and that the authors try to identify the reasons behind this large variation (for example the method may work well for only a certain range in watershed area).

2. Line 8-9 on page 2507 reads as “. . . . . first developing regional regressions relating catchment characteristics to selected FDC quantiles. . . . . ”, but does not elaborate on what type of catchment characteristics are considered here, leaving the reader guessing. Neither are they explained later, apart from within the section on the CRB where only three characteristics are discussed. The authors should present a broad range of possible characteristics which may be adopted in any other part of the world if the tool is to be reproduced.

3. Line 13 page 2507: “. . . . . selected quantiles on the FDC are estimated from regional regression equations and a continuous. . . . . ”. What is the form of these regression equations and how are the catchment characteristics related to FDC quantiles in these equations? Without this knowledge, the tool cannot be reproduced anywhere else. Although these regression equations are mentioned even later in the text at several places, nowhere are they presented. Merely referring to another pa-

GMDD

5, C824–C828, 2012

per where the method has been applied is not sufficient for a reader of this paper to understand the procedures presented here.

4. Tables 2 and 3 do present information purported to be on these regression equations, however, they are not at all helpful since (a) the equations themselves are not explained in the text, and (b) the tables are utterly confusing leaving the reader guessing as to what most of these columns stand for. The tool should be understandable to any interested party who wants to reproduce it for water management purposes.

5. Line 1 in Page 2507 first mentions the "map-correlation" method, but does not explain how the cross correlation takes place between the ungauged site and the index stream gauge. For example, what specific characteristics are correlated, and what equations are used?

6. Line 23 in page 2513 refers to "leave-one-out" cross validation, but does not explain the rationale behind it. I suggest that it is explained at least in broad terms, since this is not a standard term that one comes across every day.

7. Line 14, page 2514 says ". . . . . from the Cruise tool at high streamflow values is more of a challeng. . . . . ." I am not sure how the difference between goodness of fit values for the transformed and untransformed streamflows explains that only high values (or both high and low values for that matter) are a challenge. Might not this difference be caused by discrepancies in mid-range values too?

8. In Figure 2, text on the top graph which reads as "Flow quantiles greater than 0.01" should read as "Flow quantiles less than 0.01" if I understand the text correctly. Figure 3 is extremely small and none of the screen shots are clearly visible. I think it is better to break this figure into 2 and expand the size of each screen-shot to have more clarity. All numbers and lettering in Figure 4 is too small to read, while the comparisons between observed and generated streamflows (graphs) are not at all visible to the naked eye. Howewer, I think the figure itself represents a neat way of summarizing the goodness of fit information, if it can be made larger and the signs for different efficiency ranges

GMDD

5, C824–C828, 2012

are made distinct from each other.

9. Line 15 on page 2510 refers to a “Microsoft Excel” spreadsheet. However, as far as I understand the spreadsheet doesn’t necessarily have to be a “Microsoft Excel” one. Any spreadsheet program with capability to run macros, or perhaps a standalone code to perform the underlying procedures may be used instead. Perhaps the authors need to qualify that they have currently used a “Microsoft Excel” spreadsheet (If this journal is okay with the use of brand names), but the same functionality may be obtained by other means.

10. Other comments of minor nature are:

(a) Use the word “often” instead of “often times” (b) Line 19, page 2506: use “characteristics of” instead of “characteristics computed for” (c) Word “recursively” is spelled wrong in Fig. 2 (d) Line 21, page 2509: typo “by published Smakhtin(1999)” to be corrected as “published by” (e) Line 25 page 2509: typo “on the same day as” to be corrected as “on the same day at” (f) Line 15, page 2512: “Fig.1” should perhaps be “Fig.4”? (g) Line 10, page 2514: the word “indicating” is spelled wrong.

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