Supplemental Information: Gas exchange in large rivers controlled by largest turbulent eddies: implications for remotely sensing gas exchange via SWOT

CB Brinkerhoff, CJ Gleason, PA Raymond, MH Harlan

2021-07-27

## Contents

This supplementary information contains 4 texts, 5 figures and 2 tables. Please consult <https://github.com/craigbrinkerhoff/RSK600> for all code to build and generate results, figures, and the manuscript.

## Text S1: Deriving two expressions for turbulent kinetic energy production

Below we detail how we arrived at our two expressions for a cross-section-averaged turbulent kinetic energy production (*G*), or equation 3 in the main text. Please see Appendix A for variable descriptions and notation.

Nakagawa and Nexu (1993) outline the standard energy budget for turbulent kinetic energy (TKE), at some distance from the bed *h*, for a 2-dimensional open-channel flow (equation S1). *G* is the TKE produced, is the TKE dissipated, is TKE transported from the riverbed via turbulent diffusion, and is TKE transported via viscous diffusion. should always be positive at the free surface, but this is increasingly less true as one moves closer to the riverbed. Since we are interested in the TKE budget at the river surface, can be assumed neglible (Nakagawa and Nexu, 1993).

Nakagawa and Nexu (1993) further define *G* and the direct energy dissipation *E*, as averaged over the entire river cross-section and integrated over the entire water column, as equation S2 where *H* is the full water-column depth, *U* is the cross-sectional average flow velocity, and is the friction velocity. is defined by the water surface slope *S*, gravitational acceleration *g* and the channel hydraulic radius as .

can also be assumed neglible at the free surface (Nakagawa and Nexu, 1993). We then integrate over the full water-column depth *H* to arrive at an expression for *G* at the free surface that reduces to equation S5. Equations S3 to S4 show the algebra.

Alternatively, we can derive another expression for *G* from equation S2. This is provided as equation S8 with the necessary algebra in equations S6-S7.

## Text S2: BIKER hyperparameterization

We assign prior hyperparameters using SWOT data only. All priors are formalized within the model as truncated normal distributions of the log-transformed terms such that for , using prior hyperparameters mean (), standard deviation (), and upper () and lower bounds () for any parameter *X*.

prior hyperparameters were assigned following an updated version of the method developed by Brinkerhoff et al. (2020). They developed a set of river channel prior hyperparameters for McFLI algorithms that are entirely RS-able and reflect differential channel hydraulics as a function of river geomorphology. They used an extensive database of field measurements and machine learning to identify patterns that associate river width with the hydraulic priors needed to run McFLIs so that prior hyperparameters may be assigned to rivers using only the existing remotely sensed data. For this study, we extracted and as the 5th and 95th percentile values rather than the absolute maximum and minimum values to avoid physically impossible bounds on .

This leaves the hyperparameters to be defined. We assigned those by using a simple slope regression model to predict solely from a SWOT-observable value. It was trained on the Ulseth et al. (2019) dataset. Equation S9 assigns while was set to equation S9's log-transformed standard error after propogation of errors for both equation S9 and equation 6 in the main text. This amounted to a value of 1.12. and $\gamma\_{k\_600}}$ were set to log(0.001) m/day and log(500) m/day, respectively. Finally, we estimate using the standard model error from equation 6 in the main text and as fit on the data shown in Figure 3a.

## Text S3: BIKER validation setup

Regardless of the validation setup or SWOT error budget used, we do not have observed data for these rivers, and to our knowledge no field dataset of exists in the type of temporal and spatial frequency that SWOT (and therefore the BIKER algorithm) provides. Therefore, we take the model outlined in row 1 of Table 1 in the main text and use that to calculate the observed , that BIKER is validated against (equation S10).

It also means that, for a fair validation scheme, in equation 7 in the main text must be set to reflect only error from our assumptions about calculating *dA* and not the parameter uncertainity inherent in the model coefficient (i.e. the 56.0294 in equation 6 in the main text). Because this model is only theoretically valid when , we convientantly do not need to account for uncertanties associated with assuming that . So for this validation, is set to 0.20 (natural-log space) to reflect a small degree of uncertainity that could arise from assuming a rectangular river channel. However once SWOT launches, it should reflect the total uncertainty described in Text S3 (1.12).

