Towards Modeling Continental-Scale Inland Water Carbon Dioxide Emissions

Brian Saccardi[1†], Craig B Brinkerhoff[2†], Colin J Gleason[2], Matthew J Winnick[1]

[1]Department of Earth, Geographic, and Climate Sciences, University of Massachusetts, Amherst

[2]Department of Civil & Environmental Engineering, University of Massachusetts, Amherst

Corresponding author: Matthew J. Winnick (mwinnick@umass.edu)

[†]These authors contributed equally

Key Points:

- We develop and calibrate a river network carbon dioxide transport model for the continental United States to estimate emission fluxes
- Compared to previous methods, this model simulates 25% lower carbon dioxide emissions using the same data constraints
- Stream corridor respiration dominates over groundwater sources, but better source constraints are needed for accurate forward predictions

Abstract

Inland waters emit significant amounts of carbon dioxide (CO_2) to the atmosphere; however, the global magnitude and source distribution of inland water CO_2 emissions remain uncertain. These fluxes have previously been 'statistically upscaled' by independently estimating dissolved CO_2 concentrations and gas exchange velocities to calculate fluxes. This scaling, while robust and defensible, has known limitations in representing carbon source limitations and spatial variability. Here, we develop and calibrate a CO_2 transport model for the continental United States, simulating carbon transport and transformation in >22 million hydraulically connected rivers, lakes, and reservoirs. We estimate 25% lower CO_2 fluxes compared to upscaling estimates forced by the same observational calibration data. While precise CO_2 source distribution estimates are limited by the resolution of model parameterizations, our model suggests that stream corridor CO_2 production dominates over groundwater inputs at the continental scale. Our results further suggest that the lack of observational networks for groundwater CO_2 and scalable metabolic models of aquatic CO_2 production remain the most salient barriers to further coupling of our model with other Earth system components.

Plain Language Summary

Inland water CO_2 emissions are recognized as an important but highly uncertain component of the global carbon cycle. Estimates rely on methods that statistically upscale point observations that are unable to account for the distribution and limits of CO_2 sources. Here we present a first step towards distributed process-based models that link CO_2 fluxes to water transport in connected rivers, lakes, and reservoirs at the continental scale. We show that using the same data constraints, incorporating water transport results in a 25% reduction relative to previous methods in estimated inland water CO_2 fluxes over the continental United States. We identify barriers to monitoring and prediction that will enable the incorporation of inland water carbon into earth system models and global budgets.

1 Introduction

Inland waters, here comprising rivers, lakes, and reservoirs, are an integral component of the global carbon cycle, particularly in their role in emitting CO_2 to the atmosphere. Recent estimates of CO_2 fluxes from inland waters are on the order of 1.5 Pg-C yr^{-1} (Lauerwald et al., 2023b), roughly 15% of anthropogenic emissions (Friedlingstein et al., 2022) and similar to the net terrestrial carbon sink (Cavallaro et al., 2018; Keenan & Williams, 2018). These estimates have steadily risen over the past decade (Drake et al., 2018) with increasing satellite resolution of lotic environments (Allen & Pavelsky, 2018) and extensive sampling campaigns in tropical environments (Borges et al., 2015; Sawakuchi et al., 2017). Despite our growing knowledge, estimates remain highly uncertain due to the inherent challenges of upscaling point measurements of stream CO_2 concentrations, which can vary by orders of magnitude over short reaches (Duvert et al., 2018; Johnson et al., 2008; Lupon et al., 2019), and due to a lack of representation of inland water CO_2 fluxes in global carbon cycle models (Friedlingstein et al., 2022). Additional uncertainty is derived from systematic errors associated with physical hydraulic constraints on dissolved CO_2 concentrations (Rocher-Ros et al., 2019; Saccardi & Winnick, 2021) and the artificial separation of lotic and lentic environment flux estimates (Brinkerhoff et al., 2021). A number of studies have thus called for process-based models to advance total flux estimates and to better facilitate monitoring and prediction efforts to gauge the response of inland water carbon cycling to climate change (Battin et al., 2023; Duvert et al., 2018).

Inland water CO_2 fluxes represent the culmination of CO_2 transported from soil and groundwater environments, as well as CO_2 produced internally via respiration in aquatic and hyporheic environments as balanced by photosynthetic uptake (Duvert et al., 2018; Gómez-Gener et al., 2021; Hotchkiss et al., 2015). Evasion fluxes of CO_2 (F_{CO2}; mol m^{-2} s^{-1}) from inland waters are calculated as,

$$F_{CO2} = k_{CO2}(CO_{2(aq)} - C_{atm}) \qquad \text{(Eq. 1)},$$

where k_{CO2} is the gas exchange velocity of CO_2 (m s^{-1}), $CO_{2(aq)}$ is the dissolved CO_2 concentration (mol m^{-3}), and C_{atm} is the atmospheric-equilibrated concentration of CO_2 (mol m^{-3}). Current best estimates of global CO_2 contributions from inland waters (Butman & Raymond, 2011; Lauerwald et al., 2015, 2023a, 2023b; Liu, Kuhn, et al., 2022; Raymond et al., 2013) rely on 'statistical upscaling' methods, in which water $p$$CO_2$ is scaled using regional observational averages (Butman et al., 2016; Butman & Raymond, 2011; Lauerwald et al., 2023a, 2023b; Raymond et al., 2013) or by relating observations of CO_2 concentrations to watershed characteristics and applying those statistical relationships globally (Horgby et al., 2019; Lauerwald et al., 2015; Liu, Kuhn, et al., 2022). k_{CO2} is typically estimated by applying empirical relationships from observational studies via stream discharge and slope (Raymond et al., 2012; Ulseth et al., 2019) and scaled as a single value across entire watersheds (Butman et al., 2016; Butman & Raymond, 2011; Raymond et al., 2013). Despite the significant progress that statistical upscaling has enabled, independent treatment of $CO_{2(aq)}$ and k_{CO2} within upscaling calculations is not consistent with established hydraulic controls on k_{CO2} (Brinkerhoff et al., 2022; Raymond et al., 2012; Ulseth et al., 2019) or recent work showing $CO_{2(aq)}$ rarely reaches elevated levels when k_{CO2} is high (Rocher-Ros et al., 2019; Saccardi & Winnick, 2021) (Fig. 1). The absence of high stream $CO_{2(aq)}$ values under turbulent, high k_{CO2} conditions is due to source limitations on CO_2 inputs that are unable to keep pace with evasion rates, and studies have shown that statistical models' inability to account for these limitations may lead to overestimates in global CO_2 fluxes by as much as 50% (Rocher-Ros et al., 2019; Saccardi & Winnick, 2021). This potential error reflects that fact that under CO_2 source limitations, the product of mean k_{CO2} and mean $CO_{2(aq)}$ values (statistical upscaling methods) is higher than the mean of local k_{CO2} and $CO_{2(aq)}$ products. A recent study of global methane fluxes suggests that machine-learning algorithms are also subject to overestimating gas fluxes from turbulent reaches (Rocher-Ros et al., 2023).

The distribution of inland water CO_2 sources also represents a significant knowledge gap, both in terms of where CO_2 is emitted (i.e. rivers v. lakes/reservoirs) and the balance of terrestrial versus internal CO_2 production. Constraining the latter is particularly important to better gauge potential carbon cycle feedbacks, and previous work presents conflicting findings. Broadly, studies that focus on scaling CO_2 fluxes based on direct concentration measurements and carbonate speciation calculations identify stream corridor production as the dominant source (Butman & Raymond, 2011; Kirk & Cohen, 2023; Rasilo et al., 2017; Saccardi & Winnick, 2021), whereas stream metabolism measurements based on diel dissolved oxygen variations identify external groundwater inputs as dominating CO_2 budgets (Hotchkiss et al., 2015). Stream reach and watershed scale studies of CO_2 budgets, for example, suggest that stream corridor CO_2 sources may dominate in all but headwater stream systems (Kirk & Cohen, 2023; Rasilo et al., 2017; Saccardi & Winnick, 2021). Continental-scale CO_2 flux estimates further suggest that terrestrial sources can only account for ~25% of inland water CO_2 gas fluxes assuming relatively high groundwater pCO_2 values of 25,000 ppm (Butman & Raymond, 2011). In contrast, comparison of dissolved oxygen-based stream metabolism estimates and CO_2 fluxes at stream sites across the US suggest that terrestrial groundwater inputs dominate across all stream environments (Hotchkiss et al., 2015). Generally, oxygen-based estimates of stream net ecosystem production (NEP) of carbon are relatively low (Bernhardt et al., 2022), with global estimates of 0.27 Pg-C yr^{-1} (Battin et al., 2023) contributing only 18% of the estimated 1.5 Pg-C yr^{-1} global inland water CO_2 emissions (Lauerwald et al., 2023a). Additionally, soil respiration metrics are among the strongest statistical predictors of stream pCO_2 (Liu, Kuhn, et al., 2022).

Process-based transport models with distributed CO_2 source and sink representation, proper hydrographic representation, and explicit downstream routing have the potential to address many of these uncertainties and knowledge gaps. Specifically, transport models incorporate a hydrologic system's upstream history and have been applied at the watershed scale to predict the downstream transport of CO_2 (Brinkerhoff et al., 2021; Saccardi & Winnick, 2021), dissolved organic carbon (DOC) (Maavara et al., 2023), and other nutrients (Schmadel et al., 2018, 2019; Segatto et al., 2023). These transport models also enable explicit modeling of river corridor connectivity, including lake and reservoir connectivity, to the river network (Brinkerhoff, 2024). The latter was recently shown to exert controls on carbon/nutrient transport through inland waters (Brinkerhoff et al., 2021; Liu, Maavara, et al., 2022; Maavara et al., 2023; Schmadel et al., 2018, 2019). Likewise, considerable progress has been made in mapping hydrography globally for millions of rivers, lakes, and reservoirs (Lehner et al., 2008; Lin et al., 2021; Messager et al., 2016; R. B. Moore et al., 2019; Sikder et al., 2021; Wang et al., 2022), but the missing link to deploy these advances has been efficient computation for process-based transport models at scale. Here, we demonstrate the potential for coupled hydrologic and biogeochemical models that extend and expand upon statistical upscaling to advance our understanding of inland water CO_2 fluxes.

We calibrate and deploy a CO_2 transport model for over 22 million rivers, lakes, and reservoirs across the continental United States (CONUS) at mean annual flow for 1970-2000, which explicitly simulates advection of CO_2 from headwaters to the sea and reach-scale CO_2 production from net respiration within the stream channel (respiration – primary productivity), respiration within the stream corridor subsurface introduced via hyporheic exchange, and lateral groundwater CO_2 inputs. We assess the difference in CONUS-scale fluxes between our transport model and previous statistical upscaling techniques using identical observational constraints. We further evaluate the magnitude and uncertainties of modelled CO_2 source distributions (lotic v. lentic and external v. internal) and identify the most salient barriers towards providing robust CO_2 estimates from process-based models that must be addressed moving forward.

