

Combined surface-subsurface stream restoration structures can optimize hyporheic attenuation of stream water contaminants

Skuyler P. Herzog^{1,2}, Jason Galloway³⁻⁴, Eddie W. Banks⁵, Malte Posselt⁶, Anna Jaeger³⁻⁴, Andrea Portmann⁷, René Sahm⁸, Björn Kusebauch⁸, Jörg Lewandowski³⁻⁴, and Adam S. Ward^{2,9}*

¹Natural Resources Program, College of Forestry, Oregon State University-Cascades, Bend, Oregon 97702, United States of America

²O'Neill School of Public and Environmental Affairs, Indiana University, Bloomington, Indiana 47405, United States of America

³Department of Ecohydrology, Leibniz-Institute of Freshwater Ecology and Inland Fisheries, 12587 Berlin, Germany

⁴Geography Department, Humboldt University of Berlin, 12489 Berlin, Germany

⁵National Centre for Groundwater Research and Training, and College of Science & Engineering, Flinders University, Adelaide, SA 5001, Australia

⁶Department of Environmental Science, Stockholm University, 11418 Stockholm, Sweden

⁷Department of Civil and Environmental Engineering and Hydrologic Science and Engineering Program, Colorado School of Mines, Golden, Colorado 80401, United States of America

⁸German Environment Agency (Umweltbundesamt), Section IV 2.5 – Trace Analysis, Artificial Ponds and Streams, 12307 Berlin, Germany

⁹Biological and Ecological Engineering Department, Oregon State University, Corvallis, Oregon 97331, United States of America

*Corresponding author.

KEYWORDS: groundwater-surface water interactions; stream restoration; water quality.

ORCID IDs:

- SP Herzog 0000-0003-2905-9578
- J Galloway 0000-0002-5512-6140
- EW Banks 0000-0002-2431-7649
- M Posselt 0000-0001-8979-8044
- A Portmann
- A Jaeger 0000-0002-9141-3235
- R Sahm 0000-0001-8806-265X
- B Kusebauch
- J Lewandowski 0000-0001-5278-129X
- AS Ward 0000-0002-6376-0061

Abstract (196/200 words)

There is a design-to-function knowledge gap regarding how engineered stream restoration structures can maximize hyporheic contaminant attenuation. Surface and subsurface structures

have each been studied in isolation as techniques to restore hyporheic exchange, but surface-subsurface structures have not been integrated nor optimized. Here we used a numerical model to systematically evaluate key design variables for combined surface (i.e., weir height and length) and subsurface (i.e., upstream and downstream baffle plate spacing) structures. We also compared performance metrics that place differing emphasis on hyporheic flux versus transit timescale. We found that surface structures tended to create relatively high flux, short transit time flowpaths, whereas subsurface structures promoted moderate flux, long transit time flowpaths. Combined surface-subsurface structures provided unique hydraulic conditions and could be many times more effective than surface or subsurface structures alone. All water quality metrics were improved by the presence of an upstream plate and the absence of a downstream plate. Increasing weir length tended to improve all metrics, whereas the optimal weir height varied based on metric. We place our results in the context of stream restoration to better align specific restoration goals with appropriate performance metrics and hyporheic structure designs.

Introduction

Making hyporheic contaminant attenuation a primary goal rather than a fortuitous byproduct of stream restoration has been an objective for more than half a century¹, with increasing attention in recent years (e.g.,²⁻⁵). The impact of any hyporheic restoration structure on reach-scale water quality depends on two critical factors: (1) the magnitude of hyporheic flux^{3,6} relative to streamflow, and (2) the alignment between hyporheic transit times and the timescale(s) for the desired reaction(s)⁷⁻⁹. Unfortunately, common restoration structures often only deliver one of these factors at a time, yielding suboptimal or counterproductive water quality results¹⁰⁻¹². Practitioners seeking to attenuate contaminants via the hyporheic zone lack the information they need to tailor stream restoration designs to promote a reaction of interest.

Most techniques for increasing hyporheic exchange flux produce a concomitant reduction in hyporheic transit timescales, thereby reducing reaction progress and canceling out contaminant attenuation benefits¹²⁻¹⁴. For example, hyporheic flux can increase directly with alterations of surface roughness (e.g., increasing structure height for weirs^{13,14}, logs¹⁵, dunes¹⁶, and bars^{17,18}), overall bed slope^{2,14}, or hydraulic conductivity^{1,5,19}. Under typical conditions of shallow streambed depth (i.e., bedrock constraint), any design that increases hyporheic flux also yields a decrease in hyporheic transit times due to increasing porewater velocities or decreasing effective hyporheic depth^{12-14,20}. Conversely, strategies for increasing transit timescales often rely on reducing hyporheic fluxes and porewater velocities (e.g., shorter weir heights¹⁴, removal of surface roughness¹², lower hydraulic conductivity¹³), thereby reducing the proportion of water treated within the hyporheic zone. In other words, the only way to increase transit time along a fixed hyporheic flow tube (i.e., without altering the sediments or flowpath length) is to decrease porewater velocity by decreasing flux.

Hyporheic restoration projects typically focus on surface structures (e.g., weirs^{13,21}) or subsurface structures (e.g., low-permeability blocks in the streambed^{2,5}) rather than combining

and optimizing these elements. We predict that combined surface-subsurface restoration elements may decouple flux from transit time by lengthening hyporheic flow tubes, thereby allowing high-flux flowpaths that also have longer transit times. Specifically, flowpath geometries and transit times have been lengthened by hyporheic caps beneath a plunge pool^{22,23} and with baffle walls placed in the hyporheic zone to direct flowpaths^{1,2,24}. Expanding the hyporheic cross-sectional area^{1,12,14,22,23,25} can also increase fluxes without reducing transit times. However, it is unclear how subsurface design elements may be coupled with surface features, especially in terms of design decisions that can shift hyporheic fluxes and transit times toward optimal contaminant attenuation performance.

