Hydrology and Earth System Sciences Discussions

Hydrol. Earth Syst. Sci. Discuss., referee comment RC2
https://doi.org/10.5194/hess-2021-353-RC2, 2021

Comment on hess-2021-353

Anonymous Referee #2

Referee comment on "Synthesizing the impacts of baseflow contribution on concentration–discharge (*C*–*Q*) relationships across Australia using a Bayesian hierarchical model" by Danlu Guo et al., Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2021-353-RC2, 2021

This manuscript presents a synthesis of baseflow effects on C-Q relationships in watersheds across Australia. The authors have leveraged a Bayesian Hierarchical Model in this research. Overall, I think the research is solid, the manuscript is well written, and it can become an important contribution to the literature on riverine C-Q relationships. I provide below some comments to the author, which I hope can help improve the manuscript.

1. The use of Bayesian Hierarchical Model should be more fully justified. I am aware of the research the team has done in the past few years involving Bayesian approaches, but why is it used in this work on C-Q relationships. Please provide your reasoning in the Introduction, probably the last paragraph.

2. The authors reported that the Bayesian Hierarchical Model can explain over half of the observed variability in concentration of TSS, EC and P species across all catchments (93% for EC, 63% for TP, 63% for SRP, and 60% for TSS). I feel the intension has switched here from understanding C-Q relationship to predicting water-quality concentrations, which seems to be a distraction to me. Moreover, what's the benefit of adopting the Bayesian Hierarchical Model, given that many statistical models (e.g., WRTDS) have been developed and can probably provide more accurate estimates?

3. Figure 3b: It is not a strictly positive relationship for the entire range of BFI_m. The variability continues to increase with BFI_m up to ~ 0.5 and then starts to decrease with BFI_m. The latter part of the curve seems largely ignored in the manuscript, including Discussion. The same observation holds true for the individual constituents (Figure S5).

4. It would be interesting to investigate the effects of seasons and antecedent discharge conditions (wet vs. dry), both of which may change the response of C-Q slope to the BFI_m metric. There may be strong contrast among, for example, growing vs. non-

growing seasons. Toward the end of manuscript, the authors have briefly pointed out the possibility of season effects. I think it is probably beyond your scope to look into these effects in this paper, but I encourage the authors to provide a brief discussion to point out that the response of C-Q slope to the BFI_m metric can vary among different seasons, among different antecedent discharge conditions, and even among different periods. In the latter regard, it is reported that anthropogenic disturbances and/or management actions occurred in the catchment can cause the C-Q relationship to change. For example, Zhang (2018) provides an investigation of C-Q relationship for different river flows and years: https://doi.org/10.1016/j.scitotenv.2017.09.221.

5. The term BFI_m (median BFI) is not self-evident in the Abstract. Given the importance of this metric, I encourage the authors to define it more clearly in the Abstract.

6. The authors have used BFI_l and BFI_h to represent the variability of BFI, which makes sense to me. I may have used 2.5% and 97.5% instead but 10% and 90% are fine. By the way, have you considered using standard deviation to capture the spread, which may help shorten the manuscript in terms of text and figures presented? I think an argument can be added to the end of Section 2.1, which favors the use of BFI_l and BFI_h, that these are percentile based and hence are more robust to outliers.

7. For days with multiple samples, is it necessary to pre-calculate the average concentration? Why not keeping all the samples in the analysis? In addition, it may be helpful to provide a table that quantifies the fraction of such days in the record.

8. BFI calculation: I am curious about the use of 0.98 for alpha in the baseflow filter. Did Ladson et al. (2013) recommend this value? What is the rationale?

9. Figure 2: Please add numbers and units (even if hypothetical) on the y-axis for panels a and b.

10. Equations 2-3: Consider changing ð□□¿ð□□µð□□¹ð□□¼ð□□□ð□□□ð□□□ð□□□ð□□□ð□□ið□□□ to ð□□¿BFI_ð□□□ð□□□ð□□□ð□□□ð□□□ð□□ið□□□. At first glance, I thought this is the product of two variables (ð□□¿ and BFIclimate).

11. Section 3.3.1, including Table 1: I would like to refer back to my comment above. The NSE values do not seem to be comparable to more established approaches such as WRTDS. What is the value of showing these results? Should the baseline model or the BFI-based model be used for predicting concentrations? Why not those other established approaches?

12. Section 3.3.2: According to published literature on many catchments around the world, SRP is a minor component of TP, whereas NOx is a major component of TN. It is quite interesting that in these Australian catchments, NOx/TN is quite small. This presents a strong contrast to many regions and may be discussed with a couple of sentences.