Validation is performed using the BIKER posterior means. Validation metrics take two forms (and are detailed in Table S1). To validate across all rivers and timesteps, we used the coefficient of determination and the root mean square error RMSE. Four normalized metrics were used for by-river validation: RRMSE and NRMSE are normalized root mean square errors that have been normalized by the observed value and the mean observed value (respectively). rBIAS is a measure of prediction bias that is normalized by the mean observed value. r2 is again used on the by-river case too.

## Text S4: BIKER FCO2 workflow

To calculate fluxes and carbon efflux, we pair Beaulieu et al. (2012)'s biweekly 26 and water temperature samples with every 11th SWOT observation by date, ignoring the timesteps beyond 26. This amounts to only ~15% of the SWOT observations and we deem this acceptable. We sample every 11 days as this is the average sampling resolution for SWOT and the and water temperature data are approximately the same at a 14 day resolution. Not all of the SWOT rivers have observations for a full year, and when simulation dates are not available they are assumed to start on January 1st. We then pair the modeled values (obtained from and equation S11) with these water-side concentrations and water temperatures. In equation S11, *Sc* is the Schmidt number at the observed water temperature. This was calculated following Raymond et al. (2012) and Wanninkhof (1992). Atmospheric was assumed 390 uatm. validation was performed using the same metrics as validation (Table 1).

To calculate the bulk carbon efflux, we multiply the median value across all rivers and timesteps (in Tg-C/yr) by the total average surface area of all 47 rivers.

