7/27/2022

Dr. Isaac Santos
Global Biogeochemical Cycles

Dr. Santos,

We thank you for the reviews for our paper entitled "Remotely sensing river greenhouse gas exchange velocity using the SWOT satellite" [Paper # 2022GB007419]. We are delighted to receive comments such as "*this manuscript makes a substantial advance in scaling gas exchange"* and that it is *"timely … in the context of the upcoming SWOT mission*". We thank both reviewers for such a careful reading of both the main text and supporting information of this manuscript.

We agree with and appreciate the critiques of this paper, and accordingly have accepted nearly all reviewer comments or explained how our lack of clarity led to a comment that is not applicable. In those cases, we have edited the text to ensure our original intent is clearly communicated.

The primary changes to the manuscript include:

1) Addressing reviewer 1's comments about the statistical choices in the gas exchange model: we now additionally test a suite of log-transformed implementations of the model and ultimately implement a Bayesian regression model informed by both theory and data.
2) Incorporating Reviewer 1 and editor's comments about restructuring certain sections for clarity; most significantly, we streamline the presentation of our results and discussion.
3) Addressing Reviewer 2's comments on algorithm limitations and the global representativeness of the validation data: we add clarification paragraphs throughout the discussion that make clear the limitations and biases throughout our analysis.

In our response, we have grouped together similar reviewer comments when appropriate and give original reviewer text in regular typeface. Our responses are italicized, with tangible changes to the manuscript highlighted in dark grey. Underlined text is new, while strikethrough text is deleted. If you have any further questions, please do not hesitate to contact me.

On behalf of all authors,

Craig Brinkerhoff
Department of Civil & Environmental Engineering
University of Massachusetts, Amherst
cbrinkerhoff@umass.edu

Editor’s remarks

We thank the authors for their patience during the review process. We have now received two reviews of the manuscript. Both reviews were generally quite positive, highlighting the utility and substance of the BIKER model in advancing the scaling of atmospheric gas exchange from rivers. I agree with these perspectives that this work represents an important advancement. However, both reviewers have raised a number of criticisms that need to be addressed before this manuscript can be considered for publication. Specifically, reviewer 1 raises a number of critical questions regarding the statistical choices and justifications made by the authors and how these choices affect the performance and output of the model. Careful responses are needed to each of these points. Reviewer 2 additionally raises some more practical questions around the usability of such a model given resolution (50m) and geographic constraints. Moreover, I strongly agree with reviewer 1 that this paper is rather long and difficult to parse in many places. Their advice to restructure certain sections, particularly the results, is solid. Lastly, both reviewers raise a number of specific line-by-line comments that should be addressed as well.

Reviewer 1's remarks

This manuscript makes a substantial advance in scaling gas exchange. Most work has assumed that gas exchange is stable with place, but this one demonstrates how remote sensing of river parameters coupled with a statistical model can provide highly time resolved estimates of gas exchange. The work is well founded in theory and extends this theory to accurately (or sometimes not) scale gas exchange to many rivers through time. Lastly all the code and data are available for anyone to replicate this work.

Despite the above enthusiasm I have several comments to improve this work.

Main Comments:

1. The initial model of gas exchange is one where the authors estimated based on theory what the parameters ought to be (save an intercept term) and they showed that these prediction fit the actual data well. Except that I disagree with this point, the fit looks like it does not fit all that well. It over predicts at high k and under predicts at low k. My questions:

   a. How was this model fitted to estimate \beta_1? Least squares? If so this method of fitting implicitly assume a normal distributed likelihood, but on unlogged data, the error term is probably varies in this model. I would fit on a logged version of equation 7. The authors show the logged data and indeed these data have homogenous variance suggesting use of unlogged data would have non constant variance. Indeed the authors did not say with the stochastic part of this model is. I see no reason why the multiplicative model in eq 7 cannot be an additive model of logged data

   b. I like the theoretical justification for the parameter estimates. But why adhere so strongly to these and not simply let the data pick the parameters? The goal here is to predict gas exchange and not put the theoretical cart before the horse. Set these parameters free and there might be a much better fit of the data to the model. Given the authors expertise in Bayesian approaches, they could even put tight priors on these parameters centered on the theoretical estimates. Thus if the data say that the parameter is in fact not = 9/16 (which is to say 9/16 lies outside the posterior of the parameter estimate), then they have an interesting finding and a better predicting model. If the authors are worried about scaling the parameter uncertainty for several parameters vs one, I suggest it is not that difficult, just use the joint distribution which is to say the group of parameters at each step on the MCMC.

   L247 I disagrees that this model accurately captures, its accuracy depends on the value of k.
   L248 Given all of the Bayesian reasoning in the paper (which I strongly approve of!) it seems strange to see parameter estimate, \beta presented without any nod to the fact that this parameter is really a probability distribution and has some uncertainty interval.