2 Materials and Methods

To ask how the distributed nature of hydrography and CO_2 sources along the stream-to-ocean continuum impacts continental-scale CO_2 flux estimates, we use the same CO_2 data to drive two different models for the CONUS CO_2 emissions and compare the differences in the resulting estimates. These two models are a process-based transport model and a traditional statistical upscaling model. A more detailed description of modeling methods is included in the Supplemental Information, and all model validation and calibration performance analyses are detailed in Supplementary Figs. S9-S21 and Supplementary Table S3-S4. It is important to stress that we do not aim to reproduce CO_2 concentrations in individual rivers, nor do we aim to rectify any biases in existing CO_2 databases or statistical upscaling methods. By using the same CO_2 data for all tested models, we specifically isolate the role that heterogeneous hydrography and CO_2 sources play in continental-scale flux estimates.

2.1 Dissolved CO_2 data

All models presented are either run or calibrated using the same CO_2 data. These data are obtained from the GLORICH database (Hartmann et al., 2014), which includes 1.27 million samples from across the world. Riverine CO_2 is calculated from GLORICH measurements of alkalinity and pH, although there is a long-standing concern for overestimation of CO_2 via this approach in acidic waters as small errors in pH leads to large errors in calculated CO_2 (Abril et al., 2015; Liu et al., 2020). To address this problem in this study, we filtered GLORICH for samples with a pH >5.4, and took the median value at an individual sample location resulting in 6,324 CO_2 estimates across the CONUS. We then mapped these 6,324 samples to US regions using an inverse distance weighted approach to make an interpolated grid with 0.5x0.5 degree resolution following previous similar work (Raymond et al., 2013). This grid was cut to each CONUS region as defined by our hydrography and the mean CO_2 was calculated and used for model calibration and/or forcing. To obtain lake/reservoir CO_2 estimates, we follow the method described in Raymond et al. (2013) using CO_2 data from the GLORICH database. This method requires estimates of lake surface area and dissolved organic carbon (DOC) per region. We calculate lake surface area for each region from the global lakes and wetlands database (GLWD) (Lehner & Döll, 2004) by summing the estimated surface area of five lake size classes. We calculate each size class surface area by multiplying the estimated cumulative abundance by mean surface area. We estimate the lake DOC for each region by taking the DOC value at the river mouth with the largest discharge from the GLOBALNEWS dataset (Mayorga et al., 2010). If the region does not discharge into the ocean directly, we use the DOC of the region it discharged into. For endorheic basins, we use a median lake pCO_2 of 340 ppm following the Raymond et al. (2013) analysis.

2.2 CO_2 Transport Model

The underlying hydrology and hydrography are an extension of a previously developed river/lake/reservoir CO_2 routing framework (Brinkerhoff et al., 2021), which explicitly coupled rivers, lakes and reservoirs into a routing scheme that enabled offline solute transport modeling in the Connecticut River watershed. Here, we extend this framework to CONUS using the USGS National Hydrography Dataset High-Resolution (R. B. Moore et al., 2019) (NHD-HR), excluding the Mexican and Canadian basins that do not directly flow into CONUS. Additionally, the NHD-HR is discretized into 'reaches', or mass-conserved segments of river, lake, or reservoir. Network topology is maintained through lakes/reservoirs via artificial flowlines. We assign fractions of lakes/reservoir morphometry to the artificial flow lines to account for complex waterbodies with multiple inflows (Brinkerhoff et al., 2021). Further details on preprocessing the NHD-HR for our modeling are provided in the Supplementary Information.

The NHD-HR features a nested basin scheme. We run our analysis at the 4th level (HUC4) due to computation, data availability, and ease of interpretation. We split the 4th-level basin for coastal Washington State into two separate basins (coastal catchments on either side of the Columbia River) to ease computation requirements (the two sub-basins are added back together as a single basin in our presented results). CO_2 data are calibrated at the 2nd watershed level (HUC2), which are regional amalgamations of 4th level basins. We also run our statistical upscaling analysis at the 2nd level.

To run the model on a drainage network, we use estimates of reach-level discharge (Q, $m^3 s^{-1}$), surface area (A, m^2), hydraulic residence time (τ, s), bed slope (S, $m\ m^{-1}$), mean depth ($\underline{H}$, m), and additionally for rivers, width (W, m) and mean flow velocity ($\underline{U}$, $m\ s^{-1}$). S is provided in the NHD-HR and calculated from a digital elevation model, and for missing values we use the average S across the immediately upstream reaches. We use the mean annual discharge model provided with the NHD-HR, described in detail in the SI (R. B. Moore et al., 2019). We validate the discharge model against observed mean annual flow for 1970-2018 in reaches with corresponding stream gauges (Fig. S10). For ‘emergent’ streams we set the emergent discharge at the upstream end of the reach to reflect initial streamflow conditions for the start of the network. We use a consistent emergent stream width of approximately 30cm, identified in headwater networks around the world (Allen et al., 2018). The remaining variables are calculated based on hydraulic geometry and global database fitting as discussed in the Supplement.

We adapt a previously developed CO_2 stream network model (Saccardi & Winnick, 2021) to incorporate lakes and reservoirs, as,

$$\frac{dC}{dt} = -\underline{U}\frac{dC}{dX} + \frac{1}{A}\frac{dQ}{dX}\left(C_{gw} - C\right) - \frac{k_{CO2}}{H}\left(C - C_{atm}\right) + \frac{k_{hz}}{H}C_{hz} + F_{wc} \qquad \text{(Eq. 2),}$$

where C is the concentration of dissolved CO_2 (mol m^{-3}), x is distance along a reach (m), C_{gw} and C_{atm} are dissolved CO_2 concentrations of groundwater and atmosphere-equilibrated water (mol m^{-3}), respectively, C_{hz} is the difference in dissolved CO_2 between the stream and the hyporheic zone (mol m^{-3}), F_{wc} is the water column net respiration rate (mol m^{-3} s^{-1}), and k_{hz} is the hyporheic exchange velocity (m s^{-1}).

Based on our model framework, CO_2 sources are classified as (1) upland groundwater inputs, representing terrestrial respiration and subsurface water-rock interactions that scale with upstream contributing area and a set groundwater CO_2 concentration; (2) net respiration within the surface water column, and (3) respiration within the subsurface stream corridor environment comprising stream benthic zones, the hyporheic zone, and near-stream riparian zones and floodplains (Fig. S1). In terms of stream corridor subsurface respiration, these input fluxes are modeled via turbulent exchange across the stream’s sediment-water-interface (e.g. Grant, Azizian, et al., 2018; Winnick, 2021), where elevated CO_2 concentrations at the sediment-water-interface represent the accumulated respiration from the subsurface stream corridor. As these C_{hz} values are calibrated based on observational stream $CO_{2(aq)}$ data, along with F_{WC}, they physically represent the integrated stream corridor respiration needed to match regional CO_2 observations in excess of upland groundwater inputs.

This model, based on traditional solute transport frameworks (e.g. Bencala & Walters, 1983), represents downstream solute advection, solute inputs from lateral groundwater inputs, atmospheric equilibration, solute inputs from the subsurface stream corridor environment facilitated by hyporheic exchange, and net solute production within the water column. Within stream environments, k_{CO2} is parameterized using the empirical relationships from Ulseth et al. (2019), calculated based on channel slope and water depth. Hyporheic exchange rates (k_{hz}) in this model represent turbulence-driven exchange across the sediment-water-interface based on surface renewal theory (Grant, Gomez-Velez, et al., 2018) that dominate overall water exchange fluxes (Grant, Azizian, et al., 2018; Harvey et al., 2019). The adaption to lakes and reservoirs is achieved by rearranging Eq. 2 such that it is based on τ rather than $\underline{U}$, by incorporating alternative parameterizations for k_{CO2} and k_{hz} for lakes/reservoirs (Lorke & Peeters, 2006; Raymond et al., 2013; Read et al., 2012) (see SI for details). We note that k_{hz} for lakes/reservoirs represents benthic water-sediment fluxes (Lorke & Peeters, 2006). Lakes were assumed to be well mixed under long-term average conditions, meaning that lake stratification's influence on residence time was not considered. We also assume that benthic and atmospheric lake interfaces were both equal to the lake's surface area, acknowledging that many lakes have complicated and highly heterogenous morphologies. CO_2 is converted between partial pressure and dissolved concentration using a temperature-dependent Henry's constant. Within our modeling framework C_{hz} and F_{wc} are free parameters, and the remaining variables are either fixed or calculated based on published scaling relationships (see SI for detailed parameterizations).

We use a genetic algorithm (GA) to determine optimal parameter sets of C_{hz} and F_{wc} to match GLORICH CO_2 observations at the HUC2 scale. GAs do not rely on derivative information about one's function a priori (unlike a gradient-based optimization method). Instead, GAs use many evolutions of parameter sampling to explore the solution space stochastically, though often they take a hybrid approach that leverages a gradient search within the GA. This is particularly useful for noisy solution spaces, problems that suffer from equifinality (multiple possible solutions to the same function due to due complex interactions of system processes- e.g. Beven, 1993), or when there is little prior knowledge of what the solution space looks like. Finally, because each 'generation' of GA evolution is composed of many independent model runs, GAs are readily parallelized and allow for straightforward computational scaling as required for the scale of this study (Mitchell, 1998). Our fitness function is specified as,

$$cost = \frac{1}{|(pCO2_{lake} - cal_{lake})| + |(pCO2_{river} - cal_{river})|} \quad \text{(Eq. 3)},$$

which we sought to maximize, where $pCO_{2,lake}$ and $pCO_{2,river}$ are the model's median HUC2 lake and river CO_2 partial pressures, respectively, and cal_{lake} and cal_{river} are the upscaled CO_2 partial pressures for lakes and rivers, respectively (Extended Data Table 2). The four parameters we calibrate are river C_{hz}, lake/reservoir C_{hz}, river F_{wc}, and lake/reservoir F_{wc} from equation 6. C_{gw} is held constant at 16,000 ppm (Kessler & Harvey, 2001; Macpherson, 2009) as groundwater was found to range from ~5,000 to 30,000 in the US and shallow groundwaters; however, we note we were not able to make groundwater spatially variable due to the lack of available groundwater

pCO$_2$ data products. We note that these values are consistent with measured and calculated upland shallow groundwater pCO$_2$ in stream carbon budget studies across a range of environments (Kirk & Cohen, 2023; Lupon et al., 2019; Saccardi & Winnick, 2021). We run the GA for 500 generations but terminate after 50 successive generations with no performance improvement. Each generation is composed of 25 individual runs. We terminate the calibration once the model cost goes below 10 ppm (or equivalently, 5 ppm per river or lake/reservoir). All modeling and geospatial analyses were run in R on the Unity Cluster at the Massachusetts Green High Performance Computing Center. Calibration results by basin are presented in Supplementary Figs. S11-S21.