## Figures

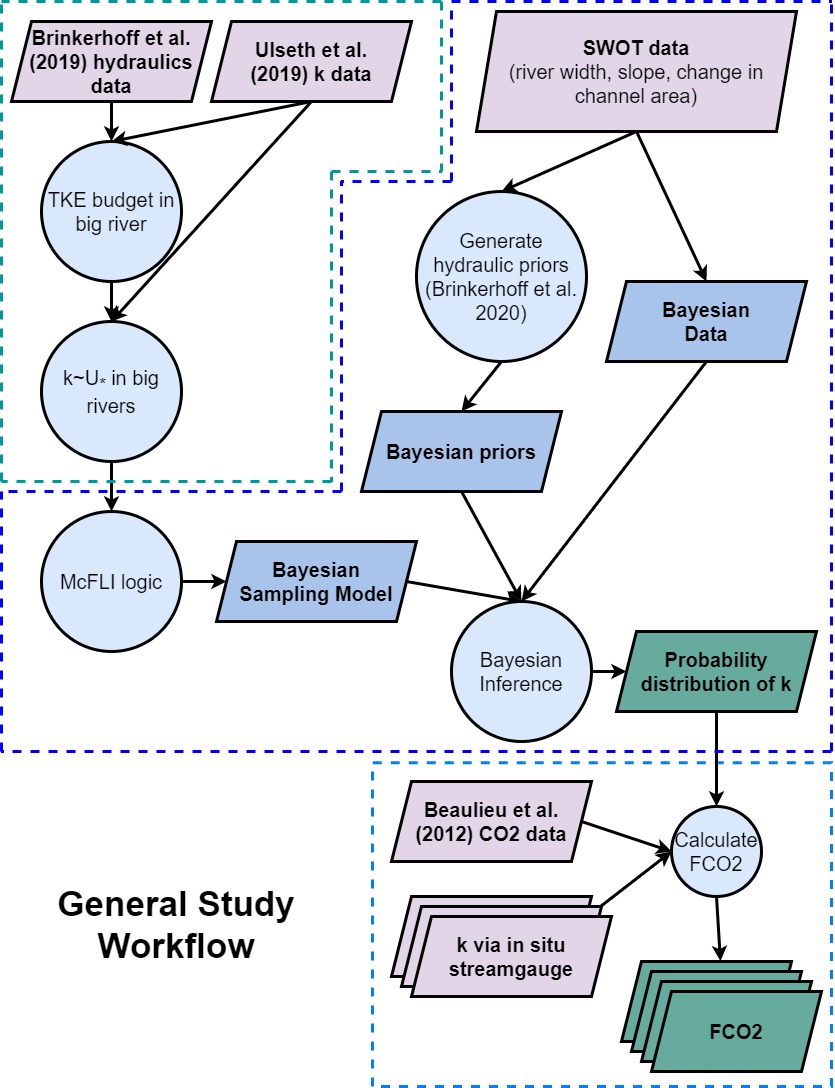


Figure S1: Flowchart of the entire study's workflow

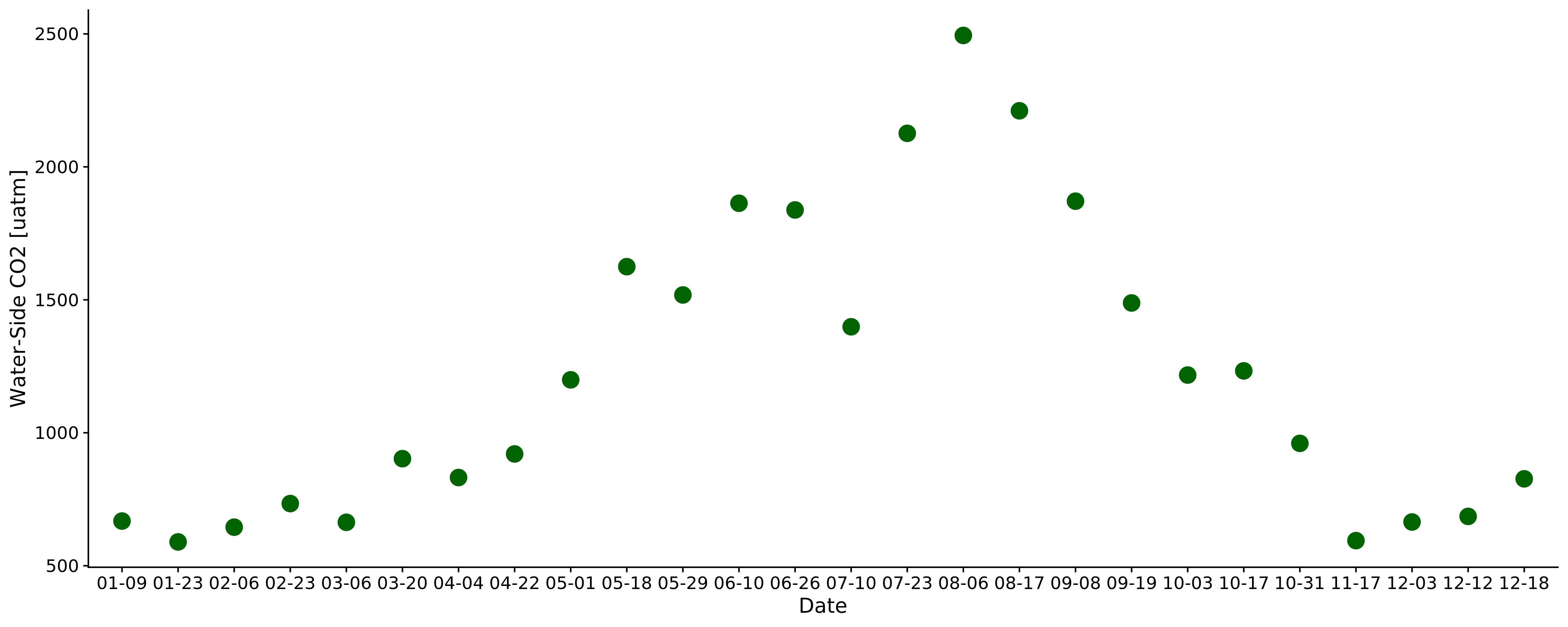


Figure S2 Timeseries of the biweekly CO2 data from Beaulieu et al. (2012). Sampling took place 2008-2009 in the Ohio River (upstream of Cincinnati, Ohio, United States). Each point here was joined to the 11-day SWOT observations used in this study (section 2.4).

