Title: Sediment-associated processes drive spatial variation in ecosystem respiration in the Yakima River basin

Authors: Matthew Kaufman[1,2+], Vanessa A. Garayburu-Caruso[1+], Brieanne Forbes[1], Xinming Lin[1], Robert O. Hall Jr.[3], Stephanie Fulton[1], Lupita Renteria[1], Yilin Fang[1], Kyongho Son[1], James C. Stegen[1,4*]

[1]Pacific Northwest National Laboratory, Richland, WA 99354
[2]Worcester State University, Worcester, MA 01602
[3]Flathead Lake Biological Station, University of Montana, Polson, MT 59860
[4]Washington State University, Pullman, WA 99163
[+] These authors contributed equally.
*Correspondence: james.stegen@pnnl.gov

Abstract: Areas where groundwater and surface water mix (i.e., hyporheic zones, HZ) contribute substantially to stream ecosystem respiration (ERtot). We rely on reactive transport models to understand HZ respiration at large scales; however, model outputs have not been evaluated with field estimates of ERtot. Here we evaluate the degree to which spatial variation in model-predicted HZ respiration can explain spatial variation in field-estimated ERtot across 32 sites in the Yakima River basin (YRB). We find that predicted HZ respiration did not explain spatial variation in ERtot. We hypothesize that ERtot is influenced by processes that integrate contributions from sediments, such as benthic algae, submerged macrophytes, and shallow HZ. Our results indicate that sediment-associated processes hydrologically connected to the active channel are primary drivers of spatial variation in ERtot in the YRB. We encourage conceptual and physical models of stream ERtot to integrate shallow hyporheic exchange with sediment-associated primary production.

Introduction

River corridors contribute substantial amounts of carbon dioxide (CO_2) to the atmosphere[1]. Specifically, areas where groundwater and surface water mix (i.e., hyporheic zones, HZ) can contribute 4-90% of total ecosystem respiration[2–7]. Field experiments have estimated HZ respiration via extrapolations of sediment microcosms[8,9], as the difference between ecosystem and benthic respiration[2], or via end-member mixing analysis using in-situ solute concentration[3]. However, quantifying HZ respiration in the field across scales is difficult due to high spatial and temporal heterogeneity in hydrologic conditions and stream geomorphology. Currently, we rely on models like the Networks with Exchange and Subsurface Storage (NEXSS)[10] and the River Corridor Model (RCM)[11–13] to predict the spatial variation of HZ respiration at watershed to basin scales. These and other watershed/basin scale models rely on physical stream characteristics (e.g., sediment texture, geomorphology, discharge) to estimate hydrologic exchange fluxes that, when paired to reaction network models, provide an estimate of HZ respiration rates[10–13]. While each of these models provide insights into spatial patterns of HZ respiration, they have not been evaluated via field estimates of stream ecosystem respiration. Therefore, the relative contributions of HZ respiration to total stream ecosystem respiration at a basin scale are still largely unexplored.

Ecosystem metabolism and carbon (C) cycling in rivers have been widely studied via two key processes: C fixation (gross primary production, GPP) and organic C mineralization (ecosystem respiration, ER)[14]. These estimates are derived from diel dissolved oxygen (DO) measurements that include the collective aerobic organisms (i.e., autotrophs and heterotrophs) and habitats

(i.e., benthic, planktonic, and hyporheic zones), which together contribute to C cycling dynamics in streams and rivers[15]. Due to the advances in sensing technologies[16] and modeling capabilities[17], river metabolism has been studied worldwide to evaluate how environmental factors like climate, elevation, and latitude impact aquatic ecosystems[1,18,19]. Studying river metabolism allows for deeper understanding about food webs, energy fluxes, C and nutrient cycling, as well as predictions of ecosystem responses to environmental change[20–23].

River metabolism is influenced by several biotic and abiotic factors like light, temperature, nutrients, and hydrogeomorphology[24,25]. A recent study across 222 rivers in the United States shows that light availability and stream flow conditions govern annual variation in both GPP and ER[26]. Additionally, a global assessment of river metabolism shows that GPP and ER increased with latitude and suggests that spatial variability of site characteristics like land use and vegetation are more influential than latitudinal temperature effects on river metabolism[27]. These studies identified metabolism drivers linked to chemical and biological processes across scales, yet they do not explicitly consider the role of hydrologic exchange, a fundamental process within the reactive transport models. Thus, we lack conceptual understanding of the degree to which spatial variation in model-predicted HZ respiration explains spatial variation in total stream ecosystem respiration (ERtot).

In this study, we first test the hypothesis that spatial variation in HZ respiration is a strong driver of spatial variation in ERtot at the basin scale. To compare these processes, we use RCM model-predictions to represent spatial variation in HZ respiration and relate these predictions to field-based estimates of ERtot across 32 sites in the Yakima River basin (YRB). Upon finding little to no relationship between predicted HZ respiration and ERtot, we explored other factors that may drive ERtot. In turn, we examined the relative influence of water column processes and found minimal contributions from water column respiration (ERwc) to ERtot. This finding allowed us to test the post-hoc hypothesis that spatial variation in ERtot is primarily modulated by processes associated with benthic/streambed sediments and the HZ, but that may not all be represented in the RCM. We refer to all sediment-associated respiration that influences field-based measures of ERtot as ERsed. Our data support the hypothesis that ERsed is a primary driver of ERtot, and we find that ERsed is primarily linked to sediment-associated GPP and aspects of physical settings associated with stream velocity and topographic gradient. Our study highlights the importance of sediment-associated processes in modulating basin scale ERtot and emphasizes the need for process-based forward-simulation capabilities that incorporate sediment-associated autotrophic and heterotrophic metabolism in benthic zones and HZs, in addition to hydrologic exchange.

Results and Discussion

Our basin-scale study provided three primary outcomes: (*i*) Spatial variation in model-predicted HZ respiration (as viewed through the river corridor model, RCM) was most likely not a dominant driver of spatial variation in field-based estimates of stream ecosystem respiration (ERtot) at the basin scale. (*ii*) Little respiration was contributed by processes within the water column, and spatial variation in water column respiration does not explain spatial variation in ERtot. (*iii*) Spatial variation in ERtot was primarily modulated by sediment-associated gross primary production (GPP) and the physical environment. Our observations do not reject the RCM model, and instead indicate there are processes driving ERtot at a basin scale beyond those present in the RCM model. This finding implies that for physical models to predict ERtot, they need to account for physical processes associated with surface-subsurface hydrologic

exchange flows and sediment-associated respiration (ERsed) fueled directly and indirectly by sediment-attached photoautotrophs (i.e., algal biofilms and submerged macrophytes).

Spatial variation in RCM-predicted hyporheic zone (HZ) respiration does not explain spatial variation in field-based ERtot estimates across the Yakima River basin (YRB).

Estimates of time averaged ERtot across sites (n = 44) ranged from -19.6 to 7.7 O_2 $m^{-2}{\cdot}d^{-1}$ (Fig. 1a). We were not able to estimate metabolism (i.e., GPP and ERtot) in 4 sites that failed to meet the assumptions of one station open channel oxygen budgets due to slower flows and lack of turbulence at the deployment sites[17]. In addition, out of the 44 sites where we were able to estimate metabolism, 12 sites displayed positive estimates of ERtot. These had DO that rarely went under 100% saturation, which indicates high gas exchange rates potentially due to turbulence and pressurized bubbles relative to the atmosphere[28]. Oversaturation hinders the estimation of ERtot using open channel DO time series because we have no simple way of knowing what the concentration of oxygen would be in the river in the absence of any biological activity[29]. We removed positive ERtot values from our analysis because they have no biological meaning, and the magnitude of these ER values reflects misspecification due to bubble mediated gas exchange and not a biological process. After removing these values, average ERtot (n = 32) ranged from -19.61 to -0.27 O_2 $m^{-2}{\cdot}d^{-1}$.

Field-based estimates of ERtot diverged from RCM-predicted HZ respiration. The fastest ERtot rates occurred in the lower region of the YRB (Fig. 1a), in locations with shallow topographic gradients and high order streams (i.e., topographic slopes of 0.0004 - 0.002 and stream orders of 5-7). The fastest HZ respiration rates predicted by the RCM were, in contrast, in upper parts of the YRB (Fig. 1b). RCM-predicted HZ respiration values (n = 46) ranged from -800 to -0.024 O2 $m^{-}2{\cdot}d^{-}1$. The RCM-predicted rates are much faster than estimates of ERtot, which may suggest a need for better parameterization of the RCM and the need for additional process representation (e.g., water-atmosphere gas exchange). Resolving uncertainties associated with the RCM is beyond the scope of the current study, yet we focus on RCM-predicted spatial patterns as the best current representation for how HZ respiration rates are likely to vary across the YRB. This approach allows for evaluating a qualitative relationship between the spatial patterns of predicted HZ respiration and ERtot.

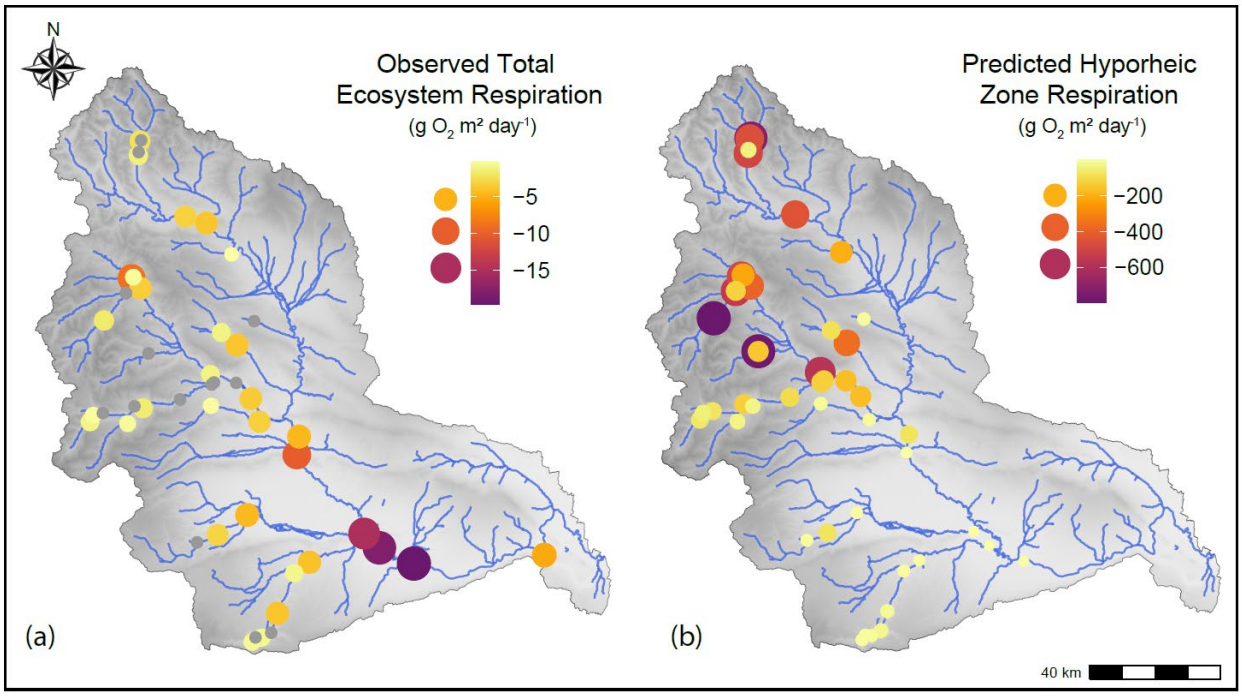

Figure 1. Map displaying Yakima River basin sites (n = 48) investigated in this study in conjunction with respiration rates. (a) Field-based estimates of total stream ecosystem respiration rates (ERtot) and (b) River Corridor Model (RCM) predicted hyporheic zone (HZ) respiration rates[13]. Color and symbol size are used to indicate magnitude of respiration rate. Grey dots in panel a indicate sites where ERtot could not be modeled or where the outputs were positive values.