We define calibration uncertainty per basin as $\delta FCO_{2,transport}$ using equation 4, where k_{median} is the median k_{CO2} across all reaches, A_{basin} is the total inland water surface area, and δpCO_2 is the calibration error for the median river/lake/reservoir. In effect, equation 4 applies the error in the median river/lake/reservoir pCO$_2$ over the network's entire surface area. We sum $\delta FCO_{2,transport}$ across all basins to obtain a CONUS uncertainty estimate (error bar in Fig. 2c).

$$\delta CO_{2,transport} = k_{median}\delta pCO_2 A_{basin} \quad \text{(Eq. 4)}$$

$$\delta pCO_2 = (1/cost)/2 \quad \text{(Eq. 5).}$$

2.3 Statistical Upscaling Model

Our 'statistical upscaling model' is informed by previous approaches to estimating inland water CO$_2$ emissions at large scales (Butman et al., 2016; Butman & Raymond, 2011; Lauerwald et al., 2023a; Liu, Kuhn, et al., 2022; Raymond et al., 2013). We calculate FCO_2 using a regionally-lumped pCO$_2$ and k_{CO2}, separately for rivers and lakes/reservoirs. This regionally homogenous FCO_2 is then applied to the region's total inland water surface area to obtain a CO$_2$ emissions estimate. Following previous methods (Butman et al., 2016; Butman & Raymond, 2011; Raymond et al., 2013), we calculate river lumped $k_{CO2,upscale}$ using mean k_{CO2} by stream order and then take the average of those values, weighted by stream order surface area (note that these approaches treat lakes/reservoirs as rivers during the stream order averaging- emissions are even higher when we remove them from the river network). This means that differences in FCO_2 estimates cannot come from different k_{CO2} equations, as k_{CO2} calculations are identical across all models. The only difference is the stream order averaging and lumping approach. We estimate this uncertainty ($\delta FCO_{2,\mathrm{upscale}}$) using equations 6-7, incorporating $k_{\mathrm{CO2,upscale}}$, and the total river surface area A_{river}. In effect, equation 6 applies the error in $k_{\mathrm{CO2,upscale}}$ over the network's entire surface area. We sum $\delta FCO_{2,upscale}$ across all regions to obtain a CONUS uncertainty estimate (error bar in Fig. 1c).

$$\delta FCO_{2,upscale} = \delta k_{CO2} pCO_2 A_{river} \quad \text{(Eq. 6)}$$

$$\delta k_{CO2} = abs\left(k_{median} - k_{CO2,upscale}\right) \quad \text{(Eq. 7).}$$

3 Results & Discussion

3.1 Continental-scale flux estimates and regional patterns

Following the calibration of our CO_2 transport model production parameters, CONUS inland water emissions are estimated as 120±23 Tg-C yr^{-1} (Fig. 1) (uncertainty from Eq.'s 6,7). This estimate is larger than several previous CONUS estimates from statistical upscaling methods (Table S1), and results from our use of explicit, high-resolution NHD-HR hydrography rather than statistical river and pond size distributions for the smallest waterbodies. Specifically, NHD-HR hydrography features exponentially more small water bodies, in particular low Strahler order streams, with higher area-normalized fluxes than accounted for in previous studies. This result demonstrates the importance of using high resolution hydrography to capture the full extent of inland water surface area, as described in previous studies (e.g. Allen & Pavelsky, 2018). To evaluate the direct impacts of incorporating transport constraints on CONUS CO_2 fluxes, we compare this estimate to one calculated using statistical upscaling techniques while applying the same gas exchange model to the same NHD-HR hydrography and interpolated average HUC2-level pCO_2 values estimates, which yields total CONUS inland water fluxes of 159±55 Tg-C yr^{-1} (Fig 1c) – a difference of 25%.

Notably, the largest differences between the transport and statistical models occurs in the East and Midwest US where the transport model estimates significantly lower fluxes ($p=0.008$ using paired samples Wilcoxon test). In the mountainous West, however, the transport model simulates slightly higher fluxes (Fig S2). Emission uncertainties due to model mechanics including calibration error for the transport model and uncertainties in stream order averaging for the upscaling model cannot alone explain the differences in flux estimates (Fig. 1c). Note that parameter uncertainty is identical between both models and so is not included here (see Methods). Instead, this difference in continental scale fluxes exclusively represents the transport model's ability to reflect source limitations that result in lower CO_2 concentrations in steep environments. This source limitation is demonstrated in Fig 1a, which plots model output distributions from the transport model, statistical model, and the global observational GLORICH dataset (Hartmann et al., 2014) in k_{CO2}-pCO_2 space overlaid on CO_2 flux contours. As also shown in Fig. 1, the transport model provides a closer match to observed k_{CO2}-pCO_2 distributions; the statistical model features higher average CO_2 values for any given k_{CO2} value (and thus, higher fluxes), which is only partially offset by the lack of representation of high pCO_2 values at low k_{CO2} (i.e. reduced y-axis range of the blue contours). Together, these analyses suggest that incorporating realistic carbon source limitations via a hydrologic routing framework results in a significant reduction in total flux estimates relative to statistical models using the same observational constraints. Our estimated 25% reduction in total fluxes, though, is less than previously hypothesized (Rocher-Ros et al., 2019).

Regionally, the transport model predicts that area-normalized inland water fluxes are highest in mountainous regions of the US (Fig. 2). This model result is driven by high k_{CO2} values associated with steep topography coupled to elevated regional pCO_2 observations in the GLORICH dataset. In the transport model, for example, rivers with slopes steeper than 0.03 account for just 11% of stream surface area but contribute 46% of river emissions. The importance of mountainous environments has been previously demonstrated via statistical upscaling estimates (Horgby et al., 2019) and our median mountainous flux rates of 5.3 kgC/m^2/yr are comparable to median fluxes measured across the Swiss Alps of 3.5 kg-C/m^2/yr (Horgby et al., 2019). We note that the continental-scale map in Fig. 2 visually overrepresents first order stream reaches with high fluxes (>10 kg-C/m^2/yr) that feature the rapid degassing of groundwater CO_2 in steep terrain. These overall large fluxes simulated in the transport model may in part be due to biases in the GLORICH dataset that may not capture the steepest and most turbulent reaches with lower pCO_2. This bias would lead to overestimates in regional pCO_2 averages in both the transport and statistical models (supplemental text 1.4 and 1.5). We also note that the hydrography underpinning our model is of a higher resolution than previous studies; we include many steep headwater streams that may lead to higher basin-aggregated flux estimates. Further, many of these headwaters are non-perennial streams (Brinkerhoff et al., 2024a), which are a known uncertainty in global inland water CO2 emission estimates (Bretz et al., 2023; Lauerwald et al., 2023b) and may be underestimated globally (Keller et al., 2021; López-Rojo et al., 2024).While mountainous environments may feature reduced organic carbon for respiration, high erosion may provide increased particulate organic carbon substrate from the terrestrial environment (France-Lanord & Derry, 1997; Hilton & West, 2020) for stream corridor respiration.

Regional patterns simulated in the transport model are susceptible to considerable uncertainty, particularly regarding the parameterization of constant groundwater pCO_2 values. We simulate a constant groundwater pCO_2 of 16,000 ppm based on a lack of robust spatial groundwater pCO_2 data products and calibrate hyporheic zone CO_2 transport and water column net respiration within both rivers and lake/reservoirs to match GLORICH pCO_2 values at the HUC4 scale (see Methods). Based on this approach, our simulations do not incorporate direct mechanistic representations of CO_2 production, but instead calibrate CO_2 production parameters (net water column CO_2 production rates and sediment-water-interface pCO_2) within a mechanistic hydrologic framework (groundwater inputs, gas exchange velocity, downstream transport, and turbulent vertical hyporheic exchange) to find the production parameters that best match regionally representative stream CO_2 observations. For example, if groundwater pCO_2 values are correlated with plant productivity via organic matter availability (Brook et al., 1983; Kessler & Harvey, 2001), we would expect lower groundwater pCO_2 values in the mountainous West. While to first order this may result in reduced simulated montane CO_2 fluxes, the model calibration would compensate for this reduced groundwater export with increased stream corridor CO_2 production to best match the observational dataset. We note, however, that constraining spatial variability in groundwater pCO_2 will provide better constraints on total inland water flux and source estimates.

3.2 Sources of inland water CO_2 emissions

Stream corridor sources of CO_2 make up the majority of emissions at the continental scale within the process-based model, especially in the West and in larger rivers (Fig 3a-c). These stream corridor sources, which include subsurface respiration within the benthic zone, hyporheic zone, and riparian subsurface, account for 84% of CO_2 emissions across CONUS, with groundwater inputs accounting for the remaining 16%. We note that as above, these values are sensitive to our assumed groundwater pCO_2; however, for groundwater sources to exceed stream corridor sources would require average groundwater pCO_2 values of >50,000 ppm across CONUS, which is not supported by estimates of spatial soil pCO_2 (Brook et al., 1983; Kessler & Harvey, 2001; Macpherson, 2009) or previous studies that have measured or calculated upland groundwater contributions to stream CO_2 budgets (Kirk & Cohen, 2023; Lupon et al., 2019; Saccardi & Winnick, 2021). Our simulated stream corridor production of CO_2 would require a terrestrial flux of organic carbon to inland waters of ~10 t C km^{-2} yr^{-1} from land surfaces to sustain. This flux is within current estimates of terrestrial dissolved organic carbon exports of 1-85 t C km^{-2} yr^{-1} in temperate and boreal regions (Hope et al., 1994; McCallister & del Giorgio, 2012; T. R. Moore, 2003; Neff & Asner, 2001), which does not include additional particulate organic carbon and riparian zone soil processes that may further contribute to these fluxes, particularly in mountainous regions where physical erosion may enhance terrestrial contributions of particulate organic carbon (Hilton & West, 2020).

Of these stream corridor CO_2 sources, subsurface respiration within the stream corridor environment, facilitated by hyporheic exchange, is the largest simulated source of CO_2 across CONUS, accounting for 82% of all carbon emitted by streams. Relative stream corridor source contributions show an east-west gradient with Western basin contributions averaging 87% compared to mean basin contributions of 57% in the East. Additionally, large rivers have greater proportional contributions from stream corridor subsurface respiration, with first through fifth orders receiving a median of 40%, 71%, 80%, 86%, and 90% of their CO_2 from these sources, respectively (Fig 3d). This is consistent with previous studies that suggest internal CO_2 production becomes increasingly important at higher stream orders (Hotchkiss et al., 2015; Saccardi & Winnick, 2021) as proportional groundwater contributions to discharge decrease with stream size. This large proportion of stream corridor CO_2 contributions aligns with upper estimates from mass balance considerations at the continental scale (~65-80%) (Butman & Raymond, 2011) and with a recent study finding that 87% of CO_2 emissions are sourced from the stream corridor in a 5th order watershed in southeastern coastal plain Florida (Kirk & Cohen, 2023). Notably, while our model parameterizes hyporheic exchange as occurring with the benthic zone of stream environments, the CO_2 exchanged may integrate respiration occurring throughout the stream corridor environment including adjacent riparian zones as represented in Kirk & Cohen (2023) and described in our Methods. Based on model structure, this hyporheic CO_2 functionally represents the excess carbon needed beyond upland groundwater inputs to match regional mean riverine CO_2 concentrations. Within the transport model, net water column respiration accounts for a relatively minor portion of total CO_2 sources at 2%. This estimate is slightly below a previous CONUS estimate of ~4% (Butman & Raymond, 2011), which may be due to our incorporation of primary production into our net water column respiration term (see Methods).

