Biogeosciences
Discussions
EGU
Open Access

Interactive comment on "A comprehensive biogeochemical record and annual flux estimates for the Sabaki River (Kenya)" *by* Trent R. Marwick et al.

Anonymous Referee #1

Received and published: 7 September 2017

General comments:

This paper presents a two-year biogeochemical record (with biweekly sampling frequencies) of the Sabaki River and A-G-S river basin in Kenya. The authors seek to provide initial baseline data given the expected changes to the AGS river basin, such as: the growing contiguous population in Nairobi with inadequate sewage facilities, the anticipated increase in dissolved nutrient export from African river basins, and the planned damming of a river within the basin. While establishing this baseline is a critical need and the data collected for this effort is impressive in scope, the paper could greatly benefit from better organization around hypotheses and re-focusing based on

the data and statistical tools needed to test these hypotheses. Three general suggestions are highlighted below:

1. Hypothesis and purpose unclear; the comprehensive nature of the paper obscures the message. The authors are encouraged to identify a story (or stories) they can tell with these data and keep to that purpose. One potential action is to split into multiple papers if data support multiple, novel stand-alone documents. Once clear hypotheses are formulated, the paper should be trimmed to focus on the objective(s).

2. Novelty; the comprehensive past and current collaborations in the basin both benefit and detract from the strength of this manuscript. It is often unclear what is new vs. repackaged (CH4, N2O) vs. re-sampled (sediment fluxes, Marwick et al. 2014a) from previous publications, which muddies the novelty & distinct advancements made by this paper. The authors are encouraged to better highlight what is new.

3. Analysis/Statistical tests; the paper is lacking quantitative analyses and acknowledgements of uncertainty. Stats should follow hypotheses to test correlations between key parameters or multivariate models of interest. Time series analysis may be used to address time-varying correlations or controls on different biogeochemical fluxes. Further, if the authors wish to make quantitative comparisons of fluxes between different studies or solutes, they must acknowledge the uncertainty of their estimates (are the differences significant or just different within a similar range of uncertainty)?

Specific comments:

— Introduction - Starts out very C-focused. Overall - the introduction does not capture the objective(s) of the paper. Since a significant portion of the introduction is focused on C cycling/dynamics and metabolism, the reader is lead to believe that those topics will be a major focus of the manuscript. — Objectives of study - Not defined in the introduction, but clearly stated on pages 8 and 9, lines 12 and 1-3, respectively. Moving this section to the introduction (or re-wording it to fit into the introduction) would give the reader a much better understanding of the purpose. — What was the reasoning for

excluding CO2 data but including CH4 and N2O (e.g., Borges 2015a)? Failure to include CO2, DO, and metabolism data is a missed opportunity if this is to be a key focus of the paper. — Discharge data: gap-filling - the gaps and potential consequences of gaps in the discharge data must be addressed before making comparisons with other flux estimates. Given a 2 month period of no measurement - how off might the authors' estimates of missing Q from past years be? Have the authors tested the robustness of their gap-filling approach with other months that were not missing from the sampling period? — Discharge data: rating curve from gauge height - Discharge during much of the study period was well below and above the 2 clusters of points used to derive the rating curve (Fig 2a-b). What certainty do the authors have in making these interpolations and extrapolations from the 2 clusters and of the flux estimate comparisons that follow (e.g., Table 1, yields on p20)? — Nairobi - Referenced throughout text with little preface as to the location of the city in relation to the study area. "Nairobi" also appears to be used in place of urban influence (see page 5, line 15). Include clear explanation that Nairobi is the dominant "urban" influence in the study system. Introduction would benefit from additional literature supporting claim of anthropogenic influence on quantities of lateral nutrient inputs (if Nairobi or flow-regulated objectives become the primary) or whatever hypotheses the authors choose to test/focus on. — Many run-on sentences make key points difficult to follow (e.g., Page 3: 7-14, Page 21: 3-8, 8-14.)

Technical comments (noted by Page:Line):

Title: the title does not adequately capture the full scope of the paper. The title only mentions the Sabaki, yet the paper broadens its study site to the Athi-Galana-Sabaki (AGS) basin.

Abstract: the abstract was heavy in numbers, and read too much like a results section. It would benefit from more conceptual information.

2: 23: Consistency of "dammed" throughout is preferred (versus alternating with "flow-

regulated" when referring to dammed rivers).

3: 3-5: Here and throughout - try to stick to ∼3 key references to make a point. Long lists are not helpful, and especially not needed if after "e.g.".

3: 5-7: Consider removing 'advancing to...global C cycle'

3: 9-11: Consider rewording "derived from heterotrophic metabolism…." in simpler terms

3: 21: Instead of "earth system domain", perhaps use biosphere?

3: 27: It is not clear why these regions would be more significant until later in the text. Reorganize and reorder.

6: 5: Figure "d" is the crop corrected vegetation, not "c"

7: 2: Here and throughout: don't need to define as physicochemical AND biogeochemistry unless the authors re-analyze data to include more processes or reactions (i.e., biogeochemistry). Otherwise delete biogeochemistry.

7: 4: What frequency were these temp, conductivity, O2, and pH data collected? These may be an interesting times series all to themselves….

7: 11: Was 2000mL of water collected at each sample or during the entire course of study?

8: 9: change 'period was provided' to 'period were provided'

8: 16-21: Perhaps this would be better suited in a discussion section than methods?

9: 15-16: Should the nutrient data collection time frame be mentioned here with discharge, or later with the nutrient data information?

11: Throughout section 3.3 - watch out for overuse of terms like "complex patterns", "complex variability", "no strong seasonal pattern", "erratic pattern", "complex variation", "highly variable", etc. These become overwhelming and the manuscript would be

much improved if they were removed and replaced with statistics.

12: 13: Here and throughout - consider re-arranging results to include fraction names with values instead of listing in separate clauses. For example: "4.0 Tg TSM yr-1, 70.6 Gg C-POC yr-1, and 24.1 Gg C-DOC yr-1."

19: 13: Opportunity to illustrate how a coarser sampling schedule may yield these differences in flux estimates: what would the authors conclude from this bi-weekly dataset if they trimmed it to the frequency of previous budget sampling intervals? Same difference or different results entirely?

21: While a very interesting side-note, this discussion using isotope values (but NOT mixing models or other quantitative tools) is a diversion from this paper as currently organized and seems better suited for a short note of its own.

Figure 1: Would be very helpful to show S19, S20, other key sampling sites - perhaps directly labeled in panel a.

Figure 6: Keep y-axis titles on the same side