Rigorously optimizing hyporheic structures requires the selection of an objective function or performance metric. There is widespread agreement that hyporheic fluxes and transit times must be considered jointly to predict contaminant attenuation outcomes^{4,8,12,14,26}. However, several different metrics are commonly used to evaluate designs, each placing a different emphasis on flux versus transit time. First, the distribution of hyporheic transit times and Damköhler numbers (the product of transit time, t , and first order reaction rate, k)⁸ can both be flux-weighted to generate closely related metrics: flux-weighted transit time distributions (FW-TTD) and Reaction Significance Factors (RSF)⁸, respectively. RSF assesses the alignment of hyporheic flow and reaction timescales as a proxy for water quality impact (e.g.,^{8,27,28}), and implicitly assigns equal weight to long and short flowpaths. Next, mass removal is a more direct metric for assessing contaminant attenuation, but can also be separated into total mass removal (i.e., mass per time) and percent removal of downwelling mass. The former is a stream-centric metric for reach-scale contaminant load reduction, whereas the latter centers on hyporheic concentration reduction to provide high water quality refugia in upwelling zones. In contrast to RSF, mass removal implicitly assigns higher weight to shorter flowpaths. For example, most hyporheic reactions are modeled as first order, which produces greater total mass removal at earlier time points. A flowpath with a transit time of $t=k^{-1}$ removes 63% of the modeled reactant, whereas increasing the transit time to $t=2k^{-1}$ and $t=3k^{-1}$ only yields incremental removal of 23% and 9%, respectively.

Optimal timescales further depend on whether a reaction begins immediately upon entry into the hyporheic zone (e.g., attenuation of dissolved oxygen, metformin, or ammonia) or after other reactants are consumed (e.g., bulk denitrification only after sufficiently low oxygen concentrations exist), or if a primary reactant degrades into a problematic secondary product that likewise must be attenuated (e.g., nitrous oxide, guanylurea, estrone). Taken together, the range of metrics and reaction pathways mean that the optimal design for one metric-reaction combination is not necessarily universal. For example, Liu and Chui¹⁴ found that total mass removal and percent mass removal led to opposing weir designs, and thus used the product of total and percent mass removal as the objective function to optimize weir design for hyporheic contaminant attenuation. Similarly, reactions with immediate onset may also lead to diametrically opposed restoration approaches compared to reactions with delayed onset¹². However, the sensitivities of structure designs to performance metrics and reaction type have not been fully explored.

Our objective is to reveal the connections between restoration design, hydraulics, and contaminant attenuation in combined surface-subsurface hyporheic structures. Specifically, we ask (1) *How can the known controls on hyporheic exchange be used to optimize contaminant attenuation by maximizing flux at the timescales required for a given process or reaction of interest?* and (2) *How does the choice of an evaluation metric influence the optimal design selection?* To answer these questions we used a numerical model, based on prior field-scale (100-m) physical flume experiments²⁹, to simulate hyporheic water quality outcomes of combined surface-subsurface architecture. We provide a systematic analysis addressing the key design questions above by simulating the interplay of surface and subsurface design variables to optimize hyporheic fluxes and transit times, along with the sensitivity of optimization outcomes to the metric of choice. This approach explores design criteria linking surface and subsurface morphological features (means) to hyporheic contaminant attenuation (ends).

Methods

Numerical Modeling

We constructed a 2-D finite element model in COMSOL Multiphysics Version 5.5 (COMSOL, Inc., Burlington, Massachusetts) to study the impact of surface and subsurface architecture on hyporheic exchange. Model construction and interpretation of results broadly follow similar applications^{5,30–32}. The model had $318,760 \pm 1,574$ (mean $\pm$ standard deviation) triangular mesh elements of 1.4 to 3.5 mm in height, with minimum mesh quality of at least 0.75. The model was motivated by a prior set of physical flume experiments investigating surface (submerged rock weirs placed *on the streambed sediment*) and subsurface (baffle wall plates buried *in the streambed sediment*) stream restoration structures to increase hyporheic attenuation of wastewater-derived contaminants²⁹. While the model was not calibrated to the flume studies because sufficient data for calibration of the surface and subsurface models were not collected during those field experiments, we take the available data as a basis to construct our simulations, linking our results to empirical studies. Thus, our simulations are heuristic, intended to provide a conceptual yet realistic representation of surface and hyporheic flow around the restoration structures without requiring calibration nor validation. The numerical model represented a longitudinal-section along the flume centerline (Figure 1). Coupled surface and subsurface flows were represented using the Free and Porous Media Flow physics interface, part of the Subsurface Flow Module in COMSOL Multiphysics. Surface flow was described by the Navier-Stokes equations and porous media flow was governed by the Brinkman equations, with continuous pressure and velocity fields across the surface water and porewater domains.

To minimize the influence of boundary conditions at up- and downstream ends of the simulation while balancing computational demand, the model domain length was manipulated during preliminary analysis. Flow patterns stabilized within one meter of each boundary condition, so the domain length was set to 5 m long (x-direction), with structures located at the model midpoint. The model geometry reflected the average water and sediment characteristics from the field flume experiments (Table 1). Boundary conditions for the surface flow domain were constant velocity at the upstream boundary, constant pressure at the downstream boundary, and no slip at the upper (water-air interface) boundary. The constant velocity at the

upstream boundary was set to the average velocity measured in the sediment bed sections (i.e., without weirs or other obstructions) of the field flume experiments. The sediment-water interface was coupled, and all other sediment boundaries were no-flow. Weirs and plates were represented as no-flow features.

