Responses to Comments on “A Bayesian approach to understanding the key factors influencing temporal variability in stream water quality: a case study in the Great Barrier Reef catchments” (Referee #2)

Anonymous Referee #2 Received and published: 19 Feb 2021

Our responses are in blue and proposed manuscript revisions underlined.

General comment

Liu et al. conducted an improved Bayesian approach to evaluate the temporal variability in stream water quality and the related key factors. This study aimed to: i) identify the key influencing factors, and ii) predict the temporal variation, taking advantages of multiple locations and multiple water quality monitoring data. In addition, authors divided the study sites into two clusters and analyzed separately, which might avoid potential uncertainty issues caused by a single model, and improve the scientific and reliability of the modelling results.

This study is an interesting topic and generally well written. It contributes to our knowledge of both the further application of the developed Bayesian model framework and the understanding the temporal water quality variability in the Great Barrier Reef catchments. In general, this piece of work could be considered for publication after some unclear concerns were addressed.

Thank you for your comprehensive review and recognition of the study contribution. The constructive comments will help us improve our manuscript after revision. We provide detailed responses to your comments and our proposed manuscript revisions in the subsequent sections.

Major comment

1. Section 2.2.2 The authors gave a detailed process of data extraction and processing. Among them, it was noticed that “The start and end points of a specific event were determined by using a local minimum method that calculates the first derivative of the streamflow record (separated from baseflow)”. Basing on your data processing method, when can be identified as the start or end points? I think more details of the key standard or parameter maybe better for the readers to further understand your approach.

Thank you for this comment. We agree with the referee that it is a bit ambiguous as to how the event is delineated (i.e., definition of start and end points of events). Here we used an automated approach developed by Tang et al. (2017), which allows us to extract runoff event on the baseflow-free hydrograph, by specifying a set of parameters (e.g., *β* filter coefficient, *ReTh* difference between two flows to set the local minima for event extraction). As illustrated in Figure 1, once the local minimum T1 is found, the next local minimum T2 is considered as the first candidate end point. *ReTh* is used to filter out any false end point, which allows the flows at the start and end of an event can be different. This Matlab toolbox directly returns the start and end points of an event, avoid time-consuming and subjective inconsistent outcomes.

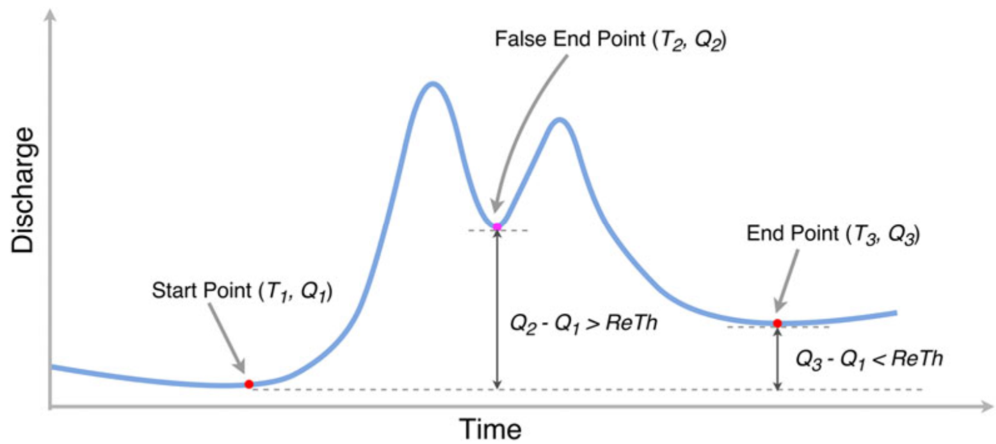

Figure 1. An example of selecting the end point of a runoff event (Tang et al., 2017).