Overall, our finding that stream corridor sources account for the majority of riverine CO_2 emissions is consistent with previous studies that explicitly estimate upland groundwater CO_2 inputs to aquatic carbon budgets (Butman & Raymond, 2011; Kirk & Cohen, 2023; Rasilo et al., 2017; Saccardi & Winnick, 2021). However, our modeled stream corridor CO_2 production rates are significantly elevated relative to dissolved oxygen-based stream metabolism methods. For example, we simulate an average CONUS stream corridor net CO_2 production rate of ~5.4 $gC/m^2/d$ compared to median US stream metabolism NEP rates of 0.54 $gC/m^2/d$ (Bernhardt et al., 2022). Similarly, our estimate that 84% of CONUS riverine emissions reflect stream corridor respiration is significantly larger than Hotchkiss et al. (2015), who estimate that internally produced CO_2 contributes 14% of emissions in small streams ($<0.01 m^3 s^{-1}$) and only 25-54% in large streams ($>100 m^3\ s^{-1}$) based on the difference between oxygen-based NEP and total CO_2 fluxes.

Interestingly, these stream metabolism estimates (e.g. Appling et al., 2018; Battin et al., 2023; Bernhardt et al., 2022; Hotchkiss et al., 2015) attribute oxygen under-saturation solely to in-stream respiration, which potentially neglects inputs of low-oxygen groundwater associated with terrestrial respiration (e.g. Hall Jr. & Tank, 2005). Hotchkiss et al. (2015) and Kirk and Cohen (2023), for example, attribute measured CO_2 emissions in excess of molar-equivalent oxygen uptake as reflecting groundwater and riparian zone CO_2 inputs, respectively, with no associated oxygen deficit. Implicitly, this assumes that stream measurements of CO_2 capture external terrestrial and near-stream inputs while oxygen measurements do not. While carbonate buffering reactions may allow for the retention of CO_2 signals from discrete groundwater inputs for longer than dissolved oxygen signals (Stets et al., 2017) and may therefore integrate more upstream heterogeneity in production/input rates (Shangguan et al., 2024), these length scales are relatively small and do not impact steady-state CO_2 versus dissolved oxygen concentrations in the case of diffuse groundwater inputs (Winnick & Saccardi, 2024). Notably, the explicit consideration of groundwater and near-stream oxygen deficits in stream metabolism budgets would likely increase the discrepancies between these carbon budgets based on dissolved oxygen versus ones that estimate upland groundwater contributions. Thus, reconciling our stream corridor respiration rates with stream metabolism measurements would require groundwater inputs to feature both extremely high pCO$_2$ (~50,000 ppm to switch from stream corridor to groundwater-dominated fluxes and likely ~100,000 ppm to match median NEP observations) and near-atmospheric dissolved oxygen, which is not consistent with terrestrial respiration.

This apparent paradox is reflective of what we see as a major gap between carbon budgets based on CO_2 measurements versus dissolved oxygen measurements, which to our knowledge has not been previously articulated. As noted above, this gap is best represented by the fact that global inland carbon fluxes estimated from oxygen variations are only ~18% of carbon fluxes estimated from CO_2 concentrations (Battin et al., 2023; Lauerwald et al., 2023b). While beyond the scope of this manuscript, this gap may reflect (1) systematic underestimates of carbon fluxes from oxygen variations, which may in part reflect metabolic study designs that seek to avoid reaches with discrete groundwater inputs; (2) systematic overestimates of carbon fluxes from CO_2 variations; or (3) processes that significantly alter molar ratios of dissolved CO_2:O_2 such as carbonate buffering, alternative metabolic pathways including nitrification, denitrification, and methanogenesis, among others. This gap warrants further investigation, though we stress that despite being significantly larger than metabolism-based NEP, our stream corridor source contributions are consistent with other CO_2 budget-based estimates (Butman & Raymond, 2011; Kirk & Cohen, 2023; Rasilo et al., 2017).

Finally, our initial modeling confirms that rivers are the major sites of emission and are responsible for 94% of all emissions in the transport model results. Headwaters (first order streams) account for 15% of the river surface area but contribute 30% of total river CO_2 emissions (Fig. 3d). Larger rivers (fifth through eleventh orders) account for 55% of the stream surface area but only contribute 34% of total river CO_2 emissions (Fig. 3d). This trend has been noted in other studies which find that first order streams are 7% of the surface area and 25% of river CO_2 emissions (Raymond et al., 2013).

Lakes and reservoirs contribute 6% of the modeled CONUS CO_2 emissions and, on average across individual basins, contribute 9% of a basin's CO_2 emissions (Fig. 4). These numbers are smaller than previous estimates, as 1) we do not include the Great Lakes in our analysis, 2) we do not rely on statistical distributions for extrapolating pond sizes instead we use the NHD-HR which includes lakes down to 1 m^2, and 3) we explicitly account for river/lake connectivity to avoid double counting of lakes as rivers. Additionally, we calibrate lakes using

HUC2 watersheds which are smaller and more representative of local conditions than earlier estimates as they do not extend to boreal and tropical regions which have elevated pCO_2 in comparison to temperate regions (Sobek et al., 2005). Lakes and reservoirs exert a significant influence on CO_2 emissions in lake-dense regions with high water tables. For example, 88% of emissions in south/central Florida and 70% in the Boundary Waters region come from lakes/reservoirs (Fig. 4). This trend is shown across CONUS, as percent lake emission contributions trend with the natural log of the total lake area per basin (R=0.45). Beyond the total lake area, field studies have noted that small lakes contribute proportionally more CO_2 emissions than larger ones (Bogard & del Giorgio, 2016; Holgerson & Raymond, 2016; Schmadel et al., 2019) due to their larger lakebed surface area to water volume ratio. This effect is simulated in our model, which explicitly represents these morphometric differences across lakes: small ponds (here defined as 0-0.1 km^2) are only 11% of the total CONUS lake/reservoir surface area but are responsible for 65% of lake/reservoir CO_2 emissions.

Despite considerable uncertainty within our CONUS CO_2 emissions estimates, CO_2 production parameterizations, and the associated breakdown of source contributions, the major takeaways from our analysis are unlikely to change. Specifically, (1) river emissions are an order of magnitude higher than lake/reservoir emissions at the continental scale, with some level of geographic variability associated with regional water table dynamics; (2) respiration within the subsurface stream corridor environment is the largest source of inland water CO_2 emissions, followed by groundwater, with net water column respiration that accounts for the balance of respiration and primary production contributing a minor proportion of total emissions. While variability in groundwater pCO_2 may alter regional partitioning estimates, average CONUS groundwater pCO_2 would have to be >50,000 ppm to account for >50% of riverine CO_2 fluxes assuming a 1:1 tradeoff in estimated groundwater v. stream corridor CO_2 inputs and even higher to reconcile stream corridor respiration rates with oxygen-based NEP measurements. These elevated values appear unrealistic for the continental scale (Brook et al., 1983; Kessler & Harvey, 2001; Macpherson, 2009); however, the fundamental mismatch between carbon budgets based on CO_2 fluxes versus those based on dissolved oxygen discussed above represents a significant uncertainty that should be investigated further. Taken together, our results suggest that the largest potential carbon cycle feedback mechanisms relate to hydraulic flow dynamics, which in turn alter terrestrial-aquatic connectivity, hyporheic exchange, and the export of terrestrial organic carbon that supports net aquatic respiration.

4 Towards forward predictive models of CO_2 emissions

Our application of a hydrologic transport framework coupled to CO_2 production rates represents a step towards fully integrating hydrologic and biogeochemical models at continental and global scales to predict inland water CO_2 fluxes. Importantly, the presented framework provides a pathway to interrogate the mechanistic impacts of hydrology on flux estimates through direct representation of groundwater inputs, advection velocities, gas exchange velocities, and hyporheic exchange rates at stream reach scales. We emphasize that our results demonstrate the impacts of representing transport dynamics on estimates of fluxes and sources given the same data constraints as statistical upscaling models, and are not yet at the level of providing robust forward predications of inland water CO_2 fluxes.

Despite this progress, our ability to apply these models globally is still limited by a few issues. First is the lack of sufficient headwater representation in global hydrography data products Spatial resolution of digital elevation models and remotely sensed imagery present a lower limit to the small streams we can observe and this has downstream effects on our ability to model solute exchange along river networks (Brinkerhoff 2024). Additionally, the majority of these small streams are non-perennial (Brinkerhoff et al., 2024; Messager et al., 2021), meaning they do not flow year-round and represent a critical nexus for terrestrial recruitment of solutes (Benstead & Leigh, 2012), including terrestrially-produced CO_2 (Gómez-Gener et al., 2016; Silverthorn et al., 2024).

Second, the largest barrier for moving towards more accurate continental-scale CO_2 flux estimates is the paucity of observational datasets. This is particularly true for streams with the steepest topography as discussed above, which may lead to overestimates in CO_2 fluxes from mountainous environments. Additionally, while recent advances have allowed for the direct measurements of pCO_2 in surface waters, most published data including the GLORICH database used in our model calibration is based on carbonate speciation calculations using measured pH and alkalinity. Previous studies have shown that these methods are subject to significant error, particularly under low pH conditions (Abril et al., 2015; Raymond et al., 2013). While we have sought to minimize this potential error via filtering (Section 2.1), a cursory comparison of GLORICH data to the direct CO_2 measurements used in Liu et al. (2022) suggests a potential overestimate of mean pCO_2 based on speciation calculations (Supplementary Information); however, differences between GLORICH and Liu et al. (2022) data are not statistically significant given the large standard deviation of GLORICH values, and this difference is not present when comparing HUC2-averaged values with the Liu et al. (2022) dataset. While the potential for artificially high pCO_2 may lead to lower total estimated fluxes as well as lower contributions from stream corridor respiration given the same parameterized groundwater CO_2 inputs, we note that these reductions in total fluxes are similar for the transport and statistical models (SI).