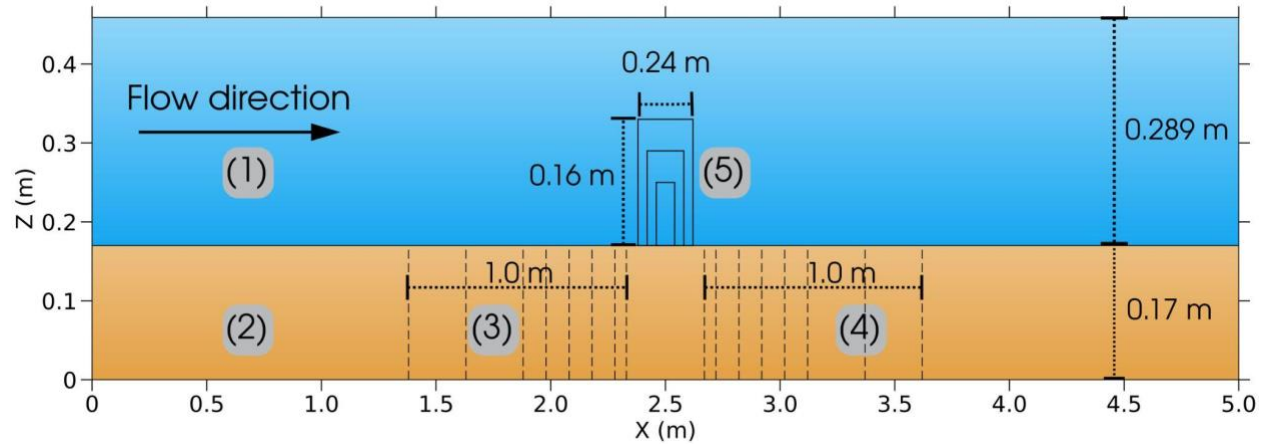


Figure 1. Model domain and design variable ranges. Showing the surface water domain (1), the sediment domain (2), the range of upstream plate positioning (3), the range of downstream plate positioning (4), and the range of weir dimensions (5). Plate positioning was measured from the closest point of the weir to the plate. In addition to the parameter ranges shown here, additional combinations of structure variables included control conditions without weirs and without plates. See Tables 1-2 for a complete list of domain parameters and design variable ranges. Flow direction is indicated by the blue arrow. Note the aspect ratio of the y-axis to the x-axis is approximately 10:3.

Table 1. Flume and sediment properties used to parameterize the model scenarios.

Parameter	Units	Value
Hydraulic conductivity	m/s	1.5×10^{-4}
Porosity	-	0.2
Slope	-	0
Sediment depth	m	0.17
Water depth	m	0.289
Influent surface velocity	m/s	0.08

Simulated Hyporheic Restoration Structures

All scenarios and modeled hyporheic restoration structures were based on prior physical flume experiments²⁹. Simulated weir heights were always within the range observed in the flume (i.e., 0.08 m to 0.16 m), which did not impact the surveyed water surface in flume experiments²⁹. Some flume scenarios also included steel plates buried in the sediment to form impermeable barriers that created hyporheic turnover. Like the physical plates in the flume, simulated subsurface plates were 0.002 m thick and designed to block the entire sediment cross-sectional

area without protruding into the surface. Thus, no portion of the plates were visible from the surface and the plate formed a continuous contact with the bottom of the flume. We used the flume dimensions as a starting point for our model experiments, but varied the surface and subsurface structure dimensions and spacings over a far broader parameter space than was possible in the physical flume studies. Specifically, the weir heights (kept within observed range from Posselt²⁹), weir lengths, and plate spacings were all varied systematically (Table 2).

Table 2. Modeled scenario variables and parameter values. Structures were named according to the formula: *Weir [Weir height]x[Weir length] Plates [Upstream plate spacing]/[Downstream plate spacing]* with all units in centimeters (e.g., *Weir 12x8 Plates 30/30*; *Weir 8x24 Plates absent/50*). We simulated the full factorial set of combinations for the set of parameter values. The 81 combinations of plate spacings without any weir fulfilled the “absent” weir height and weir length designations simultaneously.

Variable	Set of Parameter Values (cm)
Weir height	Absent, 8, 12, 16
Weir length	Absent, 8, 16, 24
Upstream plate distance from weir	5, 10, 20, 30, 40, 50, 75, 100, Absent
Downstream plate distance from weir	5, 10, 20, 30, 40, 50, 75, 100, Absent

Hyporheic Zone Response Variables

We used particle tracing to quantify hyporheic flow patterns around the structures. Particles were released along the sediment-water interface every 0.001-m along the downwelling zone. Particles were used to record vertical velocity along the sediment water interface, flowpath geometries, and transit times for each particle that downwelled upon release. Flowpath tracking ended when the particles returned to the surface water domain.

Flux-weighted transit times for each flowpath were calculated as:

$$\text{EQ 1. } \textit{flux weighted transit time} = q_i \times t_i$$

where q_i is the downwelling velocity of each particle multiplied by the sediment porosity [m/s], t_i is the transit time for each particle [s], and the FW-TTD has units of [m]. The FW-TTD was not normalized by the total downwelling flux, which varied by structure, because the absolute magnitude of the FW-TTD is an important evaluation criterion to differentiate between alternative designs.

The Reaction Significance Factor (RSF) for each flowpath was adapted from Harvey et al.⁴ as:

$$\text{EQ 2. } RSF = \left(\frac{Q_i}{Q_{stream}} \right) \times t_i \times k = \left(\frac{Q_i}{Q_{stream}} \right) \times Da_i$$

where Q_i [m^3/s] is the 3-D hyporheic downwelling flux found by multiplying the 1-D hyporheic downwelling flux of each particle [m/s] by the particle spacing (0.001 m) and representative flume width (0.98 m), Q_{stream} [m^3/s] is the inflow velocity at the upstream boundary condition (0.08 m/s) multiplied by the water depth (0.289 m) and representative flume width (0.98 m), k is a first order reaction rate constant (0.001 s^{-1} , an illustrative value chosen to align with the simulated flow timescales), and Da is the Damköhler number of each flowpath. The reaction rate was uniform with respect to space and time, which was a necessary simplifying assumption (e.g., ⁴). In our study, the only difference between FW-TTD and RSF was the scalar multiplier $4.3 \times 10^{-5} \text{ m}^{-1}$, so the two metrics show the same trends but have different absolute values. For brevity we only report RSF, which can also be thought of as a flux-weighted Damköhler number. RSF is most commonly reported in field studies that use the average transit time at a given sampling point, whereas we calculated RSF for each individual flowpath.