Spatial variation in RCM-predicted HZ respiration had a weak negative relationship with spatial variation in ERtot. Before doing associated analyses and given the quantitative difference between RCM predictions and ERtot, we standardized rates within each dataset as z-scores, whereby each dataset has a mean of zero and standard deviation of one. This does not change quantitative relationships within each dataset or the shape of the data distribution, it only puts the two datasets on the same quantitative range. Following this normalization, we regressed RCM-predicted HZ respiration z-scores against z-scores of field-based estimates of ERtot (Fig. 2). No relationship is apparent in the data (Fig. 2a), which also clearly violate normality assumptions required for conducting parametric regression. A weak negative relationship was, however, observed via Spearman rank-order correlation (Fig. 2b; $p = 0.04$). The negative direction and weakness of the relationship indicate, however, that HZ respiration—as represented in the RCM—is not a primary driver in variation of ERtot in the YRB. The lack of a strong positive relationship could be for several reasons, such as ERtot not including respiration occurring across the full HZ spatial domain that is represented in the RCM, and the RCM not including key ERtot drivers.

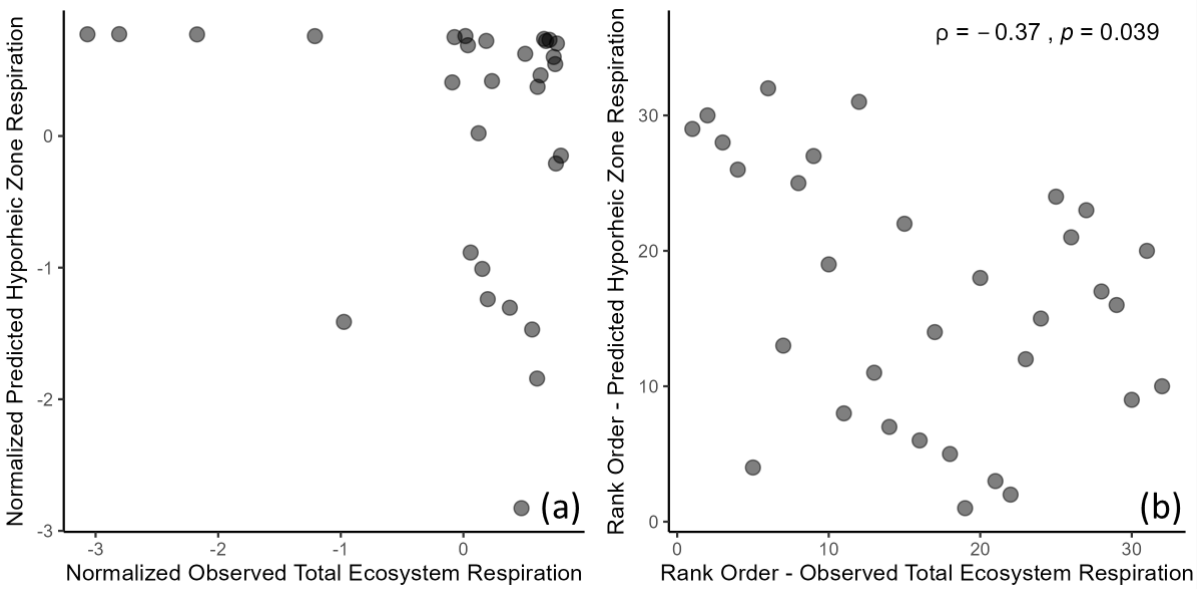

Figure 2. Normalized model-predicted hyporheic zone (HZ) respiration rates do not increase with observed field-based estimates of total stream ecosystem respiration rates (ERtot). (a) Z-score normalized data suggest no relationship. Data distributions are skewed (i.e., not normal) and thus violate assumptions of parametric regression, so no statistics are given. (b) Spearman rank-order correlation reveals a weak, but significant, negative relationship. The Spearman correlation statistics are provided on the panel.

Field estimates of ERtot are based on DO dynamics that reflect the reach-scale integrated outcome of processes occurring across all system components that are hydrologically connected to the active channel[15]. This estimate includes respiration associated with subsurface sediments along relatively short HZ flow paths. These HZ flow paths must be fully contained within the upstream reach that influences observed DO dynamics at a given monitoring location. If the HZ flow paths are longer, they should not influence observed DO and thus do not influence estimated ERtot. The length of the reach that a given DO sensor integrates varies across streams and has been estimated to be roughly three times the turnover length of O_2[15]. It is plausible that the RCM model captures respiration along flow paths that are longer than what the DO sensors integrate across. If so, this long flow path ER could lead to the lack of positive correlation between RCM-predicted HZ respiration and field-based estimates of ERtot. We suspect, however, that this might not be the primary reason for the lack of a relationship because both RCM-predicted HZ respiration and field-based ERtot account for short HZ flow paths. These shorter flow paths are likely to be within the most biogeochemically active sediments, which should in turn be the primary drivers of total HZ respiration[30]. We postulate that the lack of a relationship (Fig. 2) and divergence in spatial patterns (Fig. 1) is related to differences in which biogeochemical processes are captured in the two methods.

The RCM model represents hydrologic exchange through the HZ and sediment-associated respiration that is fueled by reactants delivered by that exchange[13]. In the model, reactants delivered to HZ sediments are not influenced by primary producers (e.g., photoautotrophic algal biofilms) and the respiration contributed by primary producers is also not accounted for. Respiration occurring by heterotrophs and photoautotrophic primary producers in the water column is also not part of the RCM. All these processes can influence observed DO dynamics[31–33] and are thus implicitly accounted for in field estimates of ERtot. This highlights that field estimated ERtot is influenced by several processes that are not represented in the RCM. This is not a critique or rejection of the RCM, rather, it is a recognition that the divergence in spatial

patterns between RCM-predicted HZ respiration and field-based estimates of ERtot are likely due, at least in part, to a mismatch in the processes captured by each method.

In addition to incomplete overlap in captured processes, RCM-predicted HZ respiration and ERtot may not be correlated if processes in the RCM do not strongly influence ERtot, if data inputs to the RCM misrepresent the system, and technical issues associated with estimating ERtot. Our estimates of ERtot largely overlap with literature values[34] (Fig. 3a) and field sites with technical issues were removed from the analysis. It is, therefore, likely that the ERtot estimates represent actual near surface benthic respiration and are not a cause for the lack of correlation. If true, the lack of correspondence in spatial patterns and the lack of direct correlation suggest that it may be useful for the RCM to expand the breadth of represented processes. Available data allow us to guide which processes could be most important to include in the RCM and, more generally, to investigate factors associated with spatial patterns in ERtot. To these ends, we next examine the relative influence of water column processes over ERtot, which provides conceptual understanding of system function and evaluates whether water column processes should be a priority for an expanded version of the RCM.

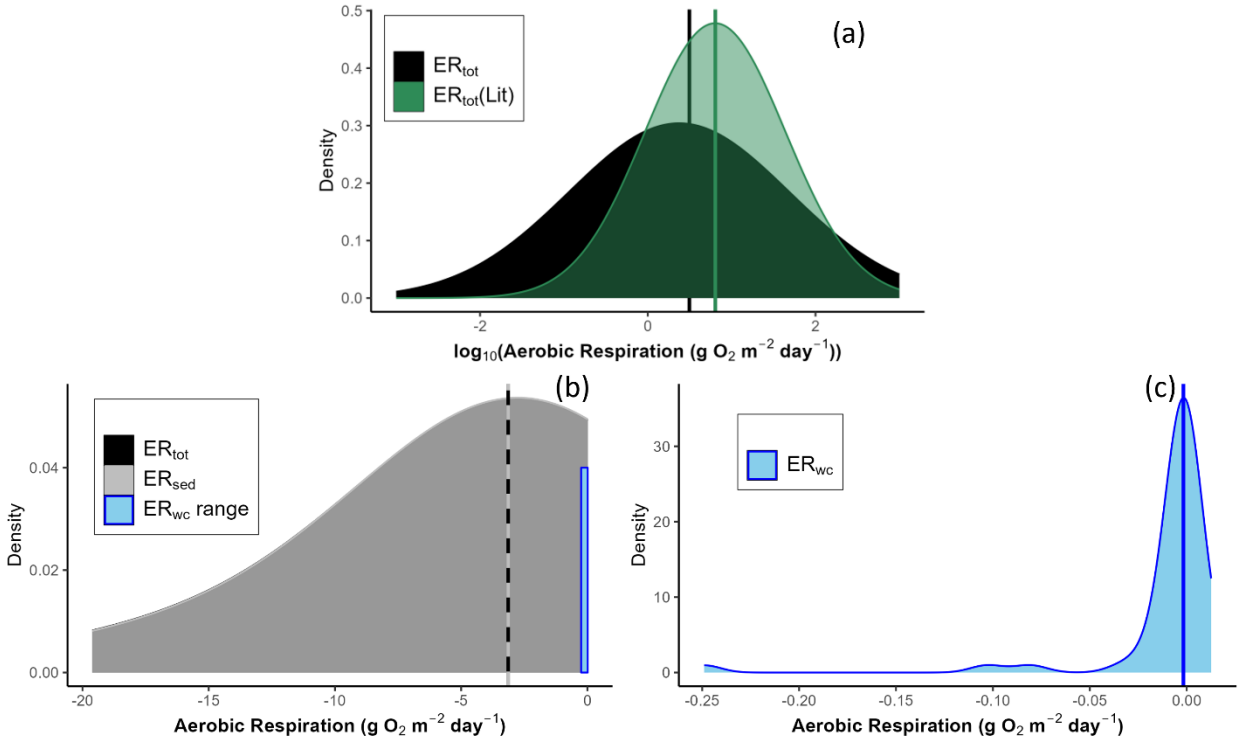

Figure 3. Density plots comparing respiration rates for different compartments within the Yakima River basin (YRB) as well as literature values. (a) Density plot for YRB total ecosystem respiration from this study (ERtot, black, n = 32) and from rivers and streams across the CONUS (n = 207) estimated by Appling et al. (2018)[34] (ERtot(Lit), green); (b) Density plot for YRB ERtot (black, n = 32) and ERsed (grey, n = 32), and the range of YRB water column respiration rates (ERwc, blue n = 48); (c) Density plot for ERwc (blue). The vertical lines are the medians for each distribution.