At present, our ability to represent groundwater CO_2 inputs is also limited by the lack of publicly available large-scale spatial groundwater chemistry data products and is thus a top priority for providing more accurate regional flux and source partitioning estimates. In particular, groundwater CO_2 and dissolved oxygen datasets will be crucial to evaluating the large discrepancies between carbon source partitioning estimates from CO_2 measurements versus stream metabolism calculations. As described above, robust spatially- and temporally-variable groundwater CO_2 datasets would allow for both more robust flux estimates and source distributions within the presented calibration framework, and could also allow for predictive forward modelling with independently validated carbon input variables. We also note that while our calibration framework is flexible to incorporate additional CO_2 inputs from connected wetland environments, provided they are adequately represented in the observational calibration datasets, these fluxes are tied via calibration to hyporheic exchange rates rather than groundwater input rates based on our current model framework. Future work is necessary to account for wetland-impacted groundwater input rates which have been shown to scale with degree of wetland connectivity across CONUS (Leibowitz et al., 2023).

While the expansion of observational datasets is vital to providing accurate and validated estimates of average inland water CO_2 emissions, forward predictions of emission fluxes will further require scalable biogeochemical models that capture spatiotemporal variability in carbon transformations. As noted, while our transport model incorporates direct estimates of advective transport, groundwater inflow rates, gas exchange, and hyporheic exchange as a function of geomorphology and flow conditions, the CO_2 concentrations associated with groundwater, hyporheic exchange, and in-stream processing are currently estimated and calibrated to observations. Recently, carbonate buffering dynamics have been incorporated into similar stream network carbon frameworks (Winnick & Saccardi, 2024), and may help to interrogate differences in oxygen- versus CO_2-based carbon budgets. However, models that can accurately predict in-stream metabolism, terrestrial carbon exports via groundwater, and hyporheic zone processing across lotic and lentic environments with limited or coarse-resolution substrate data remain elusive are an important avenue towards predicting the response of inland water CO_2 emissions to anthropogenic climate change.

Mechanistic biogeochemical models will also allow for estimating estimating temporal variability in CONUS-level CO_2 dynamics, which may allow for more accurate total flux estimates. Specifically, studies suggest variable and non-linear changes in CO_2 concentrations and fluxes in response to hydrologic changes including storm events (Aho & Raymond, 2019; Conroy et al., 2023; Crawford et al., 2017; Dinsmore et al., 2013; Dinsmore & Billett, 2008; Duvert et al., 2018). Thus, estimates of CO_2 emissions under mean annual flow conditions may not represent mean CO_2 fluxes that integrate temporal variability. Though our modeling framework can simulate the impacts of hydrologic variability on its own in terms of groundwater inputs, hyporheic exchange rates, and gas exchange rates, we cannot presently account for temporal changes in CO_2 production parameters. As it relates to observational datasets that would allow for time-dependent calibration of CO_2 production parameters, this limitation is unlikely to be addressed in the near future. Instead, the potential for providing time-variable simulations relies on either *(1)* the incorporation of process-based models for stream metabolism and groundwater CO_2 variability; or *(2)* the application of machine learning techniques to provide time-varying estimates of these parameters.

Acknowledgments

We thank the USGS and many other workers for making their data and models freely available. CB was supported by a NASA FINESST fellowship (80NSSC21K1591). This work was also support by NSF awards EAR-2103520 to MJW and EAR-2318056 to MJW and CJG, as well as NASA award 80NSSC20K1141 to CJG. We thank Eric Davidson, Robert Hall, and two anonymous reviewers for comments and suggestions that improved an earlier version of this manuscript.

Conflict of Interest

The authors declare no conflict of interests relevant to this study.

Open Research

Code used to run the model, generate results, and build figures is archived at https://zenodo.org/records/13144302 (Brinkerhoff et al., 2024b).

References

Abril, G., Bouillon, S., Darchambeau, F., Teodoru, C. R., Marwick, T. R., Tamooh, F., Ochieng Omengo, F., Geeraert, N., Deirmendjian, L., Polsenaere, P., & Borges, A. V. (2015). Technical Note: Large overestimation of pCO_2 calculated from pH and alkalinity in acidic, organic-rich freshwaters. *Biogeosciences*, *12*(1), 67–78. https://doi.org/10.5194/bg-12-67-2015

Aho, K. S., & Raymond, P. A. (2019). Differential Response of Greenhouse Gas Evasion to Storms in Forested and Wetland Streams. *Journal of Geophysical Research: Biogeosciences*, *124*(3), 649–662. https://doi.org/10.1029/2018JG004750

Allen, G. H., & Pavelsky, T. M. (2018). Global extent of rivers and streams. *Science*, *361*(6402), 585–588. https://doi.org/10.1126/science.aat0636

Allen, G. H., Pavelsky, T. M., Barefoot, E. A., Lamb, M. P., Butman, D., Tashie, A., & Gleason, C. J. (2018). Similarity of stream width distributions across headwater systems. *Nature Communications*, *9*(1), 610. https://doi.org/10.1038/s41467-018-02991-w

Appling, A. P., Hall Jr., R. O., Yackulic, C. B., & Arroita, M. (2018). Overcoming Equifinality: Leveraging Long Time Series for Stream Metabolism Estimation. *Journal of Geophysical Research: Biogeosciences*, *123*(2), 624–645. https://doi.org/10.1002/2017JG004140

Battin, T. J., Lauerwald, R., Bernhardt, E. S., Bertuzzo, E., Gener, L. G., Hall, R. O., Hotchkiss, E. R., Maavara, T., Pavelsky, T. M., Ran, L., Raymond, P., Rosentreter, J. A., & Regnier, P. (2023). River ecosystem metabolism and carbon biogeochemistry in a changing world. *Nature*, *613*(7944), Article 7944. https://doi.org/10.1038/s41586-022-05500-8

Bencala, K. E., & Walters, R. A. (1983). Simulation of solute transport in a mountain pool-and-riffle stream: A transient storage model. *Water Resources Research*, *19*(3), 718–724. https://doi.org/10.1029/WR019i003p00718

Benstead, J. P., & Leigh, D. S. (2012). An expanded role for river networks. *Nature Geoscience*, *5*(10), 678–679. https://doi.org/10.1038/ngeo1593

Bernhardt, E. S., Savoy, P., Vlah, M. J., Appling, A. P., Koenig, L. E., Hall, R. O., Arroita, M., Blaszczak, J. R., Carter, A. M., Cohen, M., Harvey, J. W., Heffernan, J. B., Helton, A. M., Hosen, J. D., Kirk, L., McDowell, W. H., Stanley, E. H., Yackulic, C. B., & Grimm, N. B. (2022). Light and flow regimes regulate the metabolism of rivers. *Proceedings of the National Academy of Sciences*, *119*(8), e2121976119. https://doi.org/10.1073/pnas.2121976119

Beven, K. (1993). Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in Water Resources*, *16*(1), 41–51. https://doi.org/10.1016/0309-1708(93)90028-E

Bogard, M. J., & del Giorgio, P. A. (2016). The role of metabolism in modulating CO2 fluxes in boreal lakes. *Global Biogeochemical Cycles*, *30*(10), 1509–1525. https://doi.org/10.1002/2016GB005463

Borges, A. V., Darchambeau, F., Teodoru, C. R., Marwick, T. R., Tamooh, F., Geeraert, N., Omengo, F. O., Guérin, F., Lambert, T., Morana, C., Okuku, E., & Bouillon, S. (2015).

Globally significant greenhouse-gas emissions from African inland waters. *Nature Geoscience*, *8*(8), Article 8. https://doi.org/10.1038/ngeo2486

Bretz, K. A., Murphy, N. N., & Hotchkiss, E. R. (2023). Carbon Biogeochemistry and Export Governed by Flow in a Non-Perennial Stream. *Water Resources Research*, *59*(9), e2022WR034004. https://doi.org/10.1029/2022WR034004

Brinkerhoff, C. B. (2024). The importance of source data in river network connectivity modeling: A review. *Limnology and Oceanography*. https://doi.org/10.1002/lno.12706

Brinkerhoff, C. B., Gleason, C. J., Kotchen, M. J., Kysar, D. A., & Raymond, P. A. (2024). Ephemeral stream water contributions to United States drainage networks. *Science*, *384*(6703), 1476–1482. https://doi.org/10.1126/science.adg9430

Brinkerhoff, C. B., Gleason, C. J., Zappa, C. J., Raymond, P. A., & Harlan, M. E. (2022). Remotely sensing river greenhouse gas exchange velocity using the SWOT satellite. *Global Biogeochemical Cycles*, *36*(10), e2022GB007419.

Brinkerhoff, C. B., Raymond, P. A., Maavara, T., Ishitsuka, Y., Aho, K. s., & Gleason, C. J. (2021). Lake Morphometry and River Network Controls on Evasion of Terrestrially Sourced Headwater CO2. *Geophysical Research Letters*, *48*(1), e2020GL090068. https://doi.org/10.1029/2020GL090068

Brook, G. A., Folkoff, M. E., & Box, E. O. (1983). A world model of soil carbon dioxide. *Earth Surface Processes and Landforms*, *8*(1), 79–88. https://doi.org/10.1002/esp.3290080108

Butman, D., & Raymond, P. A. (2011). Significant efflux of carbon dioxide from streams and rivers in the United States. *Nature Geoscience*, *4*(12), Article 12. https://doi.org/10.1038/ngeo1294

Butman, D., Stackpoole, S., Stets, E., McDonald, C. P., Clow, D. W., & Striegl, R. G. (2016). Aquatic carbon cycling in the conterminous United States and implications for terrestrial carbon accounting. *Proceedings of the National Academy of Sciences*, *113*(1), 58–63. https://doi.org/10.1073/pnas.1512651112

Cavallaro, N., Shrestha, G., Birdsey, R., Mayes, M. A., Najjar, R. G., Reed, S. C., Romero-Lankao, P., & Zhu, Z. (2018). *Second State of the Carbon Cycle Report*. U.S. Global Change Research Program. https://doi.org/10.7930/Soccr2.2018

Conroy, H. D., Hotchkiss, E. R., Cawley, K. M., Goodman, K., Hall Jr., R. O., Jones, J. B., Wollheim, W. M., & Butman, D. (2023). Seasonality Drives Carbon Emissions Along a Stream Network. *Journal of Geophysical Research: Biogeosciences*, *128*(8), e2023JG007439. https://doi.org/10.1029/2023JG007439

Crawford, J. T., Stanley, E. H., Dornblaser, M. M., & Striegl, R. G. (2017). CO2 time series patterns in contrasting headwater streams of North America. *Aquatic Sciences*, *79*(3), 473–486. https://doi.org/10.1007/s00027-016-0511-2

Dinsmore, K. J., & Billett, M. F. (2008). Continuous measurement and modeling of CO2 losses from a peatland stream during stormflow events. *Water Resources Research*, *44*(12). https://doi.org/10.1029/2008WR007284

Dinsmore, K. J., Wallin, M. B., Johnson, M. S., Billett, M. F., Bishop, K., Pumpanen, J., & Ojala, A. (2013). Contrasting CO2 concentration discharge dynamics in headwater streams: A multi-catchment comparison. *Journal of Geophysical Research: Biogeosciences*, *118*(2), 445–461. https://doi.org/10.1002/jgrg.20047