For each simulated structure, the total mass attenuated along each flowpath was calculated according to two generic first-order reactions (after ¹²). For these generic reactions, attenuation is assumed to represent permanent removal, so both terms are used interchangeably. The first reaction (hereafter ‘primary’ reaction) began immediately upon downwelling and was intended to be representative of aerobic reactions such as the attenuation of dissolved oxygen³³, metformin²⁹, or ammonia³³. The second reaction (hereafter ‘secondary’ reaction) was parameterized with a threshold behavior where reaction began after 80% of the primary mass had been consumed. The secondary reaction represented delayed onset within the hyporheic zone, such as denitrification occurring after bulk anoxic conditions develop³³, sulfate reduction following the consumption of nitrate, or guanyurea metabolism subsequent to its formation by hyporheic degradation of metformin as studied in the flume experiments by Posselt²⁹. In addition to representing different contaminants, prior research showed that primary and secondary reactions may lead to fundamentally different design criteria (e.g., $Da \ll 1$ for primary reaction but $Da \approx 1$ for secondary reaction¹²). Both reactions were modeled according to the following equation:

$$\text{EQ 3. } \dot{m} = \sum_{i=1}^n Q_i \times C_0 (1 - e^{-k \times t_i})$$

where $\dot{m}$ is mass removal rate [g/s], n is the number of downwelling particles in each simulation, Q_i is the 3-D vertical flux at each downwelling particle [m^3/s], C_0 is the initial concentration of reactant [$1 \text{ g}/\text{m}^3$], k is the first order reach rate constant [0.001 s^{-1} for both reactions], and t_i is the transit time for each particle [s]. The primary reaction used the total hyporheic transit time for each particle. As the secondary reaction did not begin until 80% of the primary mass had been consumed, 1,609 seconds (i.e., $\text{LN}(\frac{0.2}{k})$) were subtracted from each particle transit time to account for delayed onset of the secondary reaction. Any transit times that became negative after the subtraction were replaced with zero, representing flowpaths where no secondary reaction occurs.

Finally, the total mass removed was divided by the amount of mass that downwelled to calculate percent mass removal in each structure:

$$\text{EQ 4. } \% \text{ removal} = \frac{\dot{m}}{\sum_{i=1}^n Q_i \times C_0} \times 100$$

Percent removal affects the quality of upwelling water and the area of hyporheic sediment that can provide improved water quality for hyporheic organisms. For example, a structure may be designed to protect coho salmon eggs by maximizing the streambed volume with low concentrations of 6-PPD and 6-PPD quinone³⁴. Alternatively, locations with frequent lampricide applications may seek to minimize percent removal to prevent lamprey refugia³⁵. Habitat-centric metrics like percent mass removal provide an important comparison to stream-centric metrics such as total mass removal. One design optimization study of weir-driven hyporheic exchange¹⁴ assigned the same weighting to percent mass removal and total mass removal, thus we consider both metrics in our study.

In total we completed 810 unique hyporheic restoration structure design scenarios, each with 1,201 particles, resulting in more than 970,000 individual particle traces. First, we calculated the general trends for each design variable by averaging the results. For example, the mean flux and mean transit time at each weir height were calculated by averaging across all 243 combinations of weir lengths and plate spacings at each weir height. Average values for the “Weir Absent” (i.e., weir height and weir length both set to zero) scenarios include the 81 plate combinations without weirs. These analyses were repeated for each design variable. Second, we evaluated each of the 810 individual simulations according to the metrics above to determine the overlap between the general patterns that improve contaminant attenuation and the specific combinations that optimize contaminant attenuation.

Results

Structure design controls hyporheic flowpath geometries

Surface and subsurface structure manipulations both influenced hyporheic flowpath geometries (Figure 2). The control condition without plates or weirs had negligible exchange (Figure 2 A) compared to scenarios with a weir (Figure 2 B-D), with subsurface plates (Figure 2 E,I,M), or with any combination of a weir and plates (Figure 2 F-H, J-L, N-P). The tallest weir (0.16-m; Figure 2 C) had a much greater length of influence and also drove deeper hyporheic flow compared to shorter weirs (0.08-m; Figure 2 B), yet the transit time distributions were similarly short (i.e., < 400 s; indicated by red-orange flowpath color in Figure 2). Increasing weir length from 0.08 m to 0.24m (Figure 2 C vs. 2 D) increased flowpath lengths, but did not alter flowpath transit times as substantially as subsurface plates (Figure 2 E-H). In scenarios with weirs but without plates, water that entered the hyporheic zone at the upstream end of the model and persisted as underflow caused shallower, more compressed hyporheic flow cells to develop (e.g., Figure 2 B-D, where white space below the flowpaths represents down-valley underflow). By blocking all down-valley hyporheic flow from upstream, the presence of an upstream plate drove complete turnover of hyporheic water, yielding deeper hyporheic flowpaths at the structure (e.g., Figure 2 J vs. 2 B).

Exchange from submerged plates alone was deep and slow compared to weir-driven exchange (e.g., compare Figs. 2 E, I, M vs. 2 B-D). The subsurface plates also made a considerable difference to flow patterns when coupled with the weirs. Adding an upstream plate to a weir of any dimension produced deeper and longer hyporheic flowpaths than the weir alone. Many of these longer flowpaths only upwelled due to the downstream boundary condition, so they could be considered intermediate-scale flowpaths that would remain in the HZ indefinitely depending on downstream conditions. The downstream plates had a different effect, constraining the flowpaths to a greater or lesser degree depending on the location of the plates. A downstream plate at 20 cm spacing (Figure 2 I) truncated the intermediate-scale flowpaths from the scenario without a downstream plate (Figure 2 E). This effect was even more pronounced with the tighter constraint of a downstream plate at 10 cm spacing (Figure 2 M), but the flowpaths were still deeper and slower than the scenarios without plates (Figure 2 A-D). A distant or absent downstream plate (Figure 2 G) resulted in a bimodal transit time distribution between longer flowpaths associated with the upstream plate and shorter flowpaths driven by the weir. A more constraining downstream plate (Figure 2 K, O) narrowed the transit time distribution toward plug flow, especially when combined with taller weirs.