<u>To resolve this comment, in the revised manuscript, we will:</u>

<u>1) clarify the method we used to delineate flow events in ***Sect. 2.2.2 (Event mean concentration)*** and provide the key specifications (i.e., parameters) for running the Hydrun toolbox in the ***Appendix***.</u>

<u>2) provide an example hydrograph with start and end points in the ***Appendix*** (as shown in the figure below).</u>

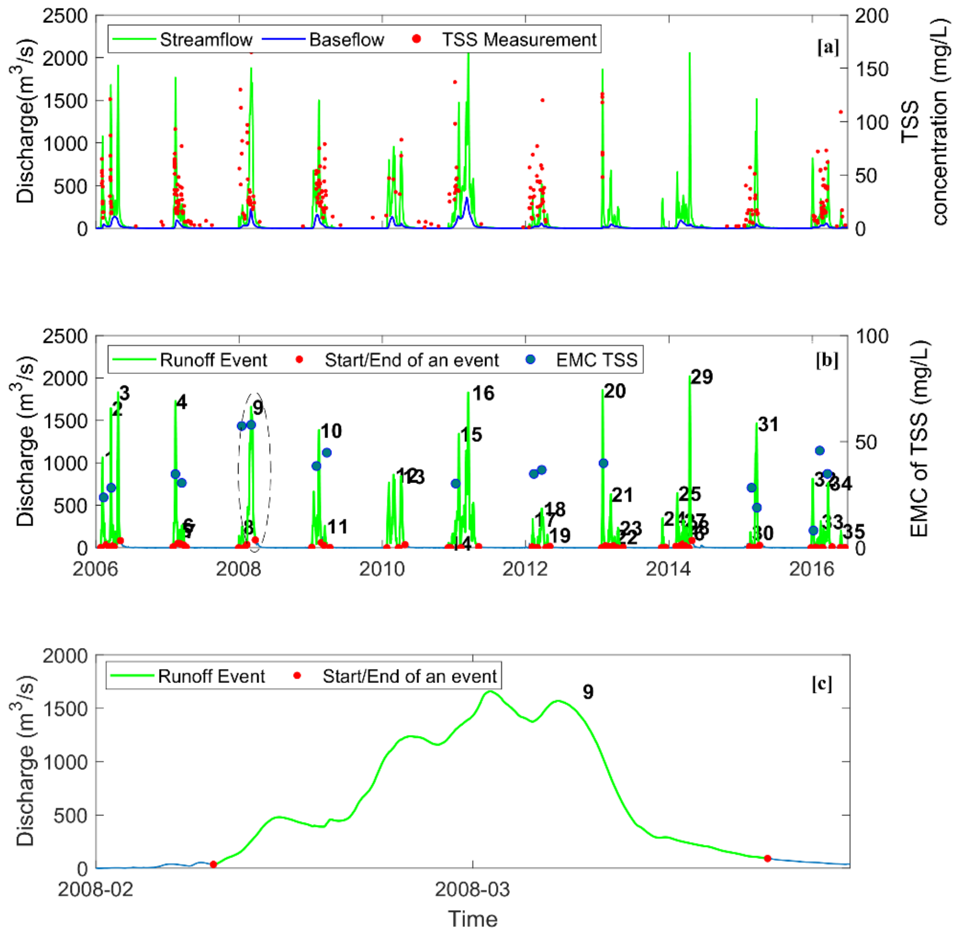

Figure 2. Delineation of runoff events and estimation of EMCs, based on the hydrograph for 105107A Normanby River at Kalpowar Crossing in the GBR catchments: (a) baseflow separation from continuous streamflow observations; (b) event identification and development of EMC, and 35 runoff events are identified with red dots representing either the start or end of a runoff event; and (c) A zoom in event #9 in 2008.

2. The authors divided the site locations of the GBR catchments into two clusters (wet and dry), and modelled separately. The advantages of the subsequent result are obvious, i.e., pertinence, reliability and so on. However, whether the strong pertinence will reduce the universality of this approach and limit its universal application? And if it is necessary to add the model and discussion of all sites?

Thank you for this comment. Application of the model based on two clusters does not limit the utility of the model. We aim to identify the key factors affecting temporal variability in water quality for two different clusters of sites, but this method can be used anywhere, e.g., a universal application for all sites. There are strong practical and conceptual reasons that we decide to model the two clusters of sites separately (i.e., we have provided detailed justifications in our reply to the Comment #3 from Referee #1). By doing this, we are not making claims that there are always variables that will be important in such catchments. Therefore, our method is universal, but our results are not.

To address this comment, in the revised manuscript, we propose to:

1) provide more details on our previous study on clustering of these catchments in ***Sect. 2.1 (Study area)***, and that differences in geographic/hydroclimatic are key factors that distinguished the two clusters of sites. Thus, there are strong practical merits in handling the clusters separately based on the clear contrast between them.

