Response to Anonymous Referee #2

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.

Thank you for your comments and review. We have revised the manuscript substantially to address these concerns, including adding new sections to the text, expanding existing sections and revising the figures to provide more detail and clarity. We address your general and specific comments related to these points 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 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 webbased software tool. . .. ."

We agree and revised the titled as, "Towards a publicly-available, map-based regional software tool to estimate unregulated daily streamflow at ungauged rivers."

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).

We agree this is not clear. We have added additional text to the methods section (Section 2) to describe how the methods can be applied to other regions as well as text as to how the tool can be viewed as a framework. Specifically, we add in Section 3:

“The software tool can be considered a general framework to provide daily streamflow time series at ungauged locations in other regions of the United States and possibly other areas. Furthermore, all data and methods underlying tool are freely available. Whereas the tool is a general framework for providing a map-based, “point-and-click” approach to estimate daily streamflow at an ungauged river location of interest, the underlying data, including the river network and catchment characteristics, are specific to the region of interest. Much like other modeling frameworks, the software tool must be calibrated based on the data available in the region of interest. Details of the functionality of the regional tool presented in this study follow. Additional details on the customization of the catchment delineation for application to other regions is discussed in Section 4.”

We also added a new Discussion section to the text with a paragraph to describe the underlying data and methods needed to develop a watershed delineation tool for other regions across the globe.

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.

We agree with this comment. We have substantially expanded the methods sections to provide equations and text for the regression equations and map correlation method used in the software tool. We have added a new section (now Section 3.2) to the text titled “Estimation of daily streamflow in the demonstration area.” This section describes how the methods introduced in Section 2 were implemented for the demonstration area. Figure 3 has been broken into two figures (now figures 3 and 4) and figure 4 (now figure 6) has been reworked so that the hydrographs are larger and more visible. Figure 1 has also been completely revised to show additional detail on the software methods.

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).*

We agree and changed this sentence in the abstract to read:

"For a demonstration region in the northeast United States, daily streamflow was, in general, shown to be reliably estimated by the software tool, with more difficulty estimating the highest and lowest streamflows that occurred over the historical period from 1960 through 2004."

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.

We agree. To address this comment, we have added more detail to this section on the regression approach, including how catchment characteristics are selected as well as present the general form of the regression equation. We also added the statement:

"In this approach, catchment characteristics (the independent variables) are regressed against the streamflow quantiles (the dependent variable) to determine which catchment characteristics have a statistically significant relation with each streamflow quantile. The catchment characteristics tested for inclusion in the regression equations are based on the availability of the spatial data layers in the particular study area of interest and, therefore, vary from region to region."

We also now added a new section (now Section 3.2), which details the specific methods used to develop the regression equations for the demonstration area. We discuss in this section the rationale for the specific basin characteristics used in the demonstration area and the details of the regression fits.

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 paper 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.*

We completely agree and address these two comments together. We have substantially expanded the text in Section 2.1 to include the general form of the regression equations and additional text to describe the regressions, included in this section is text such as:

"In this approach, catchment characteristics (the independent variables) are regressed against the streamflow quantiles (the dependent variable) to determine which catchment characteristics have a statistically significant relation with each streamflow quantile. The catchment characteristics tested for inclusion in the regression equations are based on the availability of the spatial data layers in the particular study area of interest and, therefore, vary from region to region."

In the new section 3.2, we describe the reasons for why certain catchment characteristics were selected for the demonstration region:

"Previous work in the southern portion of the study area by *Archfield et al.* [2010] showed that, from a larger set of 22 catchment characteristics, the contributing area to the streamgauge, percent of the contributing area with surficial sand and gravel deposits, and mean annual precipitation values for the contributing area are important variables in modeling streamflows at ungauged locations. For this reason, these characteristics were summarized for the study streamgauges and used in the streamflow estimation process."

We also broke out the lower portion of Figure 2 into a new figure (now Figure 5), which shows the relations between the streamflow quantiles for the high exceedence probabilities and combines table 3 and Figure 2. We simplified Table 2 and spelled out abbreviations in the column headings to make the table easier to read.

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?

We agree. We have substantially expanded this section and the description of the method, adding equations to the text and further justification for its use in the software tool.

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.

We agree. We have added a new paragraph to describe this validation approach:

"To evaluate the utility of the underlying methods to estimate unregulated, daily streamflow at ungauged locations, a leave-one-out cross validation for 31 study streamgauges (fig. 6) was applied in conjunction with the methods described in Sections 2 and 3.2. These 31 study streamgauges were selected because they have observed streamflow covering the entire 44-yr historical period of streamflow estimated by the CRUISE tool. In the leave-one-out cross validation, each of the 31 study streamgauges was assumed to be ungauged and removed from the methods described in Sections 2 and 3.2. The methods were then reapplied without inclusion of the removed site. Using the catchment characteristics of the removed site, daily streamflow was determined and compared to the observed streamflow data at the removed streamgauge. This cross-validation procedure ensured that the comparison of observed and estimated streamflow at each of the study streamgauges represented the truly ungauged case because the streamgauge was not used in any part of the methods development. This procedure was repeated for each of the 31 validation streamgauges to obtain 31 estimated and observed streamflow time series from which to assess the performance of the study methods."

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?

We agree that it is not clear why this is the case. We removed this statement from the manuscript.

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. However, 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 are made distinct from each other.

We agree with this comment and thank the reviewer for their suggestions. We made the suggested changes to the text on Figure 2. We also broke Figure 3 into two figures (now figures 3 and 4). We have now broken Figure 2 into two figures (Figures 2 and 5) – the first is a more general figure used in the methods section to describe how the flow-duration curve is estimated. The second figure (now figure 5) is referenced in the new section (Section 3.2) titled "Estimation of daily streamflow in the demonstration area," which uses the data from the demonstration area to show the relation between the streamflow quantiles for the demonstration area.

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.

We agree and added text to clarify that any macro-enabled spreadsheet program will suffice (see Section 3, end of paragraph 2):

"The spreadsheet itself, which contains the VBA source code, can be used independently of the StreamStats interface and is, therefore, able to be customized to interface with other watershed delineation tools or with any study area for which the methods in Section 2 have been applied. Additionally, any macro-enabled spreadsheet program could be used in place of the Microsoft Excel spreadsheet program."

We have also added the standard U.S. Government disclaimer in the acknowledgements sections:

"Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government."

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.

Thank for your thorough review and catching these errors. We have made all recommended changes.