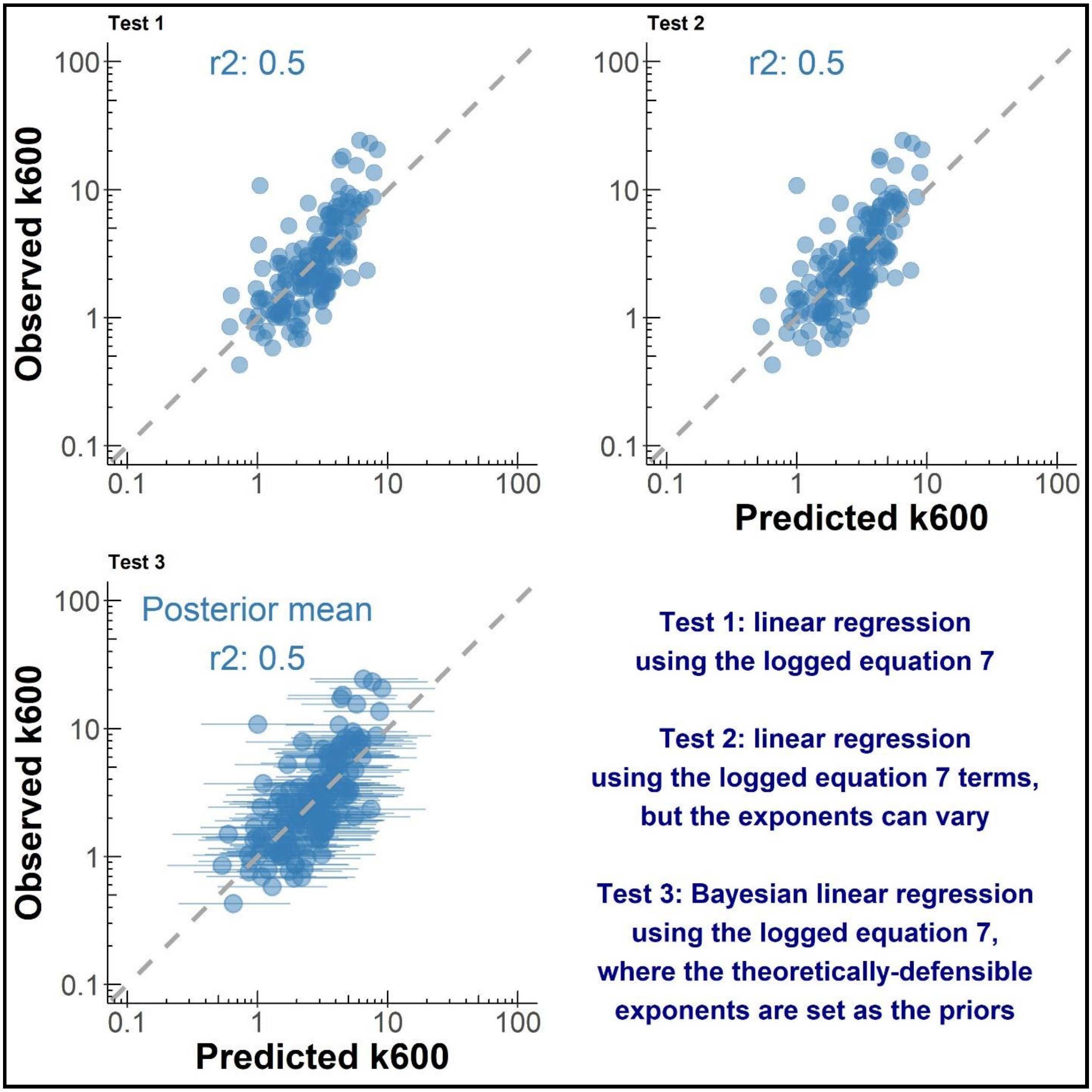
We thank the reviewer for these very helpful comments regarding a core component of BIKER. For convenience throughout this response, equation 7 (from the original version of the manuscript) is reprinted below.

$$k_{600} = \beta_1 (gS)^{\frac{7}{16}} \bar{U}^{\frac{1}{4}} H^{\frac{9}{16}} \qquad \textbf{(7)}$$

The reviewer is correct that least squares regression was used to estimate β_1 and that this is problematic if fit to data that does not exhibit homogenous variance (like these data in natural [non log] space). This is confirmed when assessing the Q-Q and homoscedasticity plots for this model. The reviewer is additionally correct to raise concerns about two things: 1) β_1 is presented with no regard for β_1 being a distribution itself and 2) the equation 7 coefficients are not informed by the data. Regarding the latter, we chose to favor theoretically defensible coefficients because of the limited data available to us: we did not want to over-fit the equation when derivable parameters specific to SWOT-observable rivers are also possible. However, as the reviewer stated, it is also worth incorporating what the data show via Bayesian tools, which we have tested here at the reviewer's suggestion.

We tested three additional regression models based on equation 7, as suggested by the reviewer. The results of these three tests follow below.