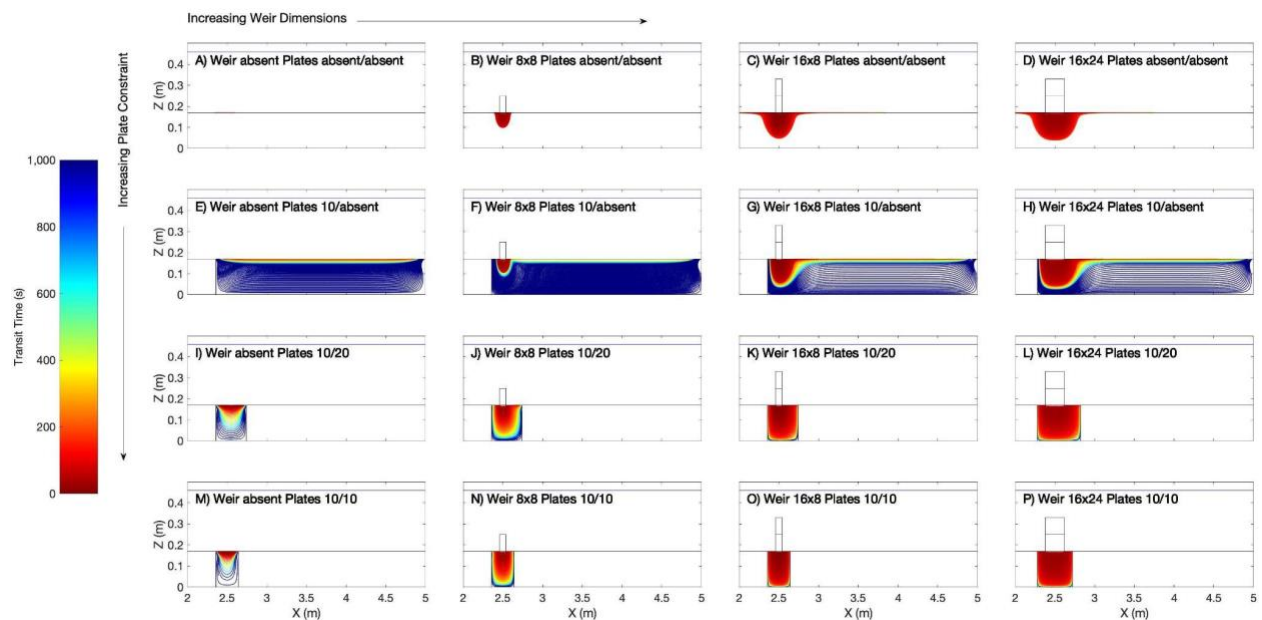


Figure 2. Simulated particle traces for select scenarios, with flowpaths color-coded according to transit time. Columns show increasing weir dimensions (weir height x length in cm), rows show increasing hyporheic plate constraint (upstream/downstream plate spacing in cm). Dark blue flowpaths indicate Damköhler numbers > 1 (i.e., transit times > 1,000 s).

In summary, surface weirs and subsurface plates both drove hyporheic exchange, but weir-driven exchange tended to be shallower and faster compared to the deeper, slower exchange driven by plates. Combinations of weirs and plates provided different hyporheic flowpath geometries and transit times than either weirs or plates in isolation.

Structure design variables control average hyporheic flowpath fluxes and transit times

Scenarios with weirs, regardless of the specific weir height or length, produced greater fluxes than scenarios without weirs (Figure 3 A,B; Figure S1). Scenarios with weirs also had longer mean transit times than scenarios without weirs, except in the case of the tallest weirs. Notably, increasing weir height increased hyporheic flux more than any other variable, but also decreased transit time. Despite this tradeoff in flux versus transit time, the net result of doubling weir height was a 6% increase in average FW-TT (see primary RSF, which only differs from FW-TT by a scalar multiple, in Figure 4 E).

In contrast to weir height, increasing weir length led to slight increases in both hyporheic flux and mean transit time (Figure 3 B). Although flux and transit time were less sensitive to weir length than to weir height, weir length created a larger increase in FW-TT by increasing both components simultaneously. For example, increasing weir length from 8-cm to 24-cm yielded a 15% increase in FW-TT. Notably, increasing weir length was also less detrimental to the secondary reaction. Whereas increasing weir height from 8-cm to 16-cm had a 33% decrease in average secondary FW-TT, increasing weir length from 8-cm to 24-cm only decreased average secondary FW-TT by 2% (see secondary RSF in Figure 4 E,F).

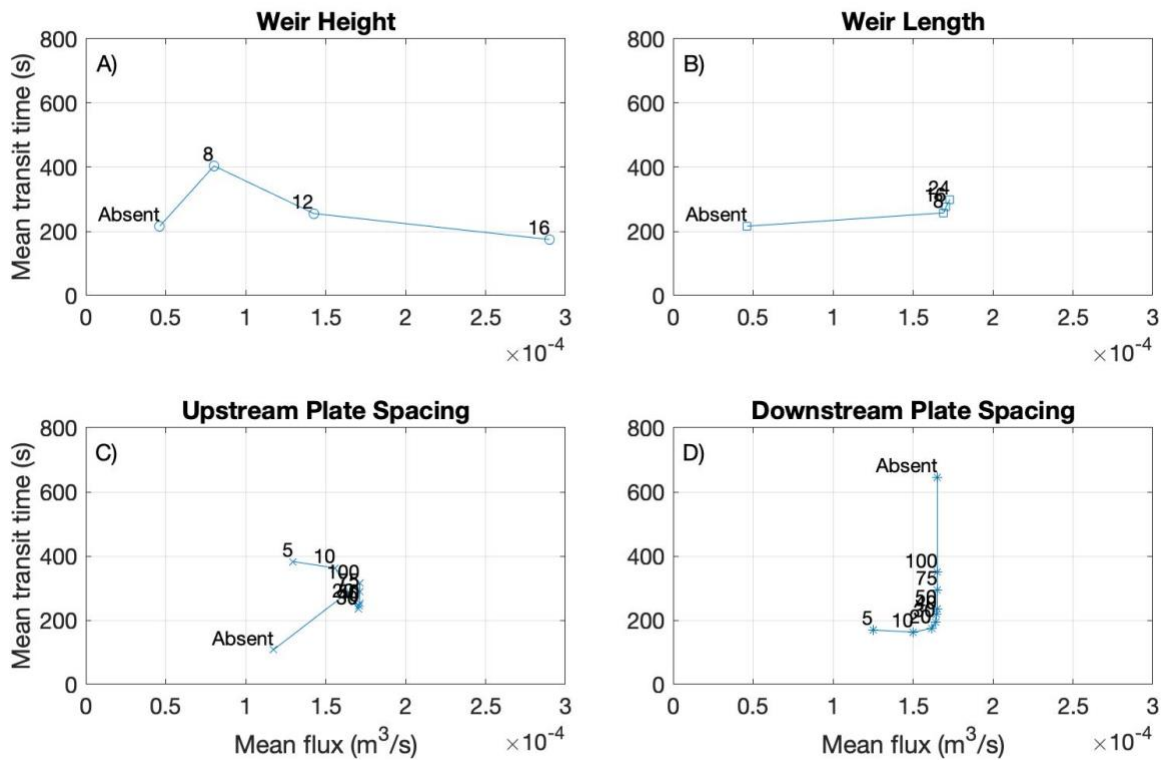


Figure 3. The effect of each design variable on total hyporheic flux (x-axis) and mean transit time (y-axis). Data labels indicate weir heights (Panel A), weir lengths (Panel B), upstream plate spacings (Panel C), and downstream plate spacings (Panel D), with all weir and plate spacing units in cm. Each data point represents the average value for the variable of interest across all combinations of other variables. Lines between data points are included to help the reader follow the sequence of data points and do not signify interpolated data points. Trends to the upper right indicate simultaneous improvements in both flux and transit time (i.e., high-flux

flowpaths with long transit times), whereas trends to the bottom left indicate decreases in both flux and transit times. Trends from upper left to bottom right, or vice-versa, indicate tradeoffs between flux and transit time.