Drake, T. W., Raymond, P. A., & Spencer, R. G. M. (2018). Terrestrial carbon inputs to inland waters: A current synthesis of estimates and uncertainty. *Limnology and Oceanography Letters*, *3*(3), 132–142. https://doi.org/10.1002/lol2.10055

Duvert, C., Butman, D. E., Marx, A., Ribolzi, O., & Hutley, L. B. (2018). CO2 evasion along streams driven by groundwater inputs and geomorphic controls. *Nature Geoscience*, *11*(11), Article 11. https://doi.org/10.1038/s41561-018-0245-y

France-Lanord, C., & Derry, L. A. (1997). Organic carbon burial forcing of the carbon cycle from Himalayan erosion. *Nature*, *390*(6655), Article 6655. https://doi.org/10.1038/36324

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Gregor, L., Hauck, J., Le Quéré, C., Luijkx, I. T., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Alkama, R., … Zheng, B. (2022). Global Carbon Budget 2022. *Earth System Science Data*, *14*(11), 4811–4900. https://doi.org/10.5194/essd-14-4811-2022

Gómez-Gener, L., Obrador, B., Marcé, R., Acuña, V., Catalán, N., Casas-Ruiz, J. P., Sabater, S., Muñoz, I., & von Schiller, D. (2016). When Water Vanishes: Magnitude and Regulation of Carbon Dioxide Emissions from Dry Temporary Streams. *Ecosystems*, *19*(4), 710–723. https://doi.org/10.1007/s10021-016-9963-4

Gómez-Gener, L., Rocher-Ros, G., Battin, T., Cohen, M. J., Dalmagro, H. J., Dinsmore, K. J., Drake, T. W., Duvert, C., Enrich-Prast, A., Horgby, Å., Johnson, M. S., Kirk, L., Machado-Silva, F., Marzolf, N. S., McDowell, M. J., McDowell, W. H., Miettinen, H., Ojala, A. K., Peter, H., … Sponseller, R. A. (2021). Global carbon dioxide efflux from rivers enhanced by high nocturnal emissions. *Nature Geoscience*, *14*(5), Article 5. https://doi.org/10.1038/s41561-021-00722-3

Grant, S. B., Azizian, M., Cook, P., Boano, F., & Rippy, M. A. (2018). Factoring stream turbulence into global assessments of nitrogen pollution. *Science*, *359*(6381), 1266–1269. https://doi.org/10.1126/science.aap8074

Grant, S. B., Gomez-Velez, J. D., & Ghisalberti, M. (2018). Modeling the Effects of Turbulence on Hyporheic Exchange and Local-to-Global Nutrient Processing in Streams. *Water Resources Research*, *54*(9), 5883–5889. https://doi.org/10.1029/2018WR023078

Hall Jr., R. O., & Tank, J. L. (2005). Correcting whole-stream estimates of metabolism for groundwater input. *Limnology and Oceanography: Methods*, *3*(4), 222–229. https://doi.org/10.4319/lom.2005.3.222

Hartmann, J., Lauerwald, R., & Moosdorf, N. (2014). A Brief Overview of the GLObal RIver Chemistry Database, GLORICH. *Procedia Earth and Planetary Science*, *10*, 23–27. https://doi.org/10.1016/j.proeps.2014.08.005

Harvey, J., Gomez-Velez, J., Schmadel, N., Scott, D., Boyer, E., Alexander, R., Eng, K., Golden, H., Kettner, A., Konrad, C., Moore, R., Pizzuto, J., Schwarz, G., Soulsby, C., & Choi, J. (2019). How Hydrologic Connectivity Regulates Water Quality in River Corridors. *JAWRA Journal of the American Water Resources Association*, *55*(2), 369–381. https://doi.org/10.1111/1752-1688.12691

Hilton, R. G., & West, A. J. (2020). Mountains, erosion and the carbon cycle. *Nature Reviews Earth & Environment*, *1*(6), Article 6. https://doi.org/10.1038/s43017-020-0058-6

Holgerson, M. A., & Raymond, P. A. (2016). Large contribution to inland water CO2 and CH4 emissions from very small ponds. *Nature Geoscience*, *9*(3), Article 3. https://doi.org/10.1038/ngeo2654

Hope, D., Billett, M. F., & Cresser, M. S. (1994). A review of the export of carbon in river water: Fluxes and processes. *Environmental Pollution*, *84*(3), 301–324. https://doi.org/10.1016/0269-7491(94)90142-2

Horgby, Å., Segatto, P. L., Bertuzzo, E., Lauerwald, R., Lehner, B., Ulseth, A. J., Vennemann, T. W., & Battin, T. J. (2019). Unexpected large evasion fluxes of carbon dioxide from turbulent streams draining the world's mountains. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-12905-z

Hotchkiss, E. R., Hall Jr, R. O., Sponseller, R. A., Butman, D., Klaminder, J., Laudon, H., Rosvall, M., & Karlsson, J. (2015). Sources of and processes controlling CO2 emissions change with the size of streams and rivers. *Nature Geoscience*, *8*(9), Article 9. https://doi.org/10.1038/ngeo2507

Johnson, M. S., Lehmann, J., Riha, S. J., Krusche, A. V., Richey, J. E., Ometto, J. P. H. B., & Couto, E. G. (2008). CO2 efflux from Amazonian headwater streams represents a significant fate for deep soil respiration. *Geophysical Research Letters*, *35*(17). https://doi.org/10.1029/2008GL034619

Keenan, T. F., & Williams, C. A. (2018). The Terrestrial Carbon Sink. *Annual Review of Environment and Resources*, *43*(1), 219–243. https://doi.org/10.1146/annurev-environ-102017-030204

Keller, P. S., Marcé, R., Obrador, B., & Koschorreck, M. (2021). Global carbon budget of reservoirs is overturned by the quantification of drawdown areas. *Nature Geoscience*, *14*(6), 402–408. https://doi.org/10.1038/s41561-021-00734-z

Kessler, T. J., & Harvey, C. F. (2001). The global flux of carbon dioxide into groundwater. *Geophysical Research Letters*, *28*(2), 279–282. https://doi.org/10.1029/2000GL011505

Kirk, L., & Cohen, M. J. (2023). River Corridor Sources Dominate CO2 Emissions From a Lowland River Network. *Journal of Geophysical Research: Biogeosciences*, *128*(1), e2022JG006954. https://doi.org/10.1029/2022JG006954

Lauerwald, R., Allen, G. H., Deemer, B. R., Liu, S., Maavara, T., Raymond, P., Alcott, L., Bastviken, D., Hastie, A., Holgerson, M. A., Johnson, M. S., Lehner, B., Lin, P., Marzadri, A., Ran, L., Tian, H., Yang, X., Yao, Y., & Regnier, P. (2023a). Inland Water Greenhouse Gas Budgets for RECCAP2: 1. State-Of-The-Art of Global Scale Assessments. *Global Biogeochemical Cycles*, *37*(5), e2022GB007657. https://doi.org/10.1029/2022GB007657

Lauerwald, R., Allen, G. H., Deemer, B. R., Liu, S., Maavara, T., Raymond, P., Alcott, L., Bastviken, D., Hastie, A., Holgerson, M. A., Johnson, M. S., Lehner, B., Lin, P., Marzadri, A., Ran, L., Tian, H., Yang, X., Yao, Y., & Regnier, P. (2023b). Inland Water Greenhouse Gas Budgets for RECCAP2: 2. Regionalization and Homogenization of Estimates. *Global Biogeochemical Cycles*, *37*(5), e2022GB007658. https://doi.org/10.1029/2022GB007658

Lauerwald, R., Laruelle, G. G., Hartmann, J., Ciais, P., & Regnier, P. A. G. (2015). Spatial patterns in CO2 evasion from the global river network. *Global Biogeochemical Cycles*, *29*(5), 534–554. https://doi.org/10.1002/2014GB004941

Lehner, B., & Döll, P. (2004). Development and validation of a global database of lakes, reservoirs and wetlands. *Journal of Hydrology*, *296*(1), 1–22. https://doi.org/10.1016/j.jhydrol.2004.03.028

Lehner, B., Verdin, K., & Jarvis, A. (2008). New Global Hydrography Derived From Spaceborne Elevation Data. *EOS Transactions*, *89*, 93–94. https://doi.org/10.1029/2008EO100001

Leibowitz, S. G., Hill, R. A., Creed, I. F., Compton, J. E., Golden, H. E., Weber, M. H., Rains, M. C., Jones, C. E., Lee, E. H., Christensen, J. R., Bellmore, R. A., & Lane, C. R. (2023). National hydrologic connectivity classification links wetlands with stream water quality. *Nature Water*, *1*(4), 370–380. https://doi.org/10.1038/s44221-023-00057-w

Lin, P., Pan, M., Wood, E. F., Yamazaki, D., & Allen, G. H. (2021). A new vector-based global river network dataset accounting for variable drainage density. *Scientific Data*, *8*(1), Article 1. https://doi.org/10.1038/s41597-021-00819-9

Liu, S., Butman, D. E., & Raymond, P. A. (2020). Evaluating CO2 calculation error from organic alkalinity and pH measurement error in low ionic strength freshwaters. *Limnology and Oceanography: Methods*, *18*(10), 606–622. https://doi.org/10.1002/lom3.10388

Liu, S., Kuhn, C., Amatulli, G., Aho, K., Butman, D. E., Allen, G. H., Lin, P., Pan, M., Yamazaki, D., Brinkerhoff, C., Gleason, C., Xia, X., & 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. https://doi.org/10.1073/pnas.2106322119

Liu, S., Maavara, T., Brinkerhoff, C. B., & Raymond, P. A. (2022). Global Controls on DOC Reaction Versus Export in Watersheds: A Damköhler Number Analysis. *Global Biogeochemical Cycles*, *36*(4), e2021GB007278. https://doi.org/10.1029/2021GB007278

López-Rojo, N., Datry, T., Peñas, F. J., Singer, G., Lamouroux, N., Barquín, J., Rodeles, A. A., Silverthorn, T., Sarremejane, R., del Campo, R., Estévez, E., Mimeau, L., Boyer, F., Künne, A., Dalvai Ragnoli, M., & Foulquier, A. (2024). Carbon emissions from inland waters may be underestimated: Evidence from European river networks fragmented by drying. *Limnology and Oceanography Letters*, *n/a*(n/a). https://doi.org/10.1002/lol2.10408

Lorke, A., & Peeters, F. (2006). Toward a Unified Scaling Relation for Interfacial Fluxes. *Journal of Physical Oceanography*, *36*(5), 955–961. https://doi.org/10.1175/JPO2903.1

Lupon, A., Denfeld, B. A., Laudon, H., Leach, J., Karlsson, J., & Sponseller, R. A. (2019). Groundwater inflows control patterns and sources of greenhouse gas emissions from streams. *Limnology and Oceanography*, *64*(4), 1545–1557. https://doi.org/10.1002/lno.11134