1) *An additive linear regression model using the logged equation 7 coefficients. This is simply the log-transformed form of equation 7 from the manuscript.*
    a. $logk_{600} = \beta_1 + \left(\frac{7}{16}\right)\log(gS) + \left(\frac{1}{4}\right)\log(\overline{U}) + \left(\frac{9}{16}\right)\log(\mathrm{H})$
    b. *In this setup, β_1 via least-squares simply equals the average residual for the above relation when fit to data.*
2) *An additive linear regression model using the logged equation 7 variables, but the coefficients can vary.*
    a. $logk_{600} = \beta_1 + (\alpha_1)\log(gS) + (\alpha_2)\log(\overline{U}) + (\alpha_3)\log(\mathrm{H})$
3) *An additive Bayesian linear regression model using the logged equation 7, where the theoretically-defensible coefficients are set as informative priors and the intercept and model uncertainty have weakly-informative priors. Posterior r^2 is calculated following Gelman et al. (2019).*
    a. $logk_{600} = \beta_1 + (\alpha_1)\log(gS) + (\alpha_2)\log(\overline{U}) + (\alpha_3)\log(\mathrm{H})$
    b. *Priors:*
        i. $\alpha_1 \sim \mathrm{N}(\frac{7}{16}, \frac{1}{8}^2)$
        ii. $\alpha_2 \sim \mathrm{N}(\frac{1}{4}, \frac{1}{8}^2)$
        iii. $\alpha_3 \sim \mathrm{N}(\frac{9}{16}, \frac{1}{8}^2)$
        iv. $\beta_1 \sim N(0, 1)$
        v. $\sigma_{LM} \sim \exp(1)$



As the reviewer anticipated, model fit is not as good as we declared in the manuscript, with r^2 of 0.50 across all three models. There is near identical model fit across all three tests, with the additive model in test 2 yielding statistical coefficients that are virtually identical to those derived from our geomorphic assumptions in test 1 (see the table below). This confirms our initial choice to use theoretically-defensible coefficients. This also means that the Bayesian model shifts very little from prior to posterior.

	Test 1	Test 2	Test 3 (posterior means)
α_1	*7/16 ≈0.44*	*0.42*	*0.43*
α_2	*1/4 ≈ 0.25*	*0.32*	*0.31*

α_3	*9/16 ≈ 0.56*	*0.50*	*0.52*
β_1 (natural log space)	*3.85*	*3.85*	*3.89*

After performing these tests, we have chosen to implement the Bayesian regression model (test 3) in the manuscript and BIKER. Aside from an explicit accounting for both our prior knowledge and what the data suggests, this model also conveniently infers the uncertainty for model k_{600} (σ_{LM}), which is a required input parameter for BIKER. Previously, we simply set a reasonable value (0.30 in natural log space). Now, we can explicitly estimate this parameter using the Test 3 model, obtaining a posterior mean log(k_{600}) uncertainty of 0.58.

We again thank the reviewer for suggesting a revisit to these model choices, as we believe this updated model to be a significantly better argument on both theoretical and empirical grounds.

Changes to the manuscript include the following
1) *Updated Equation 7, Figure 2, Figure S1, and Text S2 using the new Bayesian regression model*
2) *Added Text S3 to explain regression hyperparameterization*
3) *Added the above table to the manuscript's main text (Table 1). Equation 8 was also added to show the final fitted posterior model.*
4) *Updated Section 2.4 results and discussion when necessary*
5) *Updated all BIKER validation results presented in figures and text*
6) *Updated Text S5 to reflect an updated 'complete BIKER uncertainty' using the posterior parameter distributions*
7) *Updated all references to β_1 to make clear it is a distribution and not a discrete estimate.*

2. I am bit confused by the likelihood equation (10). i.e., I assume it is a likelihood in which guess it would read as " the data are normally distributed with a mean equal to f(\theta, X) and a standard deviation \sigma{k600}, but the way the different parameters fall on different sides of the ~ are not clearly evident to me, I request a more detailed presentation of the steps leading up to this equation. I see this is in the SI, great, I suggest it show in the paper. The table of RMSE etc can go in the SI in its place.
L386 It strikes me that the math for metrics can go in the appendix since this is stuff easily looked up on Wikipedia.

We thank the reviewer for highlighting this confusion, which was made worse by our previous presentation of equation 7 using unlogged data. We have added the necessary clarification text and equations to arrive at what is now equation 12 in the manuscript's main text (and subsequently moved the metric definitions to the supplemental text as suggested). The relevant section is reprinted below:

The heart of BIKER is its reformulation of the k_{600} model (Equation 8) as a Bayesian sampling model that is conditional on the non-remotely-sensed parameters (i.e. $f(x|\Theta)$). This approach is similar to the 'McFLI' (Mass-Conserved Flow Law Inversion) logic used in some SWOT remote sensing of discharge algorithms (Gleason et al., 2017). To start, we write k_{600} as a function of SWOT-observables W and S. This algebra is carried out using Equation 8, ~~the fitted value for $\beta 1$ from Figure 2 (62.82),~~ and Manning's equation for mean flow velocity ($\bar{U} = \frac{1}{n} R_h^{2/3} S^{1/2}$).