Respiration is dominated by processes directly or indirectly associated with sediments, not the water column.

Previous work in the YRB[35] indicated that water column respiration (ERwc) is a minor component of ERtot, but this inference was indirect as ERtot estimates were not available. We postulated that when evaluated more directly, ERwc will be a minor fraction of ERtot. If so, this would indicate that the lack of water column processes in the RCM is not a primary cause for divergence between RCM-predicted HZ respiration and ERtot. Our ERwc data, estimated via *in-situ* dark bottle incubations, were consistent with this premise. Across all sites, ERwc (n = 48) ranged from -0.249 to +0.013 g O_2 $m^{-2}{\cdot}d^{-1}$ (Fig. 3c). Contrary to the data cleaning decisions made for ERtot, we chose to interpret positive ERwc because they reflect instrument noise paired with extremely slow respiration. In this situation the rate at which O_2 was consumed was not fast enough to overcome instrument error and resulted in a small positive ER measurement. Across all 32 field locations with ERtot estimates, we found very small fractional contributions of ERwc to ERtot (2% maximum, Table 1) as well as no spatial correspondence between ERwc and ERtot (Figure S1) and thus ERwc does not explain spatial variation in ERtot. This finding is consistent with the indirect inferences from Fulton et al. (2024)[35] and strongly suggests that lack of water column respiration in the RCM is not a cause for divergence.

Table 1. Summary statistics (mean, median and range) for ERtot, ERwc, and ERsed in g O2 m^{-2} day^{-1}

	Mean	**Median**	**Range**
ER_{tot}	-4.219	-3.135	(-19.61, -0.27)
ER_{wc}	-0.017	-0.0018	(-0.249, 0.005)
ER_{sed}	-4.102	-3.135	(-19.36, -0.27)
$ER_{sed}(\%)$	99.7%	100%	(97.92%, 100%)

As discussed in Fulton et al. (2024)[35], ERwc in the YRB is slow relative to most stream and river systems[4,36,37]. This low ERwc could be due to lower amounts of total suspended solids (TSS), lower nutrient concentrations, shallower water, and shorter surface water residence time in the YRB versus to other rivers. More specifically, TSS ranged from 0.12–202.4 mg L^{-1}, total dissolved nitrogen (TN) ranged from 0.05–1.54 mg L^{-1}, dissolved organic carbon (DOC) ranged from 0.34–2.62 mg L^{-1}, and the water column depth ranged from 0.12–2.02 m. Low TSS could result in fewer particles for microbes to attach to, in addition, low nutrients constrain organismal metabolism and water column depth influences phytoplankton biomass[38–41]. Moreover potentially shorter surface water residence times in the basin would prevent plankton growth in the water column. We infer the combination of these factors lead to slow ERwc and a small fractional contribution to ERtot.

We integrated our ERwc results with previous studies that highlighted the importance of riverine hyporheic and benthic sediments in ecosystem respiration[3,4,42–44] to generate a post-hoc hypothesis that would help explain the lack of relationship between RCM-predicted HZ respiration and ERtot. We hypothesize that the spatial variation in ERtot rates is primarily modulated by metabolic processes integrated across benthic/streambed sediments and HZs that are hydrologically connected to the active channel. We define the component of ERtot that comes from these integrated sediment-associated processes as ERsed. ERsed was calculated as the difference between ERtot and ERwc at each site with negative ERtot (n = 32). ERsed values ranged from -19.4 to -0.3 g O_2 $m^{-2}{\cdot}d^{-1}$ indicating that ERsed contributed more than 97% of ERtot across all sampled stream reaches (Table 1, Fig. 3b). We emphasize that our measurement of ERsed integrates multiple ecosystem components including the HZ, benthic

biofilms, and submerged macrophytes. In contrast, ERwc is exclusively from organisms suspended in the water column.

Regardless of the governing mechanisms, during this study ERwc appeared to have little influence over ERtot in the YRB. Respiration in the water column of streams throughout the basin are therefore unlikely to contribute meaningfully to net carbon dioxide (CO_2) fluxes. However, during high-flow or high-TSS periods of the year, this result may be different as ERwc is known to increase with TSS and co-varying aspects such as nutrient concentrations[37,45,46]. Negligible influence of ERwc in the YRB also contrasts with recent observations in the mainstem Columbia River, which is immediately adjacent to the YRB. In that study, Roley et al. (2023)[47] found that most ERtot was likely from processes occurring in the water column of the Columbia River. This fact may be due to the greater depth and width of the Columbia River, relative to streams across the YRB, and from phytoplankton sourced from an upstream reservoir[47]. The contrast between the YRB and the Columbia River indicate that respiration contributions from ERwc can vary widely across streams and rivers. While ERsed contributed most ERtot in our study, we suggest that ERwc remains important to account for at certain places and times where it will constitute much of ERtot.

Spatial variability in ERsed is best explained by primary production and physical setting.

To investigate spatial drivers of ERsed, we identified variables known from previous work to influence ERtot (e.g., gross primary production; GPP) and variables that influence HZ respiration (e.g., median grain size) (see Table 4 for all variables included). We did not include RCM-predicted HZ respiration as an explanatory variable because it was strongly correlated with median grain size (i.e., d50), and because we found no relationship between RCM-predicted HZ respiration and ERtot (Fig. 2). Our analysis showed that GPP, stream slope, stream velocity, and TN explained the most variation in ERsed (Table 2). The directionality of these relationships further shows that the fastest ERsed rates (and thus ERtot rates) had the highest GPP rates (Table 2). This finding is generally consistent with the observation from Bernhardt et al. (2022)[26] who found that ERtot was most strongly linked to light availability. Additionally, faster ERsed rates were observed in reaches with shallower topographic gradients, lower velocity, and higher TN concentrations. Flow velocity plays a key role in biofilm scouring, architecture, and functional dynamics[48,49]. Moreover, stream gradient influences flow velocities along with sediment transport[50] and nitrogen can impact biofilm community structure and C uptake[51]. Together these variables can have a substantial effect on benthic metabolism and in turn ERsed.

Table 2. Regression coefficients for LASSO model explaining variation in ERsed. All input variables (Table 4) from 32 sites were log transformed and standardized.

Input Variable	Regression coefficient (β)
GPP	0.426
Slope	-0.222
Velocity	-0.186
TN	0.169
Mean Depth	0.102
PctShrb	-0.032

In addition, ERsed covaried non-linearly with GPP (Figure 4). While Akaike Information Criterion (AIC) analysis provided weak support for a segmented regression (i.e., AIC score for the segmented regression model was 144 and the AIC score for the linear regression was 147), we selected the model with lower AIC. In turn, the first regression of the segmented regression covered the lower range of GPP and displayed a less-steep relationship between ERsed and GPP. The second regression spanned the upper range of GPP, in this case the relationship was steeper whereby a unit increase in GPP led to a larger increase in ERsed[19,48]. This non-linear relationship was also linked to a shift from heterotrophy (in some reaches) at low GPP rates to net autotrophy across higher GPP. This can be seen as values falling below and then above the one-to-one line in Figure 4, moving from low to high GPP. Generally, respiration in streams is dominated by heterotrophy, under those conditions respiration is fostered by allochthonous C inputs and rivers become a net source of CO_2 to the atmosphere[15,21,26,52,53]. On the other hand, autotrophic conditions indicate that the river is fixing more C than it is respiring and thus it is storing or exporting some of this C to downstream systems[15]. Net autotrophy while less common than net heterotrophy has been observed during baseflow conditions in agricultural systems[54], in a montane stream with some seasonal and year to year variability[55] and in a tussock/pasture regularly grazed stream[56].

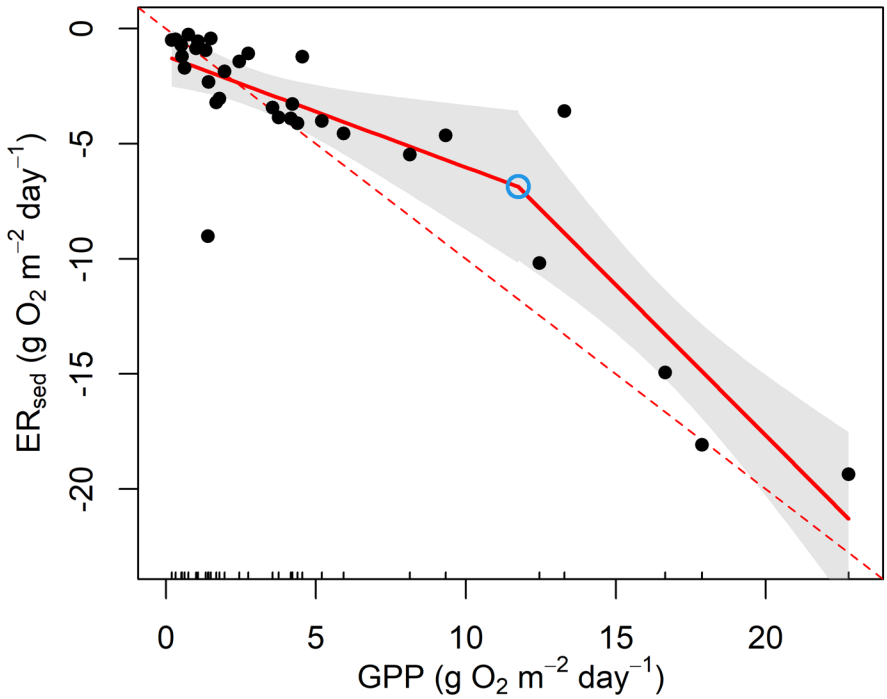

Figure 4. Scatterplot with segmented regression lines (red) depicting the relationship between GPP (x-axis) and the observed sediment respiration (ERsed , y-axis). Regression break point (blue circle) is located at 11.7 ± 2.7 g O_2 $m^{-2}{\cdot}d^{-1}$. The red dashed line is the one-to-one line with slope of -1, shaded regions represent the pointwise confidence intervals of the linear model. The first segment shows a mild negative relationship, while the second segment is much steeper, suggesting a stronger negative relationship.