Maavara, T., Brinkerhoff, C., Hosen, J., Aho, K., Logozzo, L., Saiers, J., Stubbins, A., & Raymond, P. (2023). Watershed DOC uptake occurs mostly in lakes in the summer and in rivers in the winter. *Limnology and Oceanography*, *68*(3), 735–751. https://doi.org/10.1002/lno.12306

Macpherson, G. L. (2009). CO2 distribution in groundwater and the impact of groundwater extraction on the global C cycle. *Chemical Geology*, *264*(1), 328–336. https://doi.org/10.1016/j.chemgeo.2009.03.018

Mayorga, E., Seitzinger, S. P., Harrison, J. A., Dumont, E., Beusen, A. H. W., Bouwman, A. F., Fekete, B. M., Kroeze, C., & Van Drecht, G. (2010). Global Nutrient Export from WaterSheds 2 (NEWS 2): Model development and implementation. *Environmental Modelling & Software*, *25*(7), 837–853. https://doi.org/10.1016/j.envsoft.2010.01.007

McCallister, S. L., & del Giorgio, P. A. (2012). Evidence for the respiration of ancient terrestrial organic C in northern temperate lakes and streams. *Proceedings of the National Academy of Sciences*, *109*(42), 16963–16968. https://doi.org/10.1073/pnas.1207305109

Messager, M. L., Lehner, B., Cockburn, C., Lamouroux, N., Pella, H., Snelder, T., Tockner, K., Trautmann, T., Watt, C., & Datry, T. (2021). Global prevalence of non-perennial rivers and streams. *Nature*, *594*(7863), 391–397. https://doi.org/10.1038/s41586-021-03565-5

Messager, M. L., Lehner, B., Grill, G., Nedeva, I., & Schmitt, O. (2016). Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nature Communications*, *7*(1), 13603. https://doi.org/10.1038/ncomms13603

Mitchell, M. (1998). *An Introduction to Genetic Algorithms*. MIT Press. https://mitpress.mit.edu/9780262631853/an-introduction-to-genetic-algorithms/

Moore, R. B., McKay, L. D., Rea, A. H., Bondelid, T. R., Price, C. V., Dewald, T. G., & Johnston, C. M. (2019). User's guide for the national hydrography dataset plus (NHDPlus) high resolution. In *Open-File Report* (2019–1096). U.S. Geological Survey. https://doi.org/10.3133/ofr20191096

Moore, T. R. (2003). Dissolved organic carbon in a northern boreal landscape. *Global Biogeochemical Cycles*, *17*(4). https://doi.org/10.1029/2003GB002050

Neff, J. C., & Asner, G. P. (2001). Dissolved Organic Carbon in Terrestrial Ecosystems: Synthesis and a Model. *Ecosystems*, *4*(1), 29–48. https://doi.org/10.1007/s100210000058

Rasilo, T., Hutchins, R. H. S., Ruiz-González, C., & del Giorgio, P. A. (2017). Transport and transformation of soil-derived CO2, CH4 and DOC sustain CO2 supersaturation in small boreal streams. *Science of The Total Environment*, *579*, 902–912. https://doi.org/10.1016/j.scitotenv.2016.10.187

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*(7476), Article 7476. 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*(1), 41–53. https://doi.org/10.1215/21573689-1597669

Read, J. S., Hamilton, D. P., Desai, A. R., Rose, K. C., MacIntyre, S., Lenters, J. D., Smyth, R. L., Hanson, P. C., Cole, J. J., Staehr, P. A., Rusak, J. A., Pierson, D. C., Brookes, J. D., Laas, A., & Wu, C. H. (2012). Lake-size dependency of wind shear and convection as controls on gas exchange. *Geophysical Research Letters*, *39*(9). https://doi.org/10.1029/2012GL051886

Rocher-Ros, G., Sponseller, R. A., Lidberg, W., Mörth, C.-M., & Giesler, R. (2019). Landscape process domains drive patterns of CO2 evasion from river networks. *Limnology and Oceanography Letters*, *4*(4), 87–95. https://doi.org/10.1002/lol2.10108

Rocher-Ros, G., Stanley, E. H., Loken, L. C., Casson, N. J., Raymond, P. A., Liu, S., Amatulli, G., & Sponseller, R. A. (2023). Global methane emissions from rivers and streams. *Nature*, *621*(7979), 530–535.

Saccardi, B., & Winnick, M. (2021). Improving Predictions of Stream CO2 Concentrations and Fluxes Using a Stream Network Model: A Case Study in the East River Watershed, CO, USA. *Global Biogeochemical Cycles*, *35*(12), e2021GB006972. https://doi.org/10.1029/2021GB006972

Sawakuchi, H. O., Neu, V., Ward, N. D., Barros, M. de L. C., Valerio, A. M., Gagne-Maynard, W., Cunha, A. C., Less, D. F. S., Diniz, J. E. M., Brito, D. C., Krusche, A. V., & Richey,

J. E. (2017). Carbon Dioxide Emissions along the Lower Amazon River. *Frontiers in Marine Science*, *4*. https://www.frontiersin.org/articles/10.3389/fmars.2017.00076

Schmadel, N. M., Harvey, J. W., Alexander, R. B., Schwarz, G. E., Moore, R. B., Eng, K., Gomez-Velez, J. D., Boyer, E. W., & Scott, D. (2018). Thresholds of lake and reservoir connectivity in river networks control nitrogen removal. *Nature Communications*, *9*(1), Article 1. https://doi.org/10.1038/s41467-018-05156-x

Schmadel, N. M., Harvey, J. W., Schwarz, G. E., Alexander, R. B., Gomez-Velez, J. D., Scott, D., & Ator, S. W. (2019). Small Ponds in Headwater Catchments Are a Dominant Influence on Regional Nutrient and Sediment Budgets. *Geophysical Research Letters*, *46*(16), 9669–9677. https://doi.org/10.1029/2019GL083937

Segatto, P. L., Battin, T. J., & Bertuzzo, E. (2023). A Network-Scale Modeling Framework for Stream Metabolism, Ecosystem Efficiency, and Their Response to Climate Change. *Water Resources Research*, *59*(3), e2022WR034062. https://doi.org/10.1029/2022WR034062

Shangguan, Q., Payn, R. A., Hall Jr, R. O., Young, F. L., Valett, H. M., & DeGrandpre, M. D. (2024). Divergent metabolism estimates from dissolved oxygen and inorganic carbon: Implications for river carbon cycling. *Limnology and Oceanography*, *69*(9), 2211–2228. https://doi.org/10.1002/lno.12666

Sikder, Md., Wang, J., Allen, G., Sheng, Y., Yamazaki, D., Cretaux, J.-F., & Pavelsky, T. (2021). *A Global-Scale Lake Topology for Harmonizing SWOT A Priori Lake and River Databases*. *2021*, H12I-01. AGU Fall Meeting Abstracts.

Silverthorn, T., López-Rojo, N., Sarremejane, R., Foulquier, A., Chanudet, V., Azougui, A., del Campo, R., Singer, G., & Datry, T. (2024). River network-scale drying impacts the spatiotemporal dynamics of greenhouse gas fluxes. *Limnology and Oceanography*, *69*(4), 861–873. https://doi.org/10.1002/lno.12531

Sobek, S., Tranvik, L. J., & Cole, J. J. (2005). Temperature independence of carbon dioxide supersaturation in global lakes. *Global Biogeochemical Cycles*, *19*(2). https://doi.org/10.1029/2004GB002264

Stets, E. G., Butman, D., McDonald, C. P., Stackpoole, S. M., DeGrandpre, M. D., & Striegl, R. G. (2017). Carbonate buffering and metabolic controls on carbon dioxide in rivers. *Global Biogeochemical Cycles*, *31*(4), 663–677. https://doi.org/10.1002/2016GB005578

Ulseth, A. J., Hall, R. O., Boix Canadell, M., Madinger, H. L., Niayifar, A., & Battin, T. J. (2019). Distinct air–water gas exchange regimes in low- and high-energy streams. *Nature Geoscience*, *12*(4), Article 4. https://doi.org/10.1038/s41561-019-0324-8

Wang, J., Walter, B. A., Yao, F., Song, C., Ding, M., Maroof, A. S., Zhu, J., Fan, C., McAlister, J. M., Sikder, S., Sheng, Y., Allen, G. H., Crétaux, J.-F., & Wada, Y. (2022). GeoDAR: Georeferenced global dams and reservoirs dataset for bridging attributes and geolocations. *Earth System Science Data*, *14*(4), 1869–1899. https://doi.org/10.5194/essd-14-1869-2022

Winnick, M. J. (2021). Stream Transport and Substrate Controls on Nitrous Oxide Yields From Hyporheic Zone Denitrification. *AGU Advances*, *2*(4), e2021AV000517. https://doi.org/10.1029/2021AV000517

Winnick, M. J., & Saccardi, B. (2024). Impacts of Carbonate Buffering on Atmospheric Equilibration of CO2, δ13CDIC, and Δ14CDIC in Rivers and Streams. *Global Biogeochemical Cycles*, *38*(2), e2023GB007860. https://doi.org/10.1029/2023GB007860

Supporting Information References

Abril, G., Bouillon, S., Darchambeau, F., Teodoru, C. R., Marwick, T. R., Tamooh, F., Ochieng Omengo, F., Geeraert, N., Deirmendjian, L., Polsenaere, P., & Borges, A. V. (2015). Technical Note: Large overestimation of pCO_2 calculated from pH and alkalinity in acidic, organic-rich freshwaters. *Biogeosciences*, *12*(1), 67–78. https://doi.org/10.5194/bg-12-67-2015

Allen, G. H., Pavelsky, T. M., Barefoot, E. A., Lamb, M. P., Butman, D., Tashie, A., & Gleason, C. J. (2018). Similarity of stream width distributions across headwater systems. *Nature Communications*, *9*(1), 610. https://doi.org/10.1038/s41467-018-02991-w

Boodoo, K. S., Schelker, J., Trauth, N., Battin, T. J., & Schmidt, C. (2019). Sources and variability of CO_2 in a prealpine stream gravel bar. *Hydrological Processes*, *33*(17), 2279–2299. https://doi.org/10.1002/hyp.13450

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*(16), 9637–9647. https://doi.org/10.1029/2019GL084529

Brinkerhoff, C. B., Gleason, C. J., Zappa, C. J., Raymond, P. A., & Harlan, M. E. (2022). Remotely sensing river greenhouse gas exchange velocity using the SWOT satellite. *Global Biogeochemical Cycles*, *36*(10), e2022GB007419.