Following Section 2.3, we continue to assume that the channel is hydraulically-wide ($R_h = H = \frac{A}{W}$). To leverage additional SWOT data, we use the "Durand transform" originally published by Durand et al. (2014): the wetted channel area A is further split into the ~~the~~ SWOT-unobservable portion A_0 and SWOT-observable portion dA ~~following Durand et al. (2014) and Hagemann et al. (2017)~~ where $dA_{it} = \sum_{t':W_{t'} \leq W_t} W_{t'} \delta H_{e_{t'}}$, for cross-section i and timestep t within a mass-conserved river reach.

All of ~~this~~ the above algebra simplifies to Equation 10. Conveniently, k_{600} as ~~measured~~ estimated by tracer additions to a stream is inherently a reach-scale quantity (in a mass-conserved reach). Therefore, Equations 7, 8, and 10 all yield a reach-scale k_{600} (i.e. $k_{600_i} = k_{600} \forall i$), thus ~~This~~ lowering the number of parameters BIKER must infer and making the problem much better constrained.

$$log(k_{600_t}) = 3.89 + (0.4320)log(g) + (0.5862)log(S_{i,t}) + (0.3084)log(\frac{1}{n_i}) + (0.7282)log\left(\frac{A_{0_i} + dA_{it}}{W_{it}}\right) \quad \textbf{(10)}$$

Next, Equation 10 is re-written as a Bayesian sampling model $f(x|\Theta)$, in which the joint data distribution x (i.e. SWOT observations) is sampled from the joint parameter distribution (θ). We first rearrange equation 10 to isolate x from θ (equation 11). Then, equation 12a-c re-expresses equation 11 as a normal distribution with standard deviation $\sigma_{k_{600}}$. The uncertainty expressed by $\sigma_{k_{600}}$ arises from uncertainties in 1) parameter~~s uncertainty~~ in Equation 8, 2) Manning's equation, and 3) the hydraulically wide channel assumption.

$$3.89 + (0.4320)\,log(g) + (0.5862)log(S_{i,t}) - (0.7282)log(W_{i,t}) = log(k_{600_t}) + (0.3084)log(n_i) - (0.7282)log(A_{0_i} + dA_{i,t}) \quad \textbf{(11)}$$

$$x = [3.89 + (0.4320)\,log(g) + (0.5862)log(S_{i,t}) - (0.7282)log(W_{i,t})] \quad \textbf{(12a)}$$
$$\Theta = [log(k_{600_t}) + (0.3084)log(n_i) - (0.7282)log(A_{0_i} + dA_{i,t})] \quad \textbf{(12b)}$$
$$x \sim N(\Theta, \sigma_{k_{600}}^2) \quad \textbf{(12c)}$$

3. This paper is quite long and the writing often impedes understanding. In particular paragraphs in the results section begin with instructions on where to find the data (e.g lines 422, 440, 452). This approach implicitly tells the reader "Go to figure 5 and guess for yourself what the finding is". Help the reader out. Each paragraph in the results section should begin with a statement of finding that summarizes the entire paragraph. Then describe how the data support that finding and cite the figures parenthetically. Some of what is written in the results belongs in figure captions (454-456, which I see are in the caption. Why copy a caption into the text?)

   a. Throughout the discussion, paragraphs lead with figures as the main actors and not the science. "Fig 7 confirms that...". Recast to state just the finding and relegate the contribution of the figure to this finding to some parentheses.

b. The paper has a lot of meta discourse explaining where it will go next, ok I see the point in a complicated paper such as this one, but I would try to find way to write so that the reader better understands the plan and does not need to be reminded at the top of each section. Maybe make the first figure much more detailed so that it describes the workflow in the paper.

c. "This" stands in for an unknown noun in dozens of places in the ms forcing the readers to guess at the authors meaning of "this". Make sure a noun follows every instance of "this..." in the manuscript

We thank the reviewer for such a close reading of the text and the helpful suggestions for improving the writing in the manuscript. We found this manuscript particularly cumbersome to write in our attempt to publish both theory and method together rather than salami slicing into two papers and appreciate your suggestions to streamline it. We approve of all small changes to writing style that the reviewer suggested (see below under 'Small Comments'). In response to the reviewer's broader suggestions, we have made the following changes and updates to the manuscript:

1) *Recasting paragraphs in the results and discussion sections to show the reader the results and conclusions, rather than send them to the figures.*
2) *Remove redundant text from the results section that mimics the figure captions.*
3) *Ensure specific nouns follow every instance of "this"*
4) *Add a workflow flowchart to the supplemental text (Figure S1), and point the reader there to understand the workflow of the entire paper*

4. The authors have included all of the data and have the code clearly laid out and in fact as an R package. I applaud this level of openness and detail with the analyses.
We thank the reviewer for this comment, and wholeheartedly agree!

Specific comments:

L119 ok pun here!
Further evidence of a very close read!