Given the much larger LASSO coefficient for GPP, relative to other explanatory variables (Table 2), and the steep increases in ERsed with increasing GPP (Fig. 4), we suggest there is likely a causal link between sediment-associated primary production and ERsed. Primary producers in this case are likely to include both algal biofilms attached to benthic/streambed sediments and submerged macrophytes rooted in sediments. We infer that both types of primary producers likely provide important direct and indirect contributions to ERsed. Direct contributions include phototroph respiration, and indirect contributions include providing resources (e.g., organic exudates) to sediment-associated heterotrophs[14,57]. As noted above, this link between ERsed and GPP in the YRB is conceptually consistent with Bernhardt et al. (2022)[26]'s result that light, via GPP, is a primary explanatory variable for ERtot across the continuous United States. Further consistent with Bernhardt et al. (2022)[26] is our result that ERsed rates (and thus ERtot rates) are likely influenced by variables beyond GPP and that may be linked to hydrologic exchange. In the YRB, low-gradient reaches with slow moving water had faster ERsed after controlling for GPP (Table 2). This relationship may additionally reflect a shift along the transporter-to-transformer continuum whereby residence time is long enough to enable significant reactions to occur, thereby enhancing ERsed. Together, these results point to a key role for the ratio of residence time to reaction time as classically represented in the Damköhler number. That is, we infer that ERsed hot spots[58] or control points[59] are more likely occur when the Damköhler number is high because of low gradient and slow flow causing long residence time and high GPP causing short reaction times. These factors are likely predictable from physical features of stream networks, thereby providing opportunities to develop mechanistic or data-driven models of ERsed hot spots/control points.

Discussion

This study was inspired by recent progress in modeling respiratory processes within hyporheic zones (HZ) and the common perspective that HZs are dominant control points with rivers[44,60,61]. While our study does not directly evaluate the biogeochemical contributions of HZs, we do provide evidence that the hydrologic processes that govern variation in predicted HZ respiration are not the same processes that govern ERtot and ERsed. We show that the spatial variation in basin-scale total ecosystem respiration (ERtot) is primarily modulated by processes that integrate mechanisms from benthic/streambeds and HZs that are hydrologically connected to the active channel. Further investigation is required to parse influences from benthic/streambed sediments and HZ sediments, which is extremely challenging to do *in situ*.

Identifying variables that explain spatial variability in ERsed is essential to better understand what may drive ERtot and why HZ respiration (as predicted by hydrologic exchange) wasn't related to ERtot. Figure 5 shows that sediment-associated respiration (ERsed) was fastest in lower parts of the Yakima River basin (YRB), while HZ respiration is predicted to be maximized in upper parts of the YRB. Spatial patterns in ERsed correspond to ERtot patterns given that >97% of ERtot is from ERsed. To help understand divergence in River Corridor Model (RCM) predicted HZ respiration and ERtot (and ERsed), we emphasize that sediment texture (e.g., median grain size) was the most influential variable driving spatial variation in RCM-predicted HZ respiration[13]. A preliminary evaluation across the YRB suggests weak correspondence between model-based sediment texture estimates used as inputs to the RCM and new field-based estimates of sediment texture based on machine learning analysis of *in situ* photos[62]. This weak correlation highlights potential errors and uncertainties associated with sediment texture data products used as inputs to the RCM, making unclear whether low hydrologic exchange and fine sediments mechanistically lead to fast ERsed rates. We speculate that finer sediments will present more surfaces for heterotrophic and autotrophic biofilms, but our data cannot be used to directly evaluate this possibility. We also cannot yet imagine a mechanism through which low hydrologic exchange alone would lead to fast ERsed rates. It may be that these variables are acting as proxies for other system components that are more directly and mechanistically linked to ERsed.

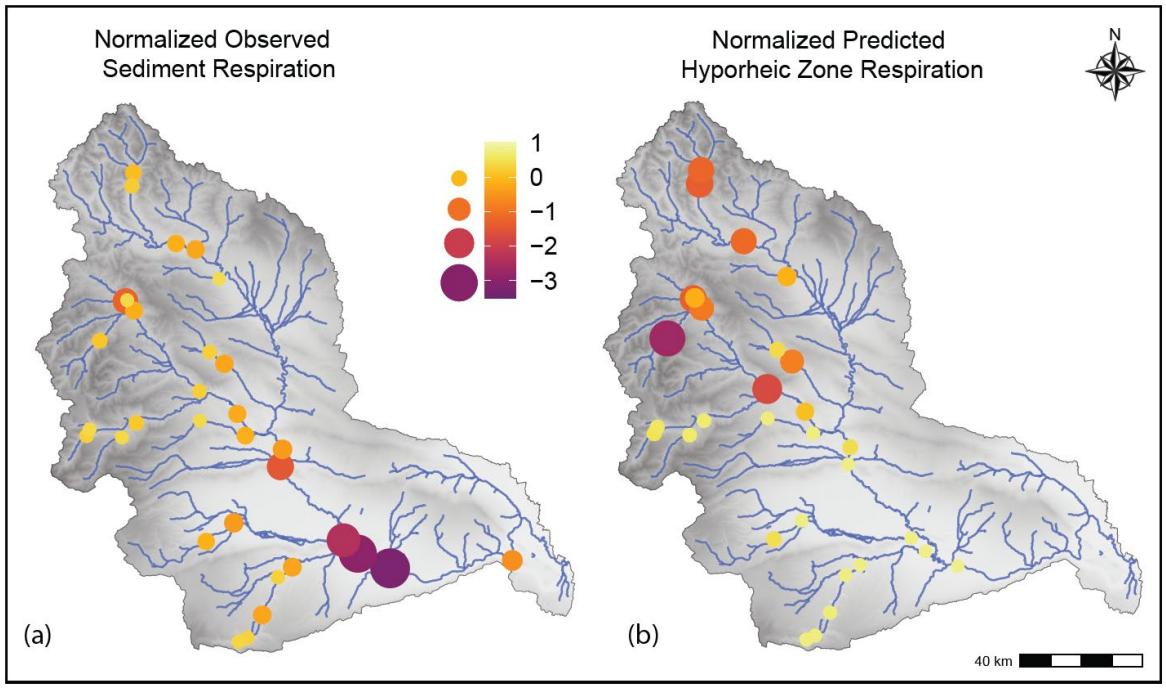

Figure 5. Map displaying normalized respiration rates for Yakima River basin sites (YRB, n = 32) investigated in this study. (a) Normalized sediment-associated respiration rates (ERsed) and (b) Normalized model-predicted hyporheic zone (HZ) respiration rates[13]. For both observed and predicted rates, color and symbol size are used to indicate respiration rate z-scores.

The RCM relies heavily on several spatially variable inputs, including median grain size (i.e., d50) and dissolved organic carbon (DOC) concentrations. These input layers, particularly d50, are generated from sparse field measurements and may not accurately reflect local conditions in the basin. If those input data are inaccurate, that will cause errors in predicted spatial patterns of hyporheic zone respiration. Additionally, the model does not consider the impact of sediment grain size on reaction rates, only on exchange rates. Full evaluation of all data inputs to physical models is well beyond the scope of this manuscript, though we encourage follow-on studies to be aware of the potential for such sources of error and to minimize them as much as possible.

Our results indicate a need to extend conceptual and physical models that aim to capture and predict active channel respiration to represent all processes that have major influences over those rates. We show that it is not enough to focus on HZ respiration as influenced primarily by hyporheic exchange processes. While we found little contribution of water column respiration (ERwc), recent work in the adjacent Columbia River found that water column processes appear to dominate both ERtot and gross primary production (GPP)[47]. Together with our study, these results indicate a need to explicitly represent primary producers and heterotrophs in both the water column and associated with sediments. Current models can be modified in these ways to generate new predictions for evaluation with further field deployments design specifically to test them. We refer to this approach as hypothesis-based model-experiment (ModEx) iteration, and this study represents a key first step. A ModEx approach goes beyond data-model integration whereby models are used to pose new hypotheses that are tested with new data, leading to modified hypotheses that guide follow-on model development, and so on. We encourage the application of this approach in other basins and across other flow conditions to evaluate the transferability of our outcomes. This is key to building predictive models that are useful across Earth's diverse rivers.

Methods

1. Site selection

We selected the Yakima River basin (YRB) as a test basin for this study. The basin drains an area of 15,941.4km^2 and spans forested mountainous regions in the west to arid valleys and plains in the east (Fig 6). The basin has a diversity of land covers and land uses dominated by dry rangeland, forest, and agriculture[63,64]. Annual precipitation ranges from 203-356 cm in the west to 25 cm in the east[65].

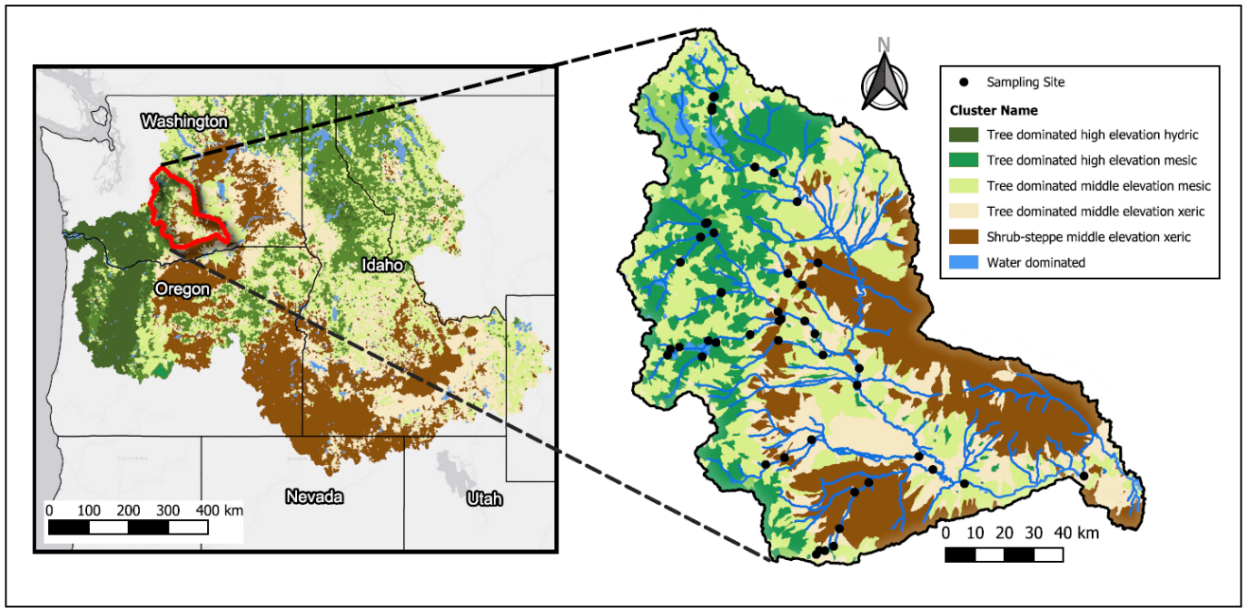

Figure 6. Yakima River basin sampling sites superimposed on a map of catchment cluster analysis results.