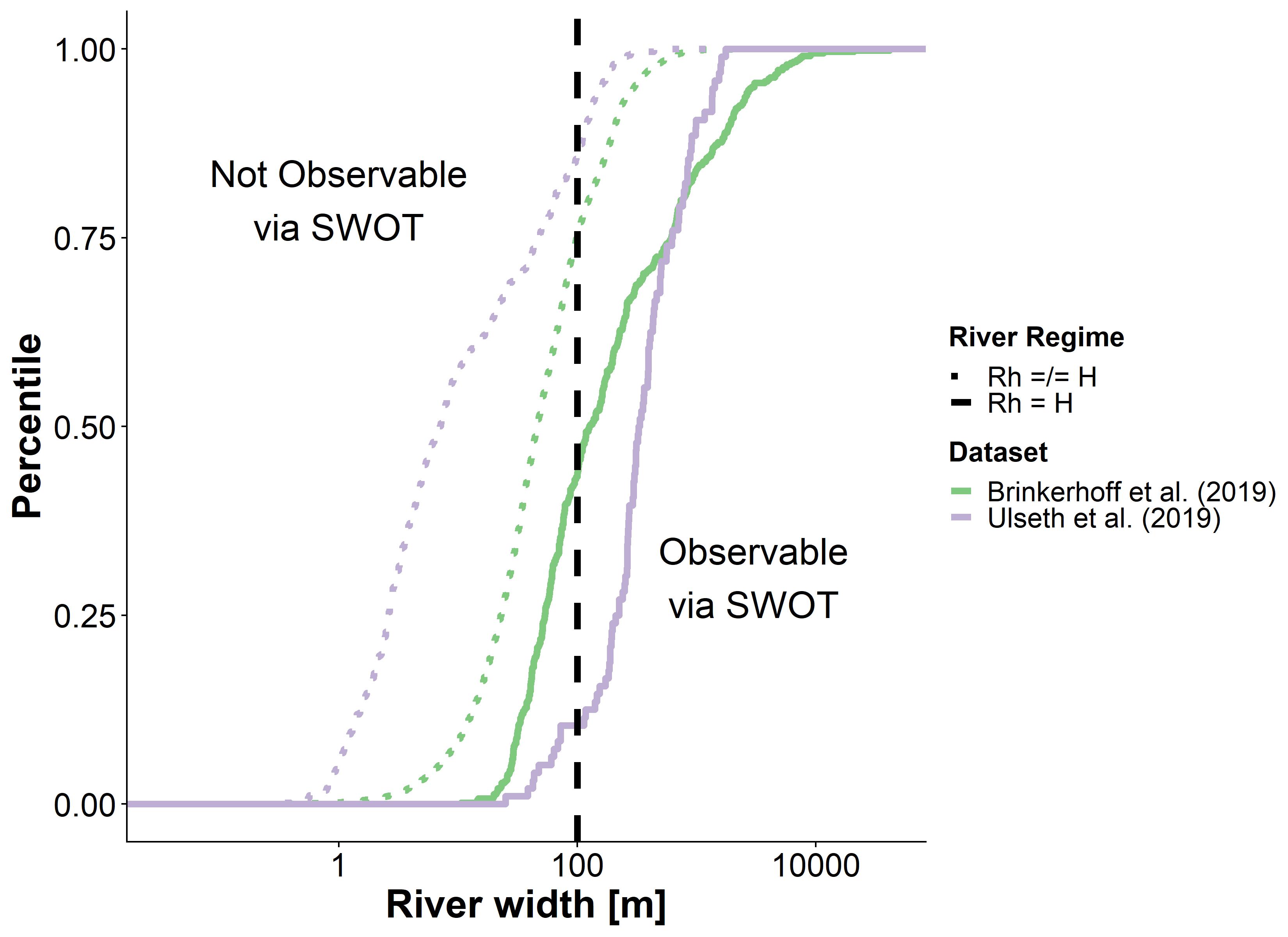


Figure S3 Cummulative density functions (CDFs) of river width measurements, as binned by their dataset and and whether or not they meet the 'large river' condition (Rh=H). The dashed black line indicates the minimum river width that the SWOT satellite will observe (100m). This brief analysis indicates that the vast majority of SWOT-observable rivers should meet the 'large river' condition.