L121 What unknown exists that this paper fills?
The reviewer is right to point out that we never explicitly addressed a specific unknown in the current literature. To amend this, we have expanded the sentence to the following:

In this context, we borrow tools from fluvial geomorphology and existing SWOT algorithms to ~~answer the following two questions~~ begin addressing the current knowledge gap in the spatiotemporal dynamics of gas exchange velocity. More specifically, we seek to answer the following two questions:

L159 of empirically estimated k using gas tracer experiments. As an aside, I realize these tracer experiments are all we have, but who knows how much error exists in any one k estimate. Having done many of such experiments myself, I suggest a lot.
We wholly agree with the reviewer on this point!

L504 Why the surprise here?

We thank the reviewer for flagging this unfocused writing on our part. The global-scope scaling equations (used to calculate the CO_2 flux) harbor significant parameter uncertainties. Therefore, it is hardly surprising that BIKER + SWOT data, which directly captures two dimensions of the river's hydraulic geometry (width and slope), can yield more accurate results than global-scope scaling relations. With that said, the reliance on hydraulic geometry within gas exchange models is still somewhat surprising to many in the SWOT hydrology community who are less familiar with these relations.

Accordingly, we have removed the "surprisingly" from the manuscript.

L520 Summarize the main points here. I suggest that the text in the current "conclusions" would work well

L623 This is not really a conclusion but rather a restatement of the justification for the study and summary of the main findings. The summary can go as the first paragraph of the discussion and reserve the conclusion for the big message of the study and to look forward. I use this paper for guidance here. https://doi.org/10.1371/journal.pcbi.1005619

We agree with the reviewer's suggestions and have added the following key discussion points to the beginning of Section 4:

In this paper, we propose that the soon-to-launch SWOT satellite will provide enough hydraulic measurements to analyze the temporal dynamics of k_{600}, and therefore allow for a global-scale analysis of spatiotemporal trends in large-river k_{600} once SWOT launches. In preparation for SWOT's launch, we developed 1) a wide-river-specific hydraulic model for k_{600} that explains 50% of variation in k_{600} and 2) the BIKER algorithm to infer k_{600} using no on-the-ground information. Validating on 47 SWOT-simulated rivers, we show strong recovery of rivers' temporal k_{600} dynamics and a hypothetical total annual carbon emission rate across all 47 rivers (section 3.3).

The conclusion now reads:

This proof-of-concept study verifies that BIKER can provide meaningful information on the spatiotemporal dynamics of gas exchange solely from SWOT data and functionally opens the door for a global-scale analysis of riverine gas exchange upon SWOT's launch (and data collection). Although BIKER results are often biased in magnitude, they strongly capture the temporal dynamics of gas exchange velocity and will provide an unprecedented amount of new information on global riverine gas exchange that should be essential for better constraining existing river $[CO_2]$ models.

L586 Figures cannot confirm, Recast

We have recast this statement to better reflect the implied usefulness of BIKER, rather than declaring it to be true.

~~Figure 7 confirms that~~ BIKER will likely be useful for ~~is quite successful at predicting~~ informing annual upscaled carbon emissions estimates from ~~the~~ river networks when coupled with $[CO_2]$ data (Figure 7).

L601 If one is going through all the trouble to put out CO2 sensors to estimate CO2 flux, then I would recommend also adding O2 sensors (they are much easier to use) and then using diel

excursion of O2 to model gas exchange rates for that reach at a daily time step for as long as the sensor is out there (Appling et al. 2018 JGR). I suggest that if conditions are right for doing so (enough GPP, low k, which is implicit in a wide river-Bernhardt et al 2022), then one will get more accurate estimates of daily k than what scaling can do. I see the beauty of the scaling here for assigning estimates of k to unvisited/unmonitored rivers.
This is a valid point raised by the reviewer, and one that we agree makes sense in an in-situ scenario where one is likely already installing O_2 sensors.

Thus, we have removed this paragraph from the main text.

Small Comments:

L20 "Two decades of research has shown"
L26 State what is unknown in this abstract
L26 unique means "only one exists". Recast
L39 natural systems? Just say rivers.
L52 Write an equation here and not use prose. Is difference y-x or x-y? Only an equation can make that point clear.
L54 equations have compound variables like CO_2 which could be C times O_2
L54 use capital \Delta and not lower case \delta to describe change in something.
L63 One cannot measure k, there is no such instrument, thus we estimate k
L67 delete "there have been over"
L82 move the eq number away from the eqn so it doesn't look like you are multiplying by 2.
L90 achieve what?
L123, 125 For sure the answer is yes, even if small, rephrase to questions with quantitative answers.
L199 non-trival = difficult. I recommend a read through of this paper to remove such useless jargon
L212 following 199 substitute rivers for systems
Fig 2 No need for big bold axis titles. The abbrev. for "day" is d not dy. Variable should be italicized.
L364 delete "is unique in that it". There are other ways to derive daily timeseries of k for a river, O2 timeseries for example, so this method is not unique.
L415 \mu atm
L422 State the findings in the topic sentence of paragraphs, not where to go look for the data.
L447 ; however,
L452 A figure caption standing in for results text here.
L458 Here is the finding buried in the middle of the paragraph
L489 One can only support a hypothesis, not confirm it.
L493 But there was a "finally" that started line 470, and now here a second one?
L514 Use similar, not comparable. Everything is comparable to some extent.
L518 If Bayes, call them credible intervals
L526 systems? meaning rivers?
L527-531 Results in discussion
L537 it's is a contraction for it is.
L551 what is left to future work?
L560 I agree with this paragraph
L568 Figure numbers again standing in for insight.