Field sites spanned a range of stream orders, land cover types, physical settings, and previously published model-predicted HZ respiration rates[13]. Since model-predictions were performed at a reach scale for the Columbia River basin, we conducted a clustering analysis following methods previously described by Fulton et al. (2024)[35] to locate regions in our test basin with representative characteristics from the Columbia River basin (Fig. 6). We then combined the clustering analysis results with logistical considerations regarding site access, safety, and model-predicted HZ respiration rates to select 48 sites that span a range of clusters, model-predicted respiration rates and stream orders across the basin (Fig. 6). Table 3 shows a summary of the cluster analysis results as well as number of sites in the YRB per cluster. Geospatial variables for each site were extracted primarily from the Environmental Protection Agency StreamCat Dataset and NHDPlusV2.1 using a custom R script[66]. Land use/land cover data for all the sites were extracted from StreamCat [67], while catchment area and slope were extracted via NHDPlusV2.1 with the R package 'nhdplusTools'[68]. The custom R script was also used to extract aridity data from the Global Aridity Index and Potential Evapotranspiration (ET0) Database: Version 3[69] with the terra::rast() and terra::extract() functions[70].

Table 3. Cluster analysis results across the Columbia River basin and Yakima River basin with similar biophysical and hydrologic characteristics and the number and percentage of sites in each.

Cluster	Cluster Name	CRB Drainage Area (%)	YRB Drainage Area (%)	No. YRB Sites Per Cluster	Percent YRB Sites Per Cluster (%)
1	Tree dominated high elevation mesic	22.7	27.4	12	25.0
2	Water dominated	3.4	2.2	0	0
3	Tree dominated high elevation hydric	7.3	2.4	2	4.2
4	Shrub-steppe middle elevation xeric	25.4	28.1	5	10.4
5	Tree dominated middle elevation mesic	17.4	17.3	15	31.3
6	Tree dominated middle elevation xeric	23.9	22.6	14	29.2

"CRB Drainage Area" is the percentage of the total drainage area of the Columbia River basin that was classified in each cluster. "YRB Drainage Area" is the percentage of the total drainage area of the Yakima River basin that was classified in each cluster. "YRB Sites Per Cluster" is the total number of field sites in the Yakima River basin (n = 47) located in each cluster. "Percent YRB Sites Per Cluster" is the percentage of the total number of sampling sites in the Yakima River basin located in each cluster.

2. Sensor deployment and data collection

During the week of July 25–28, 2022, dissolved oxygen (DO) sensors (miniDOT Logger; PME, Inc.; Vista, CA, USA) and level loggers (HOBO® Water Level Logger (13') U20L-04; Onset; Bourne, MA, USA) were deployed with sensor windows facing upstream on the thalweg of the riverbed at each site (*n* = 48) for 35 days during baseflow conditions. We were unable to deploy sensors and sample first order streams due to the lack or low flow during the summer. The sensors were secured to two bricks with zip ties using several different configurations depending on streamflow velocity, depth, and channel substrate (Figure 7b, d). DO sensors recorded DO and water temperature continuously on a 1-min interval. To minimize sensor biofouling in warmer, shallower rivers or rivers that received a lot of sunlight, 25 DO sensors were equipped with wipers (miniWIPER for miniDOT; PME; Vista, CA, USA) and anti-fouling copper plates (miniDOT Anti-fouling Copper Option; PME; Vista, CA, USA) (Fig 7d). Wipers were programmed to wipe the sensor face once every 12 hours. The remaining 23 DO sensors were equipped with full copper kits, which included a copper ring, silicone ring, and a copper mesh (Fig 7b). Non-vented level loggers recorded absolute pressure and water temperature continuously on a 15-min interval. Additional barometric pressure (BP, Rugged BaroTROLL Data Logger (9 m); In-Situ; Fort Collins, CO, USA) sensors were deployed at a subset of sites (n = 9) to collect continuous BP and air temperature data throughout the study period (Fig. 7). Sensors were deployed from a 1-m rope on a tree branch as close as possible to the DO sensor/level logger deployments and recorded BP and air temperature continuously on a 1-min interval. BP data was used convert level logger absolute pressure data to depth data (Supplementary methods).

Level loggers and BP sensors were calibrated by the manufacturer and did not need further calibration. DO sensor calibrations were verified at 100% oxygen saturation in the laboratory prior to deployment and field-based data were corrected for sensor drift by applying a correction factor using lab-generated data (Supplementary methods). Approximately mid-way through the sampling period (August 8–11, 2022), sensors were field-checked for signs of disturbance, exposure due to decreased flows, and biofouling. DO sensors and level loggers were retrieved temporarily for maintenance and data downloads and immediately redeployed in the original location or, if exposed by decreasing flows, redeployed closer to the channel thalweg. Sensors were retrieved the week of August 29–September 1, 2022, and taken to the laboratory for data download.

Figure 7. Example sensor deployment configurations on example study sites. Left panels emphasize the diversity of environmental settings covered in this study. Right panels show two different sensor configurations used to accommodate different environmental conditions. a. Bull Creek (site S63P), Yakama Reservation, Yakima County, Washington, 25 July 2022. b. DO sensor and level logger deployed side-by-side on top of two bricks and secured with zip ties, 29 August 2022. DO sensor equipped with an anti-biofouling copper plate and mesh to inhibit the growth of periphyton on the sensor face. Note the dense accumulation of periphyton and sediment deposited on the sensor bodies and bricks. c. Cle Elum River (site S24R) upstream of Cle Elum Lake, Kittitas County, Washington, 11 August 2022. d. Level logger (far right) and DO sensor equipped with a wiper kit (white sensor body on left-hand side) and anti-biofouling copper plate, 01 September 2022. The white sensor body on the left-hand side of the wiper kit controls the blue flywheel, which is equipped with a brush and programmed to wipe the sensor face 1× every 12 h. Note the lighter accumulation of periphyton and sediment on the equipment deployed in the Cle Elum River (d) compared to the equipment deployed in Bull Creek (b).

3. Channel depth, channel width, and discharge data collection

Depth data for sites located on wadeable streams (*n* = 39) was collected using wading-based depth transects. For each desired reach length (Supplementary methods), we completed 5-10 depth transects over a period of one hour (a minimum of five transects was desired) using 2-person teams, starting ~2 m upstream of the depth sensor. Transects were evenly distributed along the length of the reach. Using a wading rod, depth measurements were collected along the wetted width of the reach at five roughly equidistance points along a transect perpendicular to flow. Field personnel manually recorded the start time and the coordinates for the starting point of each transect; the end time of the last transect was also manually recorded. The channel width of each transect was measured using a range finder; mean channel width was calculated by taking the average of the range finder width measurements.

Depth data for sites located on non-wadeable rivers (*n* = 5) was collected using kayak- or jetboat-mounted sonar (HDS-7 LIVE fish finder with Active Imaging™ 3 in 1 transducer; Lowrance; Tulsa, Oklahoma, USA) or, in the case of very large or inaccessible rivers (*n* = 4), publicly available gauging station data. Kayak-mounted sonar data was collected along each non-wadeable reach by 3-person kayak teams using a sonar receiver mounted to the bottom of a fishing-style kayak. Continuous point measurements were collected by turning on the sonar at the top of each reach and floating downstream using a consistent, zig-zag pattern from bank-to-bank along the entire length of the reach. Coordinates for the top and bottom of each reach and sonar start and end times were manually recorded. HOBO level loggers mounted on a meter stick were used to record depth in reach sections too shallow to collect accurate sonar data. Mean channel width for non-wadeable rivers where depth data was collected using sonar was calculated for each reach by averaging a minimum of five width estimates from the sonar map. For two sites on the Naches River, WA, river depth and channel width were determined from a rating curve developed by the Washington Department of Ecology (https://apps.ecology.wa.gov/publications/SummaryPages/1803009.html). For three sites on the Yakima River, river depth and channel width were determined from a rating curve developed using data from the United States Geological Survey (USGS) gauging station at Union Gap, WA, USA (https://waterdata.usgs.gov/nwis/inventory?agency_code=USGS&site_no=12500450).

Stream discharge and velocity for each site were extracted from the National Oceanic and Atmospheric Administration (NOAA) National Water Model Retrospective Dataset Version 2.1 (NWM V2.1; https://registry.opendata.aws/nwm-archive/). This dataset provides a 42-y (February 1979 through December 2020) hourly simulation at all sampling locations. We computed the August monthly average discharge and velocity across these simulation years as input variables for streamMetabolizer model runs (Section 5). Median grain size (d50) data for all the sites were extracted from NEXSS[10].

4. Surface water sampling

During the week of August 8–11, 2022, surface water samples were collected in triplicate from each site for different analysis. A set of samples were collected using a 50-mL syringe and filtered into 40-mL amber glass vials (ThermoScientific™ Amber Clean Snap Vials) using a 0.22-µm sterivex filter (MilliporeSigmaTM Sterivex™ Sterile Pressure-Driven Devices; MilliporeSigmaTM; Burlington, Massachusetts, USA). Filtered samples were analyzed for dissolved organic carbon (DOC) and total dissolved nitrogen (TN). Additionally, one unfiltered water sample was collected using a pre-washed 2-L amber bottle (Nalgene™ Rectangular Amber HDPE Bottles; ThermoFisher Scientific; Waltham, Massachusetts, USA) for total suspended solids (TSS). All samples were stored on ice in the field and then refrigerated at 4° C before shipping for analysis at the Pacific Northwest National Laboratory (PNNL) Marine and Coastal Research Laboratory in Sequim, Washington (DOC, TDN) or transported to PNNL Biological Sciences Facility Laboratory in Richland, Washington (TSS). TSS samples were analyzed within one week of collection and DOC and TDN content were measured within two weeks of collection.