Brinkerhoff, C. B., Raymond, P. A., Maavara, T., Ishitsuka, Y., Aho, K. s., & Gleason, C. J. (2021). Lake Morphometry and River Network Controls on Evasion of Terrestrially Sourced Headwater CO2. *Geophysical Research Letters*, *48*(1), e2020GL090068. https://doi.org/10.1029/2020GL090068

Butman, D., & Raymond, P. A. (2011). Significant efflux of carbon dioxide from streams and rivers in the United States. *Nature Geoscience*, *4*(12), Article 12. https://doi.org/10.1038/ngeo1294

Butman, D., Stackpoole, S., Stets, E., McDonald, C. P., Clow, D. W., & Striegl, R. G. (2016). Aquatic carbon cycling in the conterminous United States and implications for terrestrial carbon accounting. *Proceedings of the National Academy of Sciences*, *113*(1), 58–63. https://doi.org/10.1073/pnas.1512651112

Cael, B. B., & Seekell, D. A. (2016). The size-distribution of Earth's lakes. *Scientific Reports*, *6*(1), 29633. https://doi.org/10.1038/srep29633

Donald D. Adams. (2005). *Diffuse Flux of Greenhouse Gases — Methane and Carbon Dioxide — at the Sediment-Water Interface of Some Lakes and Reservoirs of the World*. https://doi.org/https://doi.org/10.1007/978-3-540-26643-3_6

Efstratiadis, A., & Koutsoyiannis, D. (2010). One decade of multi-objective calibration approaches in hydrological modelling: A review. *Hydrological Sciences Journal*, *55*(1), 58–78. https://doi.org/10.1080/02626660903526292

Godsey, S. E., & Kirchner, J. W. (2014). Dynamic, discontinuous stream networks: Hydrologically driven variations in active drainage density, flowing channels and stream order. *Hydrological Processes*, *28*(23), 5791–5803. https://doi.org/10.1002/hyp.10310

Grant, S. B., Azizian, M., Cook, P., Boano, F., & Rippy, M. A. (2018). Factoring stream turbulence into global assessments of nitrogen pollution. *Science*, *359*(6381), 1266–1269. https://doi.org/10.1126/science.aap8074

Grant, S. B., Gomez-Velez, J. D., & Ghisalberti, M. (2018). Modeling the Effects of Turbulence on Hyporheic Exchange and Local-to-Global Nutrient Processing in Streams. *Water Resources Research*, *54*(9), 5883–5889. https://doi.org/10.1029/2018WR023078

Hartmann, J., Lauerwald, R., & Moosdorf, N. (2014). A Brief Overview of the GLObal RIver Chemistry Database, GLORICH. *Procedia Earth and Planetary Science*, *10*, 23–27. https://doi.org/10.1016/j.proeps.2014.08.005

Hoellein, T. J., Bruesewitz, D. A., & Richardson, D. C. (2013). Revisiting Odum (1956): A synthesis of aquatic ecosystem metabolism. *Limnology and Oceanography*, *58*(6), 2089–2100. https://doi.org/10.4319/lo.2013.58.6.2089

Horgby, Å., Segatto, P. L., Bertuzzo, E., Lauerwald, R., Lehner, B., Ulseth, A. J., Vennemann, T. W., & Battin, T. J. (2019). Unexpected large evasion fluxes of carbon dioxide from turbulent streams draining the world's mountains. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-12905-z

Kessler, T. J., & Harvey, C. F. (2001). The global flux of carbon dioxide into groundwater. *Geophysical Research Letters*, *28*(2), 279–282. https://doi.org/10.1029/2000GL011505

Kortelainen, P., Rantakari, M., Huttunen, J. T., Mattsson, T., Alm, J., Juutinen, S., et al. (2006). Sediment respiration and lake trophic state are important predictors of large CO2 evasion from small boreal lakes. *Global Change Biology*, *12*(8), 1554–1567. https://doi.org/10.1111/j.1365-2486.2006.01167.x

Lauerwald, R., Allen, G. H., Deemer, B. R., Liu, S., Maavara, T., Raymond, P., Alcott, L., Bastviken, D., Hastie, A., Holgerson, M. A., Johnson, M. S., Lehner, B., Lin, P., Marzadri, A., Ran, L., Tian, H., Yang, X., Yao, Y., & Regnier, P. (n.d.). Inland water greenhouse gas budgets for RECCAP2: 2. Regionalization and homogenization of estimates. *Global Biogeochemical Cycles*, *n/a*(n/a), e2022GB007658. https://doi.org/10.1029/2022GB007658

Lauerwald, R., Laruelle, G. G., Hartmann, J., Ciais, P., & Regnier, P. A. G. (2015). Spatial patterns in CO2 evasion from the global river network. *Global Biogeochemical Cycles*, *29*(5), 534–554. https://doi.org/10.1002/2014GB004941

Lehner, B., & Döll, P. (2004). Development and validation of a global database of lakes, reservoirs and wetlands. *Journal of Hydrology*, *296*(1), 1–22. https://doi.org/10.1016/j.jhydrol.2004.03.028

Liu, S., Butman, D. E., & Raymond, P. A. (2020). Evaluating CO2 calculation error from organic alkalinity and pH measurement error in low ionic strength freshwaters. *Limnology and Oceanography: Methods*, *18*(10), 606–622. https://doi.org/10.1002/lom3.10388

Liu, S., Kuhn, C., Amatulli, G., Aho, K., Butman, D. E., Allen, G. H., Lin, P., Pan, M., Yamazaki, D., Brinkerhoff, C., Gleason, C., Xia, X., & 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. https://doi.org/10.1073/pnas.2106322119

Lorke, A., & Peeters, F. (2006). Toward a Unified Scaling Relation for Interfacial Fluxes. *Journal of Physical Oceanography*, *36*(5), 955–961. https://doi.org/10.1175/JPO2903.1

Mayorga, E., Seitzinger, S. P., Harrison, J. A., Dumont, E., Beusen, A. H. W., Bouwman, A. F., Fekete, B. M., Kroeze, C., & Van Drecht, G. (2010). Global Nutrient Export from WaterSheds 2 (NEWS 2): Model development and implementation. *Environmental Modelling & Software*, *25*(7), 837–853. https://doi.org/10.1016/j.envsoft.2010.01.007

Moore, R. B., McKay, L. D., Rea, A. H., Bondelid, T. R., Price, C. V., Dewald, T. G., & Johnston, C. M. (2019). User's guide for the national hydrography dataset plus (NHDPlus) high resolution. In *Open-File Report* (2019–1096). U.S. Geological Survey. https://doi.org/10.3133/ofr20191096

Plummer, L. N., & Busenberg, E. (1982). The solubilities of calcite, aragonite and vaterite in CO2H2O solutions between 0 and 90°C, and an evaluation of the aqueous model for the system CaCO3-CO2-H2O. *Geochimica et Cosmochimica Acta*, *46*(6), 1011–1040. https://doi.org/10.1016/0016-7037(82)90056-4

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*(7476), Article 7476. 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*(1), 41–53. https://doi.org/10.1215/21573689-1597669

Read, J. S., Hamilton, D. P., Desai, A. R., Rose, K. C., MacIntyre, S., Lenters, J. D., Smyth, R. L., Hanson, P. C., Cole, J. J., Staehr, P. A., Rusak, J. A., Pierson, D. C., Brookes, J. D., Laas, A., & Wu, C. H. (2012). Lake-size dependency of wind shear and convection as controls on gas exchange. *Geophysical Research Letters*, *39*(9). https://doi.org/10.1029/2012GL051886

Reisinger, A. J., Tank, J. L., Rosi-Marshall, E. J., Hall, R. O., & Baker, M. A. (2015). The varying role of water column nutrient uptake along river continua in contrasting landscapes. *Biogeochemistry*, *125*(1), 115–131. https://doi.org/10.1007/s10533-015-0118-z

Rocher-Ros, G., Sponseller, R. A., Lidberg, W., Mörth, C.-M., & Giesler, R. (2019). Landscape process domains drive patterns of CO2 evasion from river networks. *Limnology and Oceanography Letters*, *4*(4), 87–95. https://doi.org/10.1002/lol2.10108

Saccardi, B., & Winnick, M. (2021). Improving Predictions of Stream CO2 Concentrations and Fluxes Using a Stream Network Model: A Case Study in the East River Watershed, CO, USA. *Global Biogeochemical Cycles*, *35*(12), e2021GB006972. https://doi.org/10.1029/2021GB006972

Ulseth, A. J., Hall, R. O., Boix Canadell, M., Madinger, H. L., Niayifar, A., & Battin, T. J. (2019). Distinct air–water gas exchange regimes in low- and high-energy streams. *Nature Geoscience*, *12*(4), Article 4. 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*(C5), 7373–7382. https://doi.org/10.1029/92JC00188

Ward, N. D., Keil, R. G., Medeiros, P. M., Brito, D. C., Cunha, A. C., Dittmar, T., et al. (2013). Degradation of terrestrially derived macromolecules in the Amazon River. *Nature Geoscience*, *6*(7), 530–533. https://doi.org/10.1038/ngeo1817

Figure Captions

Figure 1. Process-based transport model emulates the distribution of in situ data: (**A**) pCO_2 versus k_{600} for the statistical upscaling model (blue lines) and our process based transport model (orange lines), both compared against GLORICH data with data source locations mapped in (**B**). To aid in visualization, we plot these models and data as the isolines for the bivariate kernel density space, showing 5 bands of equal relative likelihood that a pCO_2-k_{600} pair falls along that isoline. This probability increases with linewidth, i.e. the thicker isolines have more data. Note the outermost region extends beyond the axis limits. For both models, we randomly sampled 1,000 reaches from each of the 206 basins. All three use the same model for k_{600} (see Methods). Grey shading is the hypothetical FCO_2 flux (at 20 degrees celsius) for all possible pairs of pCO_2 and k_{600}, i.e. FCO_2 increases towards the upper-right corner of **A.** (**C**) Comparison of total CO_2 emissions from CONUS inland waters, estimated via both models. Colors match subplot **A**. Error bars refer to model uncertainty (Eq 6,7) alone; parameter uncertainty is identical across both models and so not included here (see Main text and Methods).

Figure 2. River/lake/reservoir CO_2 emissions for United States inland waters. Area-normalized FCO_2 at mean annual flow for over 22M inland waters. Lakes/reservoirs (and their associated CO_2 fluxes) are also plotted in the two smallest-scale inset maps to highlight hydrological connectivity. Reach width in the inset maps is scaled to discharge- thicker lines have more flow. Note that at the continental scale, headwater streams with the highest overall CO_2 fluxes are visually overrepresented based on the number of individual reaches.

Figure 3. Sources of inland water CO_2 emissions. **A-C:** Percent of CO_2 lost from a basin that is attributed to stream corridor subsurface respiration (**A**), upland groundwater CO_2 (**B)**, and net water-column respiration **(C)**. **D:** Percent of CO_2 emissions attributed to the same mechanisms as **A-C** by stream order; boxplots are composed of the median percent value per basin per stream order. See Methods for these calculations at the basin-scale (**A-C**) and the reach-scale (**D**). Note we lump high stream orders (seven and above) due to the small number of basins with this many stream orders and to represent network main stems as a single boxplot. SFig. 9 separates **D** by eastern and western CONUS basins.

Figure 4. Lake and reservoir influence on inland water CO_2 emissions. Percent of CO_2 emissions via lakes/reservoirs and estimated using the process-based transport model.