*We agree with all of the reviewer's suggestions and have made the appropriate changes to the manuscript, save for "**L441** was. Past tense for all results." We prefer to use present tense for results but defer final editorial decision to the editors and typesetters.*

Reviewer 2's remarks

In the submitted manuscript, 'Remotely sensing river greenhouse gas exchange velocity using the SWOT satellite', Brinkerhoff et al., established the BIKER algorithm, a Bayesian inference of gas exchange, that allows for CO2 flux estimates from measurements of river surface width and slope. This study is timely as BIKER was developed in the context of the upcoming SWOT mission and will allow for SWOT data to be used to determine spatiotemporal dynamics of riverine gas exchange. Capturing spatiotemporal dynamics of gas exchange across varying watersheds, and in particular in larger rivers, is limited, despite being a key term in determining greenhouse gas exchange from inland waters. Thus, if BIKER can be used efficiently when SWOT begins collecting data, a step towards reducing uncertainty in greenhouse gas exchange estimates from inland waters will be achieved. However, there are few points on the limitations in the use of BIKER that need to be addressed/clarified before its implementation.

Main comments:

1) Since SWOT is limited to 50m width, this should be put into context, i.e. what % of the Earth's river/streams will be covered at 50m width?

The reviewer is absolutely correct that BIKER (and SWOT more generally) is fundamentally limited by the spatial resolution of the satellite measurements. We therefore follow the reviewer's suggestion to contextualize the SWOT river network within the global river network. To do so, we compared SWOT hydrography's mean annual surface area and network length (Altenau et al. 2021) with the most current estimates of the total global mean annual river surface area and network length (Liu et al 2022).

	Surface Area [km^2]	**Length [km]**
SWOT River Network (Altenau et al. 2021)	343,483	1,417,667
Most recent global hydrography estimate Liu et al (2022)	811,000	443,509,286
% that SWOT will observe	42%	0.3%

SWOT will directly observe much of the global river surface area (42% observable) but will not observe the vast majority of the network (0.3% observable). This is simply due to the scaling properties of river networks, where small streams (not SWOT-observable) constitute the vast majority of the network. While small streams in aggregate exert a significant influence on GHG emissions from river networks (Liu et al 2022; Raymond et al 2013), BIKER will be capable of inferring k_{600} for nearly half of the global freshwater air/water interface and provide a significant amount of new information for global river biogeochemical models. In that context, we have added this table to the supplemental text (Tabel S4) and introduce it in section 4.2 as the following:

However, SWOT's relatively coarse spatial resolution limits BIKER's use to large rivers. This limitation means SWOT cannot see the vast majority of the global river network (which are too narrow for SWOT), though it is likely to observe much of its air/water interface at which gas exchange occurs (rivers wide enough for SWOT to observe). To confirm this hypothesis, we *obtained the global estimates for SWOT-observable surface area and length (at mean annual streamflow- Altenau et al 2021) and compared them to the most recent estimates of global river surface area and length (Liu et al. 2022- Table S4). We found that 42% of the global riverine surface area is SWOT-observable, while only 0.3% of the network length is SWOT-observable. While small streams in aggregate exert a significant influence on GHG emissions from river networks (Liu et al 2022; Raymond et al 2013), BIKER will still be capable of inferring k_{600} for much of the global freshwater air/water interface.*

2) How does the model factor in rivers with large seasonally dynamic river widths (e.g. northern ice melt period)? Based on the "test data" used, coverage is limited in certain regions, e.g. Arctic which experience dramatic shifts in river hydrology during the ice-melt period. Does the BIKER algorithm capture rivers that experience extreme events?

We thank the reviewer for this comment. Given its direct reliance on SWOT hydraulics data, BIKER has no problem capturing large seasonally dynamic rivers, and one of the exciting opportunities for BIKER is quantifying how much GHG fluxes vary in the large, seasonally dynamic rivers in the Arctic. However, while SWOT will provide river widths and water surface slopes during overbank events, the physics underpinning BIKER do not theoretically allow for the inversion of flood events. Using BIKER across a full hydrograph including floods would necessitate the identification (via a heuristic such as the river width return period of 2 years) and removal of overbank events from the SWOT timeseries. This is true for many of the SWOT discharge algorithms as well. To clarify this in the main text, we add the following to section 4.2:

Finally, because of its reliance on Manning's equation and hydraulic geometry (section 3.2.1.), BIKER cannot invert overbank flow events, similar to many SWOT discharge algorithms. This concept is an important distinction that must be accounted for when BIKER is run on actual SWOT data, though future work should also look to couple floodplain flow laws with BIKER to capture gas exchange in seasonally-inundated floodplains.