Dissolved organic carbon (DOC) and TN were measured using a Shimadzu TOC-L Total Organic Carbon Analyzer. DOC was measured as non-purgeable organic carbon and TN was measured by converting nitrogen in the samples to nitrogen gas and detected via chemiluminescence. Visual checks of calibration curves, samples, blanks, and check standard peaks were performed before exporting data. Concentrations below the limit of detection (LOD) of the instrument, or below the standard curve, were flagged[71]. If the concentration of an analyte

was below the LOD, the flag for that sample was replaced with a value corresponding to half of the lowest concentration measured for that analyte. The coefficient of variation (CV) for each replicate set was calculated. Replicates with CV of 30 % or higher were flagged and the outlier sample was identified by calculating the distance between each of the replicate samples. Replicates with the highest distance to the other two replicates were flagged as outliers and removed from the analysis. Mean values for each site were then calculated from the remaining replicates and inputted as predictor variables in the regression analysis (Section 8).

Total suspended solids (TSS) samples were filtered in the laboratory through a pre-weighed and pre-combusted 4.7 cm, 0.7 µm GF/F glass microfiber filter (Whatman™ glass microfiber filters, Grade 934-AH®; MilliporeSigma; Burlington, Massachusetts, USA). After water filtration, the filter and filtration apparatus were rinsed with 30 mL of ultrapure Milli-Q water (Milli-Q® IQ Water Purification System; MilliporeSigma; Burlington, Massachusetts, USA) to ensure that all residue was captured by the filter. The filter was placed in foil and oven dried overnight at 45° C. The filter was allowed to cool in a desiccator prior to weighing. TSS was calculated as the difference between the weight of the filter before and after filtration of the water sample divided by the volume of water filtered[72]. Similar to the approach applied to DOC/TN, if the TSS of a sample was below the LOD, the flag for that sample was replaced with a value corresponding to half of the LOD for the TSS method.

5. Total Ecosystem Respiration (ERtot) estimates

We estimated ERtot using the streamMetabolizer R package[17]. Model inputs consisted of time series data for DO, water temperature, barometric pressure (BP), adjusted time series of depth (Supplementary methods), site ID, site latitude and longitude and a site-specific prior estimate of gas exchange rate coefficient (K_{600}, d^{-1}). Gas exchange (K_{600}, d^{-1}) is a free parameter estimated within streamMetabolizer via Bayesian inference and therefore requires a prior estimate for K_{600}. For this, we first estimated a prior k_{600}($m \cdot d^{-1}$), value using the empirical relationship described in Model Equation 7 from Table 1 in Raymond et al. (2012)[73] as well as site specific slope, discharge, and velocity (Section 1 and 3). k_{600} values were then divided by the reach-averaged depth (Supplementary methods) to obtain K_{600}.

Before running the model, the 15-min time series BP and reach-averaged depth were interpolated (zoo::na.approx function)[74] to match the time interval of the water temperature and DO. Then the temperature, DO, and BP 1-min time series data were down-sampled to a 15-min resolution. Down-sampling improved both solution time and K_{600} pooling. DO time series was visually inspected and days with odd DO signals or biofouling were removed from the timeseries prior running the model. Additionally, photosynthetically active radiation (PAR) was estimated using the streamMetabolizer::calc_light() function[17], which is based on solar time, latitude, and longitude. Saturated DO concentration was calculated for each DO time series using water temperature and BP times series data and the Garcia and Gordon (1992)[75] equation (streamMetabolizer::calc_DO_sat function)[17].

For most of the sites, the Bayesian model was configured with normal K_{600} pooling, both observation and process error, K600_daily_meanlog_meanlog as the K_{600} natural log of the K_{600} prior estimate, a K_{600} scale value of 0.7, K_{600} daily sigma value of 0.02, 1000 burn-in steps, and 1000 saved steps. Some sites failed to converge with the settings mentioned, thus the K_{600} daily sigma value was changed to 0.01 or 0.05 depending on the site. Additionally, the model failed to converge at one site (S63P) due to noisy DO data, which was ameliorated by Fourier-smoothing the DO data using the R package fda.usc with a basis of 361[76]. The Fourier-smoothed data was then compared to the input data; where the two datasets diverged by more

than 1%, the input data was replaced by the Fourier-smoothed data. All other processing steps were identical across sites.

StreamMetabolizer outputs were comprised of ERtot (g O_2 m^{-2}·d^{-1}), GPP (g O_2 m^{-2}·d^{-1}), and K_{600} (d^{-1}) as mean daily estimates over the deployment time[77]. Finally, we calculated the mean GPP and ER for the deployment time for each site.

6. Water column respiration (ERwc) and sediment-associated respiration (ERsed) estimates

We estimated water column respiration in triplicate for 90 min at each site during sampling week using the *in situ* dark bottle incubation approach previously described by Fulton et al. (2024)[35]. Briefly, a DO sensor (miniDOT Logger; Precision Measurement Engineering, Inc., Vista, CA, USA) and a small, battery-powered mixing device (i.e., toy boat motor propellor) were placed inside a 2-L dark bottle in triplicate (Nalgene™ Rectangular Amber HDPE bottles; ThermoFisher Scientific, Waltham, Massachusetts, USA). DO sensors and mixing devices were placed in a cooler with blue ice packs to keep them cool and minimize the time needed at each site for the sensors to equilibrate with river water temperatures. Bottles with the sensor and mixing device were rinsed three times with river water, and filled with river water from the thalweg, ensuring no air bubbles were trapped inside. Once full, the bottles were deployed on the riverbed as close to the thalweg as possible, depending on river depth and access to the shoreline, and weighted down with bricks or larger rocks. The sensors recorded DO concentration and temperature at 1-min intervals during the 90-min incubation period.

We extracted the raw DO and temperature sensor data for each site and plotted each variable against incubation time for each set of triplicate incubations (n = 144). The plots were visually inspected to a) confirm that temperature sensors were at equilibrium with the river temperature when the 90-min incubation test period began and b) identify data gaps, outliers, and other data anomalies. Following the visual inspection of plots, we applied the DO correction factor (as described in Supplementary Methods) and removed the first and last 5 min of the 90-min incubation time series data to account for data anomalies due to emptying and refreshing river water in the bottles at the beginning and end of the incubation period.

Water column respiration rates (ERwc) for individual triplicate incubation samples were calculated as the slope of the linear regression model (R stats::lm function)[78] between DO and incubation time in units of d^{-1}. The volumetric daily rate was then converted to daily areal units (mg O_2 m^{-2}·d^{-1}) by multiplying by the reach-averaged depth (Supplementary methods). Only sample observations with a normalized root mean square error (NRMSE) ≤ 0.03 were used for further analysis[79]. All triplicate samples met the NRMSE criteria. Median, mean, and range of ERwc values for each site were calculated. More than one-third of ERwc values were slightly positive (i.e., $ER_{wc} > 0$). Although positive respiration rates are biologically unrealistic, and values greater than 0 mg O_2 m^{-2}·d^{-1} but less than 0.5 mg O_2 m^{-2}·d^{-1} are difficult to distinguish from zero, thus we retained positive ERwc values.

Sediment-associated respiration (ERsed) estimates were performed by combining streamMetabolizer outputs of mean total ecosystem respiration (ERtot) with in situ dark bottle incubations to isolate water column respiration (ERwc). Respiration directly and indirectly associated with sediments (ERsed) was found by the difference between ERtot and ERwc. These, ERsed estimates include sediment-associated respiration from all ecosystem components besides the water column.

7. Model-predicted hyporheic zone respiration

Model-predicted HZ respiration values were previously published in Son et al., 2022[13], and the data is publicly available on ESS-DIVE[80]. Respiration values reported by Son et al. (2022)[13] correspond to the amount of oxygen respiration in the HZ via vertical and lateral hyporheic exchange between the stream and the HZ, respectively. Therefore, total model-predicted HZ respiration is a sum of vertical and lateral respiration. The respiration is scaled by the stream surface area. The respiration values reported in Son et al., 2022[13] are in units of g CO_2 $m^{-2} \cdot d^{-1}$, however those units were converted for the manuscript. The original model outputs have units of g O_2 $m^{-2} \cdot d^{-1}$. Therefore, we selected the original model outputs for our comparisons as units match the units from ERtot field estimates.

8. Data analysis

To study how physical variables covary with in ERsed at different sites, a LASSO (Least Absolute Shrinkage and Selection Operator) regression model was built using measured and estimated variables (Table 4) as inputs, and ERsed as the target variable. LASSO regression is a sort of linear regression method for modeling relationship between a target variable and one or more exploratory (input) variables by fitting regularized least squares model. It aims to identify the variables and corresponding regression coefficients that lead to a model that minimizes the regularized least squares by imposing a constraint on the model parameters, which 'shrinks' the regression coefficients towards zero. The objective function for LASSO regression is known as L1 loss. The equation for L1 loss is as follows:

$$\min_{\beta_0,\beta} \sum_{i=1}^{n} (y_i - (\beta_i x_i^T + \beta_0))^2 + \lambda \sum_{j=1}^{p} |\hat{\beta}_j| \qquad \text{(Eq. 1)}$$

where n is the number of instances (sites), p is the number of physical variables using as the inputs, y_i is the target variable ERsed, $x_i \coloneqq (x_1, x_2, \dots, x_p)^T$, $\beta \coloneqq (\beta, \beta_2, \dots, \beta_p)$ is the coefficient vector, λ is the parameter that controls the strength of the regularization. Variables with a regression coefficient of zero after shrinkage are excluded from the model and thus neglected for the evaluation of model output. The optimal λ value is determined by using k-fold cross-validation approach[81]. A particular advantage with this technique is that it reduces overfitting without restricting a subset of the dataset to sole use for internal validation. Prior to conducting the LASSO regression, log-transformation with the base-10 logarithm was made to target variable and physical variables with skewed distribution, to meet the normality assumption. Additionally, standardization was applied to achieve a mean of zero and standard deviation of 1 (i.e. z-score) for all variable to ensure that each variable is equally and fairly treated without penalizing a variable simply due to the magnitude or scale of its nature.

To investigate the correlation between RCM-predicted HZ respiration and ERtot, Pearson as well as Spearman's rank correlation were used. Pearson correlation coefficient (r) measures the linear relationship between two continuous variables, assuming that the variables are normally distributed and have a linear relationship. The Spearman's rank-order correlation (ρ) is the non-parametric version of the Pearson correlation, it measures the strength and direction of association between two ranked variables or non-normally distributed continuous variables. In this study, we use Pearson correlation to investigate the correlation between the normalized RCM-predicted HZ respiration and ERtot (with mean of zero, standard deviation of one), and Spearman's rank-order correlation to investigate their rank-order correlation. All analyses were performed using R Statistical Software (v4.1.3)[78].