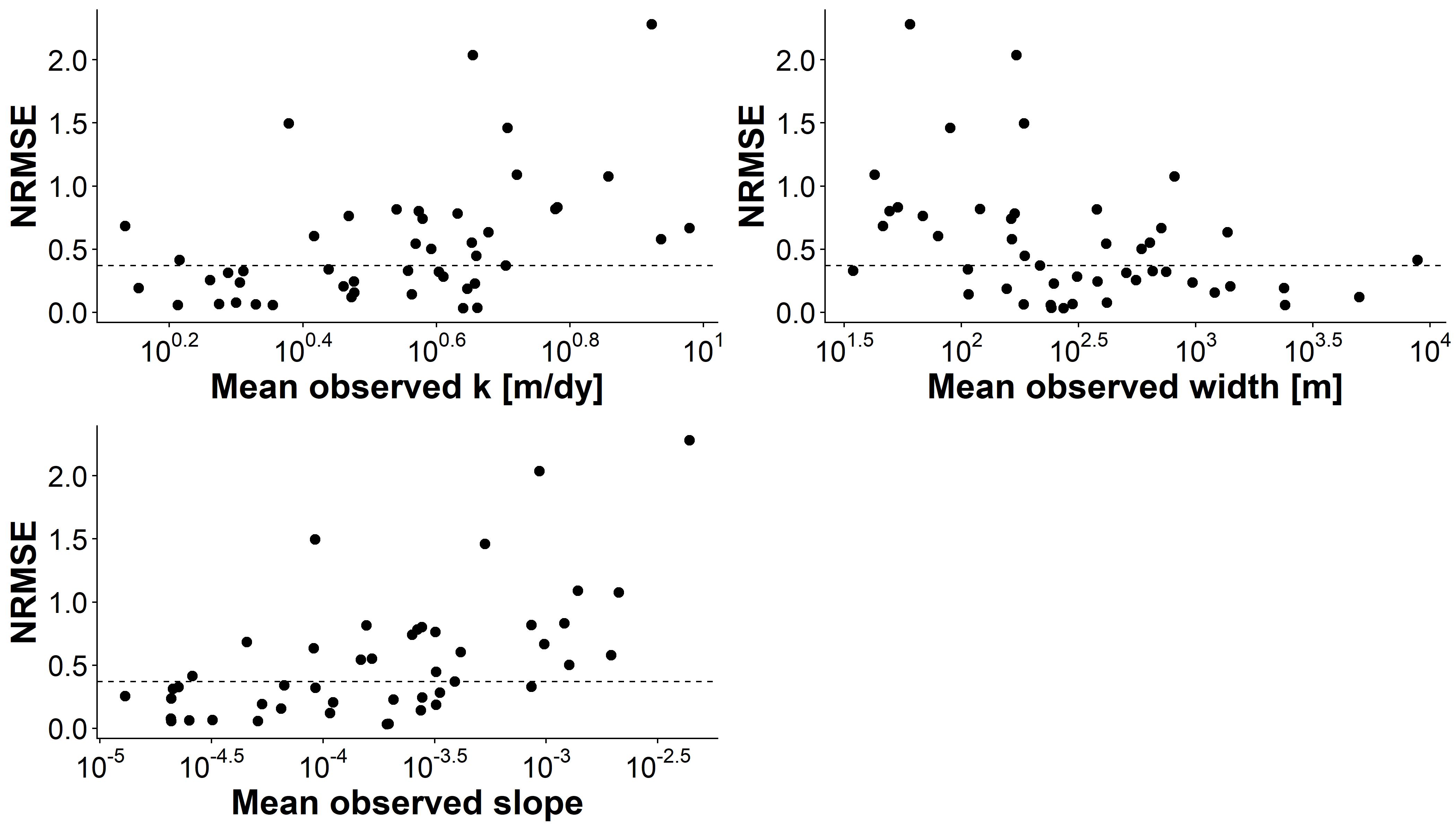


Figure S4 Mean river properties (observed k600, width, and slope) for the 47 rivers versus their BIKER performance as defined by NRMSE (see Table S1 for this metric's definition). The dashed black line is the median NRMSE score. The poorly performing rivers are universally the narrowest and steepest of the set.