3) The validation (Section 3) is based on a limited sample size (47 SWOT simulated rivers) and this sample size is further reduced to infer biases in variation (e.g. Figure 6). I understand the limitation in available data but the limited sample size and the implications it has on the output needs to be addressed further in the discussion.

L279-282 The geographic bias in the data (as shown in Figure S2) should be addressed here. Most points are in North America, with none covering the Arctic and limited coverage in South America, yet BIKER will be run on all river networks shown in the map. This needs to be stated here and addressed in the discussion later on.

L288 Why were these 16 rivers chosen? Later down (L 291) its stated they are from Frasson et al. 2021. Still should explain or characterize these 16 rivers and discuss any limitations in using such a small sample size.

L470-491 This conclusion is drawn from a very small sample size is it appropriate to subset down to 18 and 10 rivers and draw such conclusion? The limitations of such an approach should be addressed here.

We thank the reviewer for these comments, which are crucial to understanding the BIKER validation in this manuscript as an initial 'proof-of-concept'. We acknowledge that our limited validation data limits our ability to validate BIKER ahead of SWOT's launch, however we are inherently limited by the time-consuming processing to prepare these hydraulic models as 'simulated SWOT data' (Frasson et al. 2021; Durand et al. 2016). Therefore, we have added the following to section 3.1 to be more explicit about limitations in our validation set:

There is considerable geographic bias in our validation rivers, with rivers only present in North America, Western Europe, and Bangladesh. Further, no Arctic rivers are included. We acknowledge that this bias limits our ability to validate BIKER across many environments ahead of SWOT's launch. However, it is a sufficient validation set for a first proof-of-concept study consistent with the hydrology literature for SWOT. Further, the data requirements to create these test cases are strict and the processing time is enormous.

Specifically regarding the 16 rivers used to assess the influence of SWOT measurement error, lines 292-293 in the original manuscript state that Frasson et al. (2016) were limited to only these 16 rivers due to how computationally expensive the error modeling is. To further clarify this point, we expand this sentence to the following:

This error modeling is non-trivial and computationally expen where hydraulic data is concurrently available.sive, and so Frasson et al. (2021) were limited to only 16 test cases with SWOT measurement errors. Likewise, we stick to these 16 rivers for the same reason. These rivers are detailed in Figure S3. Given that it is only 16 synthetic rivers, caution should be used in over generalizing our results beyond our proof of concept.

Specific comments:

L159-164 Provide additional detail about the 763 measurements (e.g. range of the river width, range in k600). Do the k600 measurements cover different times of the year or are they from 763 distinct rivers? What is the geographic distribution, are extreme environments such as the spring melt period in the Arctic included? Table S1 only list the studies, it doesn't give any information about the diversity/range in the measurements.

We thank the reviewer for this suggestion, as the geographic and hydrologic extent of this dataset is not currently clear in the manuscript. We have added the following data description to section 2.1:

The 763 estimates cover different times of year and hydrological events. They include both individual estimates and repeat estimates in over 500 river reaches across the United States, Wales, Switzerland, and Austria. They span a wide variety of environments from temperate higher-order rivers to small mountain streams and represent a full range of river flows (width ranges from 0.26m to 1,742m, discharge ranges from 8e-4 m^3/s to 489 m^3/s, and k_{600} ranges from 0.1 m/dy to 4,118 m/dy). While there are still geographic and hydrologic biases in this dataset, it is to our knowledge the largest such dataset of field-estimated, reach-scale k_{600} where hydraulic data is concurrently available.

L226-227 How does the model function in low slope (i.e. flat) environments? As stated this is only valid for a "hydraulically-wide channel" but are there other constraints, e.g. detection limit of slope, flow depth and velocity?

While there should be no detection limit for the gas exchange model itself (equation 7 in the submitted version of the manuscript), BIKER could be limited by the resolution of the input hydraulic data being used. This is true if running BIKER via SWOT data: SWOT will have a slope detection limit of approximately 1.7 cm/km (Biancamaria et al. 2016). We have already accounted for this slope detection limit within this study's validation but neglected to mention it in the manuscript. So, the following is added to section 3.1:

Also note that SWOT water surface slope measurements will have a lower detection limit of 1.7 cm/km (Biancamaria et al. 2016), and therefore any slope measurement in our data less than this threshold was reassigned this minimum value.

L578-582 How likely is it that a "sufficient variability of 20% CV" will be captured in SWOT rivers? Need to address the limitations of rivers that have < 20% CV.