Table 4. Measured variables used as inputs in the LASSO regression model. All NHDPlus, StreamCat, NLCD, and ET0 database data were extracted using the custom R script[66].

Variable	Long name	Unit	Data Source	Data Source Citation
HOBO Temp	HOBO water temperature	C°	Field measured	Delgado et al. (2023)[72]
Mean Depth	average water depth	cm	Field estimated	Delgado et al. (2023)[72]
Slope	stream slope	NA	NHDPlus	McKay et al. (2012)[68]
Velocity	stream velocity	$m\ s^{-1}$	National Water Model	Cosgrove et al. (2024)[82]
Discharge	discharge	$m^3\ s^{-1}$	National Water Model	Cosgrove et al. (2024)[82]
DOC	non-purgeable organic carbon	$mg\ L^{-1}$	Field measured	Forbes et al. (2023)[71]
TN	total dissolved nitrogen	$mg\ L^{-1}$	Field measured	Forbes et al. (2023)[71]
TSS	total suspended solids	$mg\ L^{-1}$	Field measured	Delgado et al. (2023)[72]
totdasqkm	total drainage area	km^2	NHDPlus	McKay et al. (2012)[68]
PctFst	percentage of forest	%	StreamCat and NLCD	Hill et al. (2016)[67] Jin et al. (2023)[83]
PctShrb	percentage of shrub	%	StreamCat and NLCD	Hill et al. (2016)[67] Jin et al. (2023)[83]
Aridity	aridity	NA	ET0 Database	Zomer et al. (2012)[84]
D50	median grain size	m	NEXSS model output	Kaufman et al. (2023)[77]
GPP	gross primary production	$gO_2\ m^{-2}{\cdot}day^{-1}$	streamMetabolizer output	Kaufman et al. (2023)[77]

We further used a segmented regression, also known as piecewise regression[85], to explore the relationship between average ERsed and GPP. We applied this technique because a non-linear relationship between ERsed and GPP was observed, thus a segmented regression was used to identify the breakpoint(s) and model their correlation at the different segments using the segmented::segmented() function in R[86].

Data Availability:

Data were published previously in Forbes et al. (2023)[71], Delgado et al. (2023)[72], and Kaufman et al. (2023)[77] in the ESS-DIVE repository and are licensed for reuse under the Creative Commons Attribution 4.0 International License.

Code Availability

Scripts necessary to reproduce the primary results of this article are available at https://github.com/river-corridors-sfa/SSS_metabolism.

References:

1. Battin, T. J. *et al.* River ecosystem metabolism and carbon biogeochemistry in a changing world. *Nature* **613**, 449–459 (2023).
2. Naegeli, M. W. & Uehlinger, U. Contribution of the Hyporheic Zone to Ecosystem Metabolism in a Prealpine Gravel-Bed-River. *Journal of the North American Benthological Society* **16**, 794–804 (1997).
3. Battin, T. J., Kaplan, L. A., Newbold, J. D. & Hendricks, S. P. A mixing model analysis of stream solute dynamics and the contribution of a hyporheic zone to ecosystem function*. *Freshwater Biology* **48**, 995–1014 (2003).
4. Ward, N. D. *et al.* Velocity-amplified microbial respiration rates in the lower Amazon River. *Limnology and Oceanography Letters* **3**, 265–274 (2018).
5. Fuss, C. & Smock, L. Spatial and temporal variation of microbial respiration rates in a blackwater stream. *Freshwater Biology* **36**, 339–349 (1996).
6. Kaplan, L. A. & Newbold, J. D. 10 - Surface and Subsurface Dissolved Organic Carbon. in *Streams and Ground Waters* (eds. Jones, J. B. & Mulholland, P. J.) 237–258 (Academic Press, San Diego, 2000). doi:10.1016/B978-012389845-6/50011-9.
7. Jones, J. Factors controlling hyporheic respiration in a desert stream. *Freshwater Biology* **34**, 91–99 (1995).
8. Pusch, M. & Schwoerbel, J. Community respiration in hyporheic sediments of a mountain stream (Steina, Black Forest). *Archiv fur Hydrobiologie* **130**, 35–35 (1994).
9. Fellows, C. S., Valett, M. H. & Dahm, C. N. Whole-stream metabolism in two montane streams: Contribution of the hyporheic zone. *Limnology and Oceanography* **46**, 523–531 (2001).
10. Gomez-Velez, J. D., Harvey, J. W., Cardenas, M. B. & Kiel, B. Denitrification in the Mississippi River network controlled by flow through river bedforms. *Nature Geosci* **8**, 941–945 (2015).
11. Fang, Y. *et al.* High-Performance Simulation of Dynamic Hydrologic Exchange and Implications for Surrogate Flow and Reactive Transport Modeling in a Large River Corridor. *Frontiers in Water* **2**, (2020).
12. Fang, Y. *et al.* A multirate mass transfer model to represent the interaction of multicomponent biogeochemical processes between surface water and hyporheic zones (SWAT-MRMT-R 1.0). *Geoscientific Model Development* **13**, 3553–3569 (2020).
13. Son, K., Fang, Y., Gomez-Velez, J. D. & Chen, X. Spatial Microbial Respiration Variations in the Hyporheic Zones Within the Columbia River Basin. *Journal of Geophysical Research: Biogeosciences* **127**, (2022).
14. Odum, H. T. Primary Production in Flowing Waters. *Limnology and Oceanography* **1**, 102–117 (1956).
15. Hall, R. O. & Hotchkiss, E. R. Chapter 34 - Stream Metabolism. in *Methods in Stream Ecology (Third Edition)* (eds. Lamberti, G. A. & Hauer, F. R.) 219–233 (Academic Press, 2017). doi:10.1016/B978-0-12-813047-6.00012-7.
16. Almeida, G. H., Boëchat, I. G. & Gücker, B. Assessment of stream ecosystem health based on oxygen metabolism: Which sensor to use? *Ecological Engineering* **69**, 134–138 (2014).
17. Appling, A. P., Hall, R. O., Yackulic, C. B. & Arroita, M. Overcoming Equifinality: Leveraging Long Time Series for Stream Metabolism Estimation. *Journal of Geophysical Research: Biogeosciences* **123**, 624–645 (2018).
18. Tromboni, F. *et al.* High rates of daytime river metabolism are an underestimated component of carbon cycling. *Commun Earth Environ* **3**, 1–10 (2022).
19. Marzolf, N. S. & Ardón, M. Ecosystem metabolism in tropical streams and rivers: a review and synthesis. *Limnology and Oceanography* **66**, 1627–1638 (2021).

20. Kaylor, M. J., White, S. M., Saunders, W. C. & Warren, D. R. Relating spatial patterns of stream metabolism to distributions of juveniles salmonids at the river network scale. *Ecosphere* **10**, e02781 (2019).
21. Hotchkiss, E. R. *et al.* Sources of and processes controlling CO2 emissions change with the size of streams and rivers. *Nature Geosci* **8**, 696–699 (2015).
22. O'Donnell, B. & Hotchkiss, E. R. Resistance and resilience of stream metabolism to high flow disturbances. *Biogeosciences* **19**, 1111–1134 (2022).
23. Gómez-Gener, L. *et al.* Towards an improved understanding of biogeochemical processes across surface-groundwater interactions in intermittent rivers and ephemeral streams. *Earth-Science Reviews* **220**, 103724 (2021).
24. Bernot, M. J. *et al.* Inter-regional comparison of land-use effects on stream metabolism. *Freshwater Biology* **55**, 1874–1890 (2010).
25. Song, C. *et al.* Continental-scale decrease in net primary productivity in streams due to climate warming. *Nature Geosci* **11**, 415–420 (2018).
26. Bernhardt, E. S. *et al.* Light and flow regimes regulate the metabolism of rivers. *Proceedings of the National Academy of Sciences* **119**, e2121976119 (2022).
27. Gurung, A., Iwata, T., Nakano, D. & Urabe, J. River Metabolism along a Latitudinal Gradient across Japan and in a global scale. *Sci Rep* **9**, 4932 (2019).
28. Hall Jr, R. O., Kennedy, T. A. & Rosi-Marshall, E. J. Air–water oxygen exchange in a large whitewater river. *Limnology and Oceanography: Fluids and Environments* **2**, 1–11 (2012).
29. Hall Jr., R. O. *et al.* Turbidity, light, temperature, and hydropeaking control primary productivity in the Colorado River, Grand Canyon. *Limnology and Oceanography* **60**, 512–526 (2015).
30. Roche, K. R. & Dentz, M. Benthic Biolayer Structure Controls Whole-Stream Reactive Transport. *Geophysical Research Letters* **49**, e2021GL096803 (2022).
31. Schindler, D. E., Jankowski, K., A'mar, Z. T. & Holtgrieve, G. W. Two-stage metabolism inferred from diel oxygen dynamics in aquatic ecosystems. *Ecosphere* **8**, e01867 (2017).
32. Hotchkiss, E. R. & Hall, Jr., Robert O. High rates of daytime respiration in three streams: Use of δ18OO2 and O2 to model diel ecosystem metabolism. *Limnology and Oceanography* **59**, 798–810 (2014).
33. del Giorgio, P. A., Cole, J. J. & Cimbleris, A. Respiration rates in bacteria exceed phytoplankton production in unproductive aquatic systems. *Nature* **385**, 148–151 (1997).
34. Appling, A. P. *et al.* The metabolic regimes of 356 rivers in the United States. *Scientific Data* **5**, 180292 (2018).
35. Fulton, S. G. *et al.* Yakima River Basin Water Column Respiration is a Minor Component of River Ecosystem Respiration. *EGUsphere* **2024**, 1–27 (2024).
36. Ward, N. D., Sawakuchi, H. O., Richey, J. E., Keil, R. G. & Bianchi, T. S. Enhanced Aquatic Respiration Associated With Mixing of Clearwater Tributary and Turbid Amazon River Waters. *Front. Earth Sci.* **7**, (2019).
37. Reisinger, A. J. *et al.* Water column contributions to the metabolism and nutrient dynamics of mid-sized rivers. *Biogeochemistry* **153**, 67–84 (2021).
38. Walters, E., Graml, M., Behle, C., Müller, E. & Horn, H. Influence of particle association and suspended solids on UV inactivation of fecal indicator bacteria in an urban river. *Water, Air, & Soil Pollution* **225**, 1–9 (2014).
39. Frost, P. C., Evans-White, M. A., Finkel, Z. V., Jensen, T. C. & Matzek, V. Are you what you eat? Physiological constraints on organismal stoichiometry in an elementally imbalanced world. *Oikos* **109**, 18–28 (2005).
40. Jäger, C. G., Diehl, S. & Schmidt, G. M. Influence of water-column depth and mixing on phytoplankton biomass, community composition, and nutrients. *Limnology and Oceanography* **53**, 2361–2373 (2008).