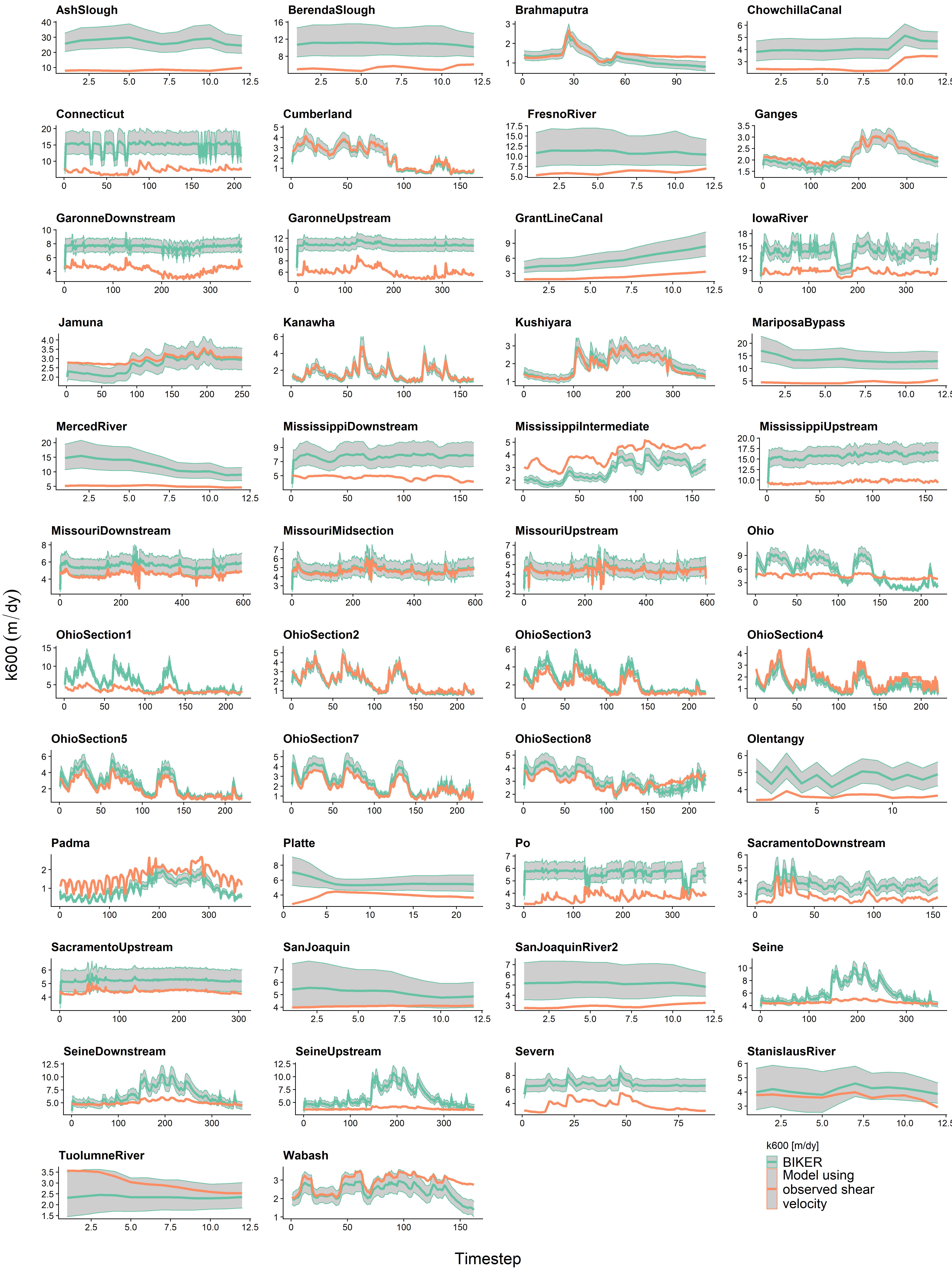


Figure S5 All timeseries of k per river

## Tables

*Table S1: Validation metrics used in this study, where r is the correlation coefficient, Nt is number of observations and i is the specific observation. σ refers to the variance of the sample and μ refers to the mean of the sample. As is standard, a carrot accent indicates the predicted value.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Description** | **Acronym** | **Definition** | **Ideal Score** | **Possible Range** | **Validation Scheme** |
| Coefficient of determination |  |  | 1 | 0 to 1 | All rivers and all timesteps |
| Root-mean-square-error | RMSE |  | 0 | 0 to ∞ | All rivers and all timesteps |
| Relative root-mean-square error | RRMSE |  | 0 | 0 to ∞ | By river |
| Normalized root-mean-square error | NRMSE |  | 0 | 0 to ∞ | By river |
| Relative bias | rBIAS |  | 0 | -∞ to ∞ | By river |

*Table S2: Details on the 3 depth hydraulic geometry models used to estimate FCO2 from the SWOT rivers (section 2.4).*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Name** | **Equation** | **Description** | **Reference** |  |
| Brinkerhoff 2019 |  | 530,945 measurements made across the United States at streamgauges | this study; Brinkerhoff et al. (2019) |  |
| Raymond 2012 |  | 1,026 measurements across the United States | Raymond et al. (2012) |  |
| Raymond 2013 |  | Average of the Raymond 2012 equation and one using 9,811 measurements at US streamgauges | Raymond et al. (2013) |  |

## References

Beaulieu, J.J., Shuster, W.D., Rebholz, J.A., 2012. Controls on gas transfer velocities in a large river. Journal of Geophysical Research: Biogeosciences 117. <https://doi.org/10.1029/2011JG001794>