Upstream plates improved both flux and mean transit time compared to scenarios without a plate (Figure 3 C). Notably, the results for plates 20-100 cm away from the weir were clustered together and did not trend in any consistent direction, indicating that the exact plate spacing was not sensitive as long as the plate was not too close to the weir (Figure 3 C). As the upstream plate approached the weir, the plate constrained the downwelling zone and led to lower flux but longer mean transit times. The net result of upstream plate constraint was a slightly lower FW-TT (see primary and secondary RSF in Figure 5 E,F).

Downstream plates were the only design variable that decreased both flux and mean transit time when added to a structure (Figure 3 D). Flux was relatively insensitive to the downstream plate spacing unless the plate was very close to the weir (i.e., within 20-cm in our model). In contrast, mean transit time was more sensitive to the downstream plate spacing than to any other variable. Notably, removing the downstream plate increased mean transit time without affecting flux, so FW-TT increased steadily with increasing downstream plate spacing.

Taken together, these average patterns suggest that hyporheic flux would be maximized by a taller, longer weir with a moderate-to-large upstream plate spacing and no downstream plate (e.g., Weir 16x24 Plates 100/absent). In contrast, transit times would be maximized by a very different design: a shorter height, longer length weir with a confining upstream plate and no downstream plate (e.g., Weir 8x24 Plates 5/absent). Alternatively, an illustrative balance of flux and transit time could be achieved by a medium height, longer length weir with moderate upstream plate spacing and no downstream plate (e.g., Weir 12x24 Plates 50/absent).

General trends of structure design on hyporheic contaminant attenuation performance metrics

Reaction metrics were always higher for the primary reaction than for the secondary reaction (Figures 4, 5) due to the latter's delayed onset. Primary and secondary reactions also had opposing trends with weir dimensions. For example, 16-cm weirs increased total primary mass removal by 26% (Figure 4 A) and primary RSF by 6% (Figure 4 C) compared to 8-cm weirs. Scenarios without weirs performed the worst, on average, for both metrics. Also, 24-cm long weirs consistently outperformed 8-cm weirs for total primary mass removal (+20%) and primary RSF (+15%). In contrast, primary percent mass removal and all secondary reaction metrics were highest for scenarios without weirs. Primary percent removal and all secondary metrics continued to decline as weir height increased (Figure 4 A-C). Increasing the length of weirs improved percent removal for both reactions, but not enough to match scenarios without weirs (Figure 4 D). Longer weirs also decreased secondary total mass removal (-1%) and secondary RSF (-2%), although to a much lesser extent than did weir height (-29% and -33%, respectively).