This is an excellent question and highlights a point we should make clearer in the manuscript. It is currently impossible to know how likely a 'sufficient variability in k_{600}' will be in SWOT rivers. This is one of the questions we hope to answer once BIKER is run on real SWOT data and we are able to infer the temporal variation in k_{600} for each SWOT-observable river. It is also a question that the SWOT discharge community seeks to answer during the calibration and validation phases of the satellite mission. However, for now we can try to contextualize this concern using a satellite much older than SWOT but with an approximately similar temporal resolution of 16 days: the Landsat program. Authors have recently shown that Landsat captures 97% of streamflow percentiles at 90% of United States gauges (Allen et al. 2020), implying sufficient hydraulic visibility for the Landsat record and, likely, SWOT as well. To clarify, we have added the following statement to section 4.2:

Although it is presently impossible to know whether SWOT will achieve 'sufficient hydraulic visibility' over its lifetime, recent similar work using the Landsat archive suggests that most rivers' full flow regime will be observed by the SWOT satellite: Allen et al. (2020) found that the Landsat record observed 97% of streamflow percentiles in 90% of United States streamgauges. Landsat has an average temporal resolution of 16 days, which is approximately similar to the repeat cycle for SWOT. Further, SWOT will penetrate clouds and provide even more data on cloudy days (unlike Landsat's optical sensor). With that said, the nominal lifespan of SWOT is only three years, within which certain streamflow magnitudes may not be experienced and reduce the chance that 'sufficient hydraulic visibility' is achieved.

L 630 What is meant by nearly entirely observable?

This refers to the fact that some parameters in equation 7 must still be inferred because SWOT fill not directly measure them (thus necessitating the Bayesian algorithm). However, we agree with the reviewer that it is needlessly confusing and, in the context of reviewer 1's comments on restructuring the conclusion, amend this sentence (now in the Discussion) to the following:

We developed 1) a wide-river-specific hydraulic model for k_{600} ~~that is nearly entirely SWOT observable and~~ that explains 50% of variation in k_{600}

Small Comments:

KP 1 Define what BIKER stands for.
KP 2 What makes the algorithm robust?
KP 3 How will BIKER allow for novel study of spatiotemporal gas exchange?
L24 Define SWOT
L25 upon launch and subsequent data collection
L27 include the range of "SWOT-observable rivers"
L47-49 could report carbon flux values for ocean and forest uptake here to have a better idea of the comparison.
L50 how is CO2 evasion "better constrained". Although details are mentioned in the following paragraph it would be helpful already here to briefly state here what is meant by better constrained.
L83 Upscaling can't be performed in the literature… perhaps, "upscaling has been performed using various techniques."
L95 define the gas exchange of oxygen (kO2)
Section 3 Header "Equation 7" should be removed from the heading, instead something along the lines of "BIKER algorithm development and validation"
L 507 Add units (Tg-C/yr)
L 631 could change "using no on-the ground information" to "using only remotely sensed information"
L 636 Could add a last sentence to state that it will help constrain the contribution of inland waters to GHG emissions.

We agree with all of the reviewer's suggestions and have made the appropriate changes to the manuscript.

References

Allen, G. H., Yang, X., Gardner, J., Holliman, J., David, C. H., & Ross, M. (2020). Timing of Landsat overpasses effectively captures flow conditions of large rivers. *Remote Sensing*, *12*(9), 1510.

Altenau, E. H., Pavelsky, T. M., Durand, M. T., Yang, X., Frasson, R. P. D. M., & Bendezu, L. (2021). The Surface Water and Ocean Topography (SWOT) Mission River Database (SWORD): A global river network for satellite data products. *Water Resources Research*, *57*(7), e2021WR030054.

Biancamaria, S., Lettenmaier, D. P., & Pavelsky, T. M. (2016). The SWOT mission and its capabilities for land hydrology. In *Remote sensing and water resources* (pp. 117-147). Springer, Cham.

Durand, M., Gleason, C. J., Garambois, P. A., Bjerklie, D., Smith, L. C., Roux, H., ... & Vilmin, L. (2016). An intercomparison of remote sensing river discharge estimation algorithms from measurements of river height, width, and slope. *Water Resources Research*, *52*(6), 4527-4549.

Durand, M., Neal, J., Rodríguez, E., Andreadis, K. M., Smith, L. C., & Yoon, Y. (2014). Estimating reach-averaged discharge for the River Severn from measurements of river water surface elevation and slope. *Journal of Hydrology*, *511*, 92-104.

Frasson, R. P. D. M., Durand, M. T., Larnier, K., Gleason, C., Andreadis, K. M., Hagemann, M., ... & David, C. H. (2021). Exploring the factors controlling the error characteristics of the Surface Water and Ocean Topography mission discharge estimates. *Water Resources Research*, *57*(6), e2020WR028519.

Gelman, A., Goodrich, B., Gabry, J., & Vehtari, A. (2019). R-squared for Bayesian regression models. *The American Statistician*.

Gleason, C., Garambois, P. A., & Durand, M. (2017). Tracking river flows from space. *EOS Earth & Space Science News*, (98).

Liu, S., Kuhn, C., Amatulli, G., Aho, K., Butman, D. E., Allen, G. H., ... & Raymond, P. A. (2022). The importance of hydrology in routing terrestrial carbon to the atmosphere via global streams and rivers. *Proceedings of the National Academy of Sciences*, *119*(11), e2106322119.

Raymond, P. A., Hartmann, J., Lauerwald, R., Sobek, S., McDonald, C., Hoover, M., ... & Guth, P. (2013). Global carbon dioxide emissions from inland waters. *Nature*, 503(7476), 355-359.