Brinkerhoff, C.B., Gleason, C.J., Feng, D., Lin, P., 2020. Constraining Remote River Discharge Estimation Using Reach-Scale Geomorphology. Water Resources Research 56, e2020WR027949. <https://doi.org/10.1029/2020WR027949>

Brinkerhoff, C.B., Gleason, C.J., Ostendorf, D.W., 2019. Reconciling at-a-Station and at-Many-Stations Hydraulic Geometry Through River-Wide Geomorphology. Geophysical Research Letters 46, 9637–9647. <https://doi.org/10.1029/2019GL084529>

Nakagawa, H., Nexu, I., 1993. Turbulence in Open Channel Flows. Routledge.

Raymond, P.A., Hartmann, J., Lauerwald, R., Sobek, S., McDonald, C., Hoover, M., Butman, D., Striegl, R., Mayorga, E., Humborg, C., Kortelainen, P., Dürr, H., Meybeck, M., Ciais, P., Guth, P., 2013. Global carbon dioxide emissions from inland waters. Nature 503, 355–359. <https://doi.org/10.1038/nature12760>

Raymond, P.A., Zappa, C.J., Butman, D., Bott, T.L., Potter, J., Mulholland, P., Laursen, A.E., McDowell, W.H., Newbold, D., 2012. Scaling the gas transfer velocity and hydraulic geometry in streams and small rivers. Limnology and Oceanography: Fluids and Environments 2, 41–53. <https://doi.org/10.1215/21573689-1597669>

Ulseth, A.J., Hall, R.O., Boix Canadell, M., Madinger, H.L., Niayifar, A., Battin, T.J., 2019. Distinct airWater gas exchange regimes in low- and high-energy streams. Nature Geoscience 12, 259–263. <https://doi.org/10.1038/s41561-019-0324-8>

Wanninkhof, R., 1992. Relationship between wind speed and gas exchange over the ocean. Journal of Geophysical Research: Oceans 97, 7373–7382. <https://doi.org/10.1029/92JC00188>