2) add discussion in ***Sect. 4.2 (Predicting temporal variations in water quality)***, to further clarify that applying our modelling approach to two clusters of sites does not limit the utility of the method.

3. The authors targeted nine common water quality indicators, including sediments, nutrients and salinity. But in the nutrients part, they only focused on N and P, without any constituents about Carbon studied. Why? Please explain it.

Water quality parameters included in the Queensland Government's Loads Monitoring Program are those that enter the Great Barrier Reef lagoon from inland catchments, and suspended solids, nutrients and pesticide are the focus of this program. Carbon from inland entering the GBR lagoon is not as harmful as sediments and nutrients for the coral reef ecosystem, thus carbon is not monitored. We could not include carbon in our analysis.

4. 2.2.2 again "The event-mean concentration (EMC) was then calculated for each event that had at least two samples on each of the rising and falling limbs of the hydrograph." Table C2 showed the Number of EMCs for each constituent. So what is the approximate amount of data per event? Why you set two samples as the minimum limitation? whether two samples are too few?

Thanks for your comment. First, we would like to highlight that we set the minimum of two on both rising and falling limbs of the hydrograph, then we have minimum 4 samples per event. Second, in Bartley et al. (2012), they review water quality data in Australian catchments and use minimum 3 samples over an event as threshold to calculate EMC (Tables 4 to 6, Bartley et al. (2012)). Therefore, our four samples per event is above the standard set by that review paper. In addition, we have calculated that, on average, there are 14 samples per event across nine constituents (ranging from 12 for DOP to 16 for EC), therefore our calculated EMCs are reliable.

To address this comment, we propose:

1) clarify the choice of four samples (2 on both rising and falling limbs) per event in ***Sect. 2.2.2 (Event mean concentration).***

2) provide a summary table (Table 1) that indicates number of samples per event for each constituent in the ***Appendix***.

Table 1. Average number of samples per event for each constituent

TSS	PN	NO_X	NH_4	DON	FRP	DOP	PP	EC
15	14	14	14	14	15	12	14	16

5. I also noticed that you normalized the data of each event first and then calculate the Event mean concentration. If this process is necessary?

Thank for your comment. We think the referee may misunderstood our method. We normalized the EMC rather than the original water quality data. We will revise our manuscript to clarify this in ***Sect. 2.2.2 Event mean concentration***. Also, normalization of the predictand is necessary to facilitate the fitting process and fulfill the statistical assumption of our model; we use Bayesian linear regression with the response variable sampled from a normal distribution (Atkinson, 2020; Castillo et al., 2015; Hoeting et al., 2002). We will incorporate in ***Sect. 2.2.2 (Event mean concentration)***.

Minor comments:

6. Fig 1d six_NRM regions.

Thank you for this comment. We will incorporate this suggestion in the revised manuscript.

7. Table 1 Delete the comma at the end of the sentence in the item “Land use/land cover” of Cluster 2.

Thank you for this comment. We will incorporate this suggestion in the revised manuscript.

8. L274 Delete the full name of “MCMC”, which has appeared in the line 261.

Thank you for this comment. We will incorporate this suggestion in the revised manuscript.

Reference

Atkinson, Anthony B. (2020). The box-cox transformation: Review and extensions. *Statistical science*.

Bartley, Rebecca, Speirs, William J, Ellis, Tim W, & Waters, David K. (2012). A review of sediment and nutrient concentration data from Australia for use in catchment water quality models. *Marine pollution bulletin, 65*(4-9), 101-116.

Castillo, Ismaël, Schmidt-Hieber, Johannes, & Van der Vaart, Aad. (2015). Bayesian linear regression with sparse priors. *Annals of Statistics, 43*(5), 1986-2018.

Hoeting, Jennifer A, Raftery, Adrian E, & Madigan, David. (2002). Bayesian variable and transformation selection in linear regression. *Journal of Computational and Graphical Statistics, 11*(3), 485-507.

Tang, Weigang, & Carey, Sean K. (2017). HydRun: A MATLAB toolbox for rainfall–runoff analysis. *Hydrological Processes, 31*(15), 2670-2682.