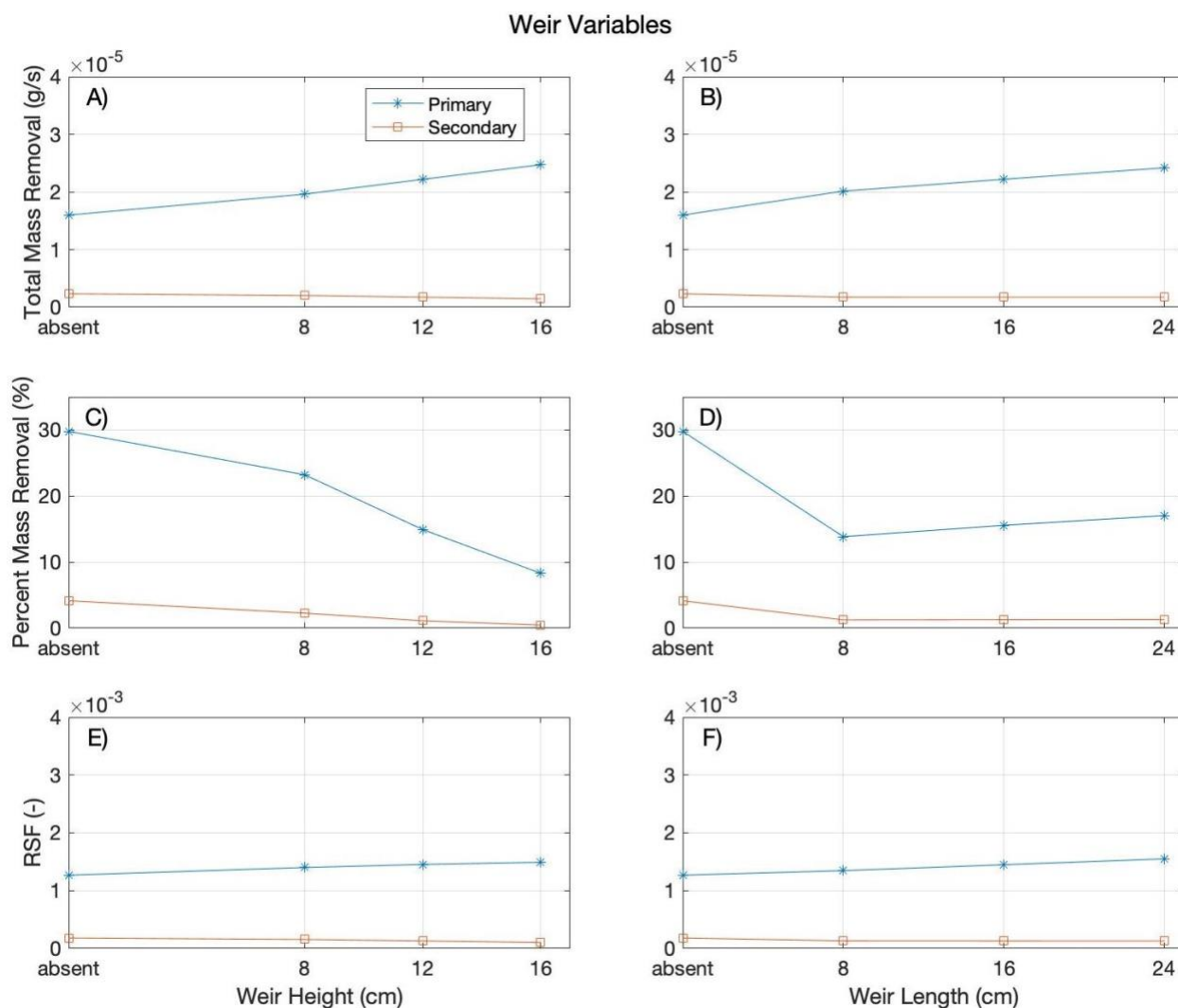


Figure 4. Average total mass removal, percent mass removal, and RSF metrics for primary and secondary reactions as a function of weir dimensions. Each data point represents the average value for the variable of interest across all combinations of other variables. For example, the average total mass removal for each weir height (e.g., data points in Figure 4 A) is a composite of all 243 weir lengths and plate spacings. Likewise, the “absent” scenario includes the 81 plate combinations without weirs. Lines between data points are included to help the reader follow the sequence of data points and do not signify interpolated data points. Y-axes are the same as Figure 5 to facilitate comparison between weir and plate variables.

The presence of an upstream plate also always improved structure performance, and the trends were notably similar for all metrics and reactions (Figure 5 A-C). Although the optimal spacing was the furthest distance considered in this study (100-cm), intermediate locations were still relatively effective. For example, reducing the upstream plate spacing 50% from 100-cm to 50-cm only reduced total and percent primary mass removal by 20%, primary RSF by 30%, and all metrics for the secondary reaction by approximately 40%.

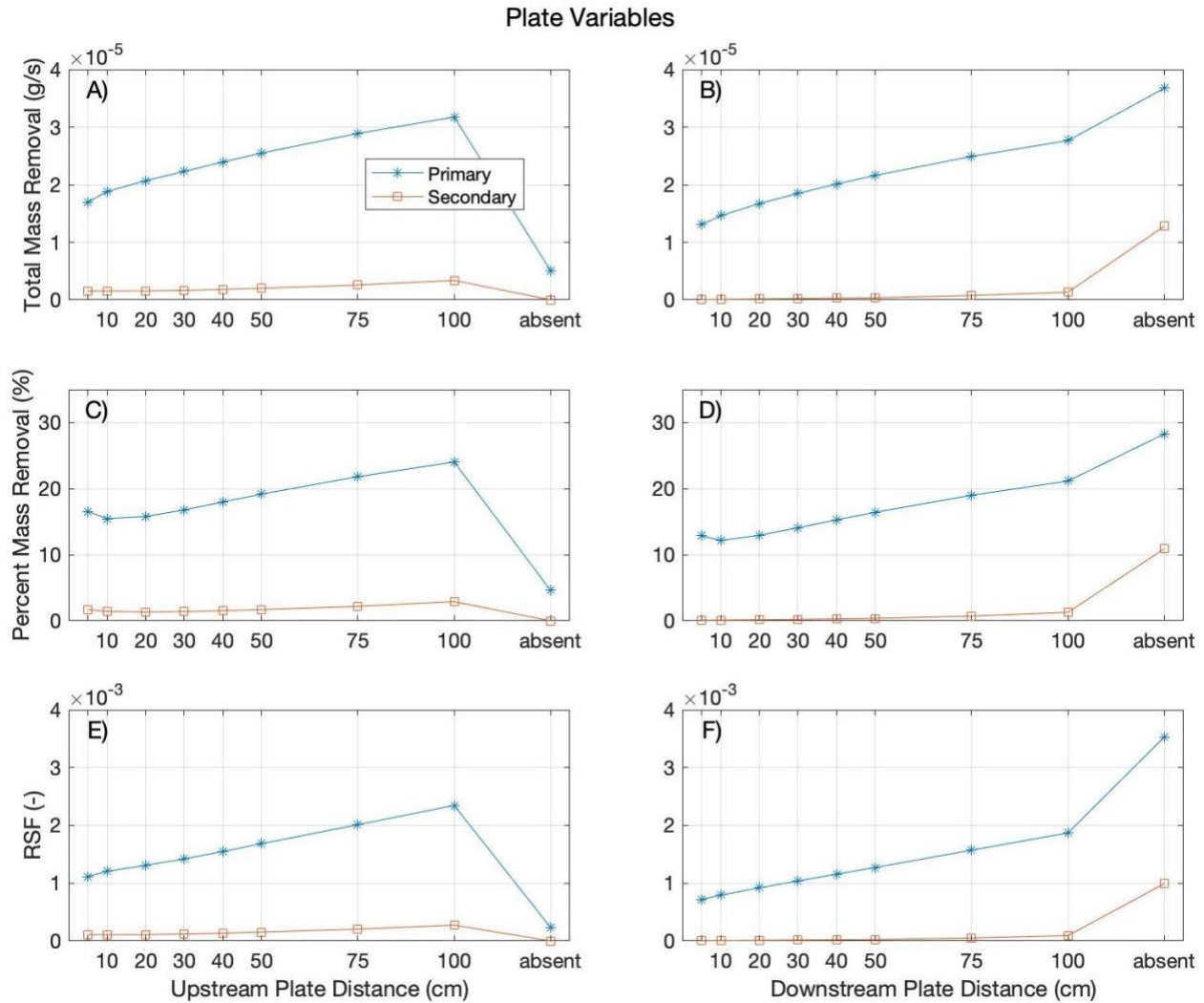


Figure 5. Average total mass removal, percent mass removal, and RSF metrics for primary and secondary reactions as a function of plate spacings. For example, the average total mass removal for each upstream plate spacing average (e.g., data points in Figure 5A) is a composite of all weir heights, weir lengths, and downstream plate spacings, including the control condition without weirs or plates. Likewise, the “absent” scenario includes both the control condition and all combinations of weirs and downstream plates. Lines between data points are included to help the reader follow the sequence of data points and do not signify interpolated data points. Y-axes are the same as Figure 4 to facilitate comparison between weir and plate variables.

The downstream plate also produced consistent trends across all metrics and reactions, but in this case performance was always maximized by the absence of a downstream plate (Figure 5 B,D). A moderately spaced downstream plate performed reasonably well for the primary reactions. For example, a downstream plate at 50-cm spacing reduced primary total and percent mass removal by approximately 40%, and primary RSF by 64% compared to scenarios with no downstream plate. The closer a downstream plate came to a weir, the greater the constraint and the worse the performance across all metrics considered, especially for the secondary reaction. Indeed, a downstream plate at 50-cm spacing reduced all secondary

reaction metrics by 97-98%. On average, performance metrics were more sensitive to plate spacing than to weir dimensions (Figures 4, 5).

In summary, the optimal weir height depends on the reaction and metric of interest, while weir length and plate spacings do not. Scenarios without a weir performed better than those with a weir for primary percent removal and all secondary reactions. However, longer weirs performed better than shorter weirs for primary reaction metrics and approximately equal to shorter weirs for secondary reaction metrics. Further, moderately-to-longer spaced upstream plates and the absence of downstream plates increase all metrics, particularly for the secondary reaction. Specifically, total mass removal and RSF for the primary reaction would be maximized by Weir 16x24 Plates 100/absent, whereas all other reactions would select for Weir absent Plates 100/absent.

Specific design optimization can greatly improve hyporheic contaminant attenuation

Specific combinations of weir and plate dimensions had a substantial impact on performance metrics, and the best performing structures were well predicted by the general patterns. The only substantial deviation from general trends was that longer weir lengths often outperformed shorter weir lengths, even for percent removal and secondary reactions where average trends showed a slight negative correlation. The general patterns also concealed the tremendous range in individual structure performance, especially for the secondary reaction. For example, 25.3% of structures did not exceed the minimum transit time requirement of the secondary reaction and thus had no secondary removal at all.

The structure with the highest total flux (Weir 16x24 Plates 100/absent) also had the greatest total primary total mass removal and primary RSF. Notably, this structure provided a 38% improvement over any plate combination without a weir (i.e., subsurface structure only), a 318% increase in primary mass removal over any weir without plates (i.e., surface structure only), and more than a 5-fold increase compared to any of the shortest length weirs without plates (i.e., the standard practice of only optimizing structures according to weir height, as in ^{12,14}). Primary RSF was especially sensitive to plate spacing, and was respectively 3%, 750%, and 1,196% higher for Weir 16x24 Plates 100/absent than for subsurface structures only, surface structures only, and the shortest length weirs without plates.

The structure with the highest mean transit time (Weir 8x16 Plates 5/absent) did exceed 73% of the maximum value for any metric. Instead, percent removal and all secondary reactions were maximized by a very different set of structures (Weir absent Plates 5-100/absent). This set of structures clustered together, with slight differences due to an artifact of the model setup. Specifically, particles were only released in the first 2.5m of the model (i.e., upstream from the weir), so scenarios with upstream plates but no weirs had their downwelling zones artificially truncated at $x = 2.5\text{m}$. An upstream plate without a weir creates a downwelling zone approximately 30cm in length (Figure S1 A), so this artifact only impacted a few scenarios (Weir absent Plates 5-30/absent), which hereafter are treated as equivalent to the comparable non-impacted structure (i.e., Weir absent Plates 100/absent).

Weir absent Plates 100/absent had 63% removal for the primary reaction and 34% removal for secondary reaction. Despite not having any surface element, this structure was 35% more effective for primary percent mass removal than was the best combined surface-subsurface structure (Weir 8x24 Plates 5/absent) and 564% more effective than the best performing structure without plates (Weir 8x24 Plates absent/absent). For the secondary percent mass removal, Weir absent Plates 100/absent was 56% more effective than the best combined surface-subsurface structure (Weir 8x24 Plates 100/absent). All scenarios without plates did not remove any secondary mass. In contrast, Weir absent Plates 100/absent only outperformed the best combined surface-subsurface structure (Weir 8x24 Plates 100/absent) by 3% and 10% for each metric, respectively.

Multiple combinations, including the top performing structures above, provided relatively high performance across all metrics. For example, Weir absent Plates 100/absent maximized primary percent removal and all secondary metrics, while also producing 97% of the maximum RSF and 72% of the maximum total mass removal for the primary reaction. Even the largest weirs generated substantial secondary mass removal when paired with Plates 100/absent. For example, Weir 16x24 Plates 100/absent still generated 83% of the maximum secondary mass removal and 70% of the maximum observed secondary RSF. Indeed, the combination of long weirs and Plates 100/absent generated relatively high values for most metrics. Weirs 8x24, 12x24, and 16x24 with Plates 100/absent were within 64-97%, 34-100%, and 15-100%, respectively, of the maximum observed value for all six metrics. Thus, design optimization for a single metric does not need to come at the expense of all other metrics.

Discussion

How can the known controls on hyporheic exchange be leveraged to maximize flux at the timescales required for a given process or reaction of interest?

Adding a weir improved both flux and transit time compared to scenarios without weirs, regardless of the specific weir dimensions (Fig. 3 A-B). Adding an upstream plate likewise improved both flux and transit time across all plate spacings considered in this study (Fig. 3 C). For both weirs and upstream plates, the specific parameter values could be modified to skew hyporheic conditions toward greater flux, greater transit time, or a balance of both. Adding a downstream plate, however, decreased both flux and transit time compared to no downstream plate (Fig. 3 D). Thus, weir dimensions and upstream plates provided useful design decisions to increase fluxes of specific transit times, whereas downstream plates did not.

Surface weirs were generally best suited to drive high exchange fluxes of relatively short transit times, whereas subsurface plates were better at creating moderate fluxes of longer transit times (Figures 2, S1). The combination and optimization of surface-subsurface structures produced a broad spectrum of fluxes, flowpath lengths, and transit times, which overlapped with a range of natural hyporheic features (e.g., Figure 6 in ⁵). The design levers that maximize

flux are increasing weir height, increasing weir length, adding an upstream plate of increasing spacing, and removing a downstream plate (Table 3). These general design trends are also predictive of specific design optimizations, as the structure with the greatest flux was Weir 16x24 Plates 100/absent.

In contrast, transit times can be maximized by reducing weir height, adding