41. Reynolds, C. S. Hydroecology of river plankton: the role of variability in channel flow. *Hydrological Processes* **14**, 3119–3132 (2000).
42. Newcomer, M. E. *et al.* Influence of Hydrological Perturbations and Riverbed Sediment Characteristics on Hyporheic Zone Respiration of CO2 and N2. *Journal of Geophysical Research: Biogeosciences* **123**, 902–922 (2018).
43. Roche, K. R. *et al.* Benthic biofilm controls on fine particle dynamics in streams. *Water Resources Research* **53**, 222–236 (2017).
44. Krause, S. *et al.* Ecohydrological interfaces as hot spots of ecosystem processes. *Water Resources Research* **53**, 6359–6376 (2017).
45. Honious, S. A. S., Hale, R. L., Guilinger, J. J., Crosby, B. T. & Baxter, C. V. Turbidity Structures the Controls of Ecosystem Metabolism and Associated Metabolic Process Domains Along a 75-km Segment of a Semiarid Stream. *Ecosystems* **25**, 422–440 (2022).
46. Ellis, E. E. *et al.* Factors controlling water-column respiration in rivers of the central and southwestern Amazon Basin. *Limnology and Oceanography* **57**, 527–540 (2012).
47. Roley, S. S., Hall Jr., R. O., Perkins, W., Garayburu-Caruso, V. A. & Stegen, J. C. Coupled primary production and respiration in a large river contrasts with smaller rivers and streams. *Limnology and Oceanography* **n/a**, 15 (2023).
48. Hall, R. O., Tank, J. L., Baker, M. A., Rosi-Marshall, E. J. & Hotchkiss, E. R. Metabolism, Gas Exchange, and Carbon Spiraling in Rivers. *Ecosystems* **19**, 73–86 (2016).
49. Battin, T. J., Kaplan, L. A., Newbold, J. D., Cheng, X. & Hansen, C. Effects of Current Velocity on the Nascent Architecture of Stream Microbial Biofilms. *Appl Environ Microbiol* **69**, 5443–5452 (2003).
50. Cohen, S., Wan, T., Islam, M. T. & Syvitski, J. Global river slope: A new geospatial dataset and global-scale analysis. *Journal of hydrology* **563**, 1057–1067 (2018).
51. Anderton, C. R. *et al.* Nitrogen Source Governs Community Carbon Metabolism in a Model Hypersaline Benthic Phototrophic Biofilm. *mSystems* **5**, e00260-20 (2020).
52. Duarte, C. M. & Prairie, Y. T. Prevalence of Heterotrophy and Atmospheric CO2 Emissions from Aquatic Ecosystems. *Ecosystems* **8**, 862–870 (2005).
53. Munn, M. D., Konrad, C. P., Miller, M. P. & Jaeger, K. A comparison of spatial and temporal drivers of stream metabolism. *Freshwater Biology* **68**, 1751–1764 (2023).
54. Griffiths, N. A. *et al.* Agricultural land use alters the seasonality and magnitude of stream metabolism. *Limnology and Oceanography* **58**, 1513–1529 (2013).
55. Summers, B. M. *et al.* Long-term data reveal highly-variable metabolism and transitions in trophic status in a montane stream. *Freshwater Science* **39**, 241–255 (2020).
56. Young, R. G. & Huryn, A. D. Effects of Land Use on Stream Metabolism and Organic Matter Turnover. *Ecological Applications* **9**, 1359–1376 (1999).
57. Hoellein, T. J., Bruesewitz, D. A. & Richardson, D. C. Revisiting Odum (1956): A synthesis of aquatic ecosystem metabolism. *Limnology and Oceanography* **58**, 2089–2100 (2013).
58. McClain, M. E. *et al.* Biogeochemical hot spots and hot moments at the interface of terrestrial and aquatic ecosystems. *Ecosystems* 301–312 (2003).
59. Bernhardt, E. S. *et al.* Control Points in Ecosystems: Moving Beyond the Hot Spot Hot Moment Concept. *Ecosystems* **20**, 665–682 (2017).
60. Wymore, A. S., Ward, A. S., Wohl, E. & Harvey, J. W. Viewing river corridors through the lens of critical zone science. *Front. Water* **5**, (2023).
61. Lewandowski, J. *et al.* Is the Hyporheic Zone Relevant beyond the Scientific Community? *Water* **11**, 2230 (2019).
62. Regier, P. *et al.* Machine Learning Photogrammetric Analysis of Images Provides a Scalable Approach to Study Riverbed Grain Size Distributions. (2023).
63. Qiu, J. *et al.* Implications of water management representations for watershed hydrologic modeling in the Yakima River basin. *Hydrology and Earth System Sciences* **23**, 35–49 (2019).

64. Vaccaro, J. J. & Olsen, T. D. *Estimates of Ground-Water Recharge to the Yakima River Basin Aquifer System, Washington, for Predevelopment and Current Land-Use and Land-Cover Conditions*. (2007).
65. Vano, J. A. *et al.* Climate change impacts on water management and irrigated agriculture in the Yakima River Basin, Washington, USA. *Climatic Change* **102**, 287–317 (2010).
66. Willi, K. & Ross, M. R. V. Geospatial Data Puller for Waters in the Contiguous United States (Version v1). (2023).
67. Hill, R. A., Weber, M. H., Leibowitz, S. G., Olsen, A. R. & Thornbrugh, D. J. The Stream-Catchment (StreamCat) Dataset: A Database of Watershed Metrics for the Conterminous United States. *JAWRA Journal of the American Water Resources Association* **52**, 120–128 (2016).
68. McKay, L. *et al. NHDPlus Version 2: User Guide*. (2012).
69. Zomer, R. & Trabucco, A. Global Aridity Index and Potential Evapotranspiration (ET0) Database: Version 3. 6086884631 Bytes [object Object] https://doi.org/10.6084/M9.FIGSHARE.7504448.V6 (2022).
70. Hijmans, R. J., Bivand, R., Pebesma, E. & Sumner, M. D. terra: Spatial Data Analysis. (2024).
71. Forbes, B. *et al.* WHONDRS River Corridor Dissolved Oxygen, Temperature, Sediment Aerobic Respiration, Grain Size, and Water Chemistry from Machine-Learning-Informed Sites across the Contiguous United States (v3). [object Object] https://doi.org/10.15485/1923689 (2023).
72. Delgado, D. *et al.* Spatial Study 2022: Surface Water Samples, Cotton Strip Degradation, and Hydrologic Sensor Data across the Yakima River Basin, Washington, USA (v2). (2023).
73. Raymond, P. A. *et al.* Scaling the gas transfer velocity and hydraulic geometry in streams and small rivers. *Limnology and Oceanography: Fluids and Environments* **2**, 41–53 (2012).
74. Zeileis, A. & Grothendieck, G. **zoo** : *S3* Infrastructure for Regular and Irregular Time Series. *J. Stat. Soft.* **14**, (2005).
75. Garcia, H. E. & Gordon, L. I. Oxygen solubility in seawater: Better fitting equations. *Limnol. Oceanogr.* **37**, 1307–1312 (1992).
76. Febrero-Bande, M. & De La Fuente, M. O. Statistical computing in functional data analysis: The R package fda. usc. *Journal of statistical Software* **51**, 1–28 (2012).
77. Kaufman, M. H. *et al.* Spatial Study 2022: Water Column, Sediment, and Total Ecosystem Respiration Rates across the Yakima River Basin, Washington, USA. (2023).
78. R Core Team. R: The R Project for Statistical Computing. R Foundation for Statistical Computing (2022).
79. Shcherbakov, M. V. *et al.* A survey of forecast error measures. *World applied sciences journal* **24**, 171–176 (2013).
80. Son, K., Fang, Y., Gomez-Velez, J. D. & Chen, X. Model Inputs, Outputs, and Scripts associated with: “Spatial microbial respiration variations in the hyporheic zones within the Columbia River Basin”. [object Object] https://doi.org/10.15485/1962818 (2023).
81. Bengio, Y. & Grandvalet, Y. No Unbiased Estimator of the Variance of K-Fold Cross-Validation. *J. Mach. Learn. Res.* **5**, 1089–1105 (2004).
82. Cosgrove, B. *et al.* NOAA’s National Water Model: Advancing operational hydrology through continental-scale modeling. *J American Water Resour Assoc* 1752–1688.13184 (2024) doi:10.1111/1752-1688.13184.
83. Jin, S. *et al.* National Land Cover Database 2019: A Comprehensive Strategy for Creating the 1986–2019 Forest Disturbance Product. *J Remote Sens* **3**, 0021 (2023).
84. Zomer, R. J., Xu, J. & Trabucco, A. Version 3 of the Global Aridity Index and Potential Evapotranspiration Database. *Sci Data* **9**, 409 (2022).
85. McZgee, V. E. & Carleton, W. T. Piecewise Regression. *Journal of the American Statistical Association* **65**, 1109–1124 (1970).

86. Muggeo, V. M. Segmented: an R package to fit regression models with broken-line relationships. *R news* **8**, 20–25 (2008).

Acknowledgements:
This work was supported by the River Corridor Science Focus Area (SFA) at the Pacific Northwest National Laboratory (PNNL). The SFA is supported by the United States Department of Energy, Office of Biological and Environmental Research (BER), Environmental System Science (ESS) Program. PNNL is operated by Battelle Memorial Institute for the United States Department of Energy under contract no. DE-AC05-76RL01830. We thank the United States Forest Service, Washington Department of Natural Resources, Washington Department of Fish and Wildlife, Confederated Tribes and Bands of the Yakama Nation, and Cowiche Canyon Conservancy for access to field locations where these samples were collected. We also thank the Confederated Tribes and Bands of the Yakama Nation Tribal Council and Yakama Nation Fisheries for working with us to facilitate sample collection and optimization of data usage according to their values and worldview.

Author Contributions:
JCS, MK, BH and VAGC conceptualized the study. LR and SF conducted the field work and analyzed samples. KS and YF configured RCM model outputs for this study. XL and BF processed the data and performed analyses with feedback from JCS, MK, BH and VAGC. MK and VACG carried out metabolism analysis with feedback from BH. BF managed the data. JCS, MK, SF, BF and VAGC drafted the initial manuscript and all authors contributed to revisions.

Competing interests
All authors declare that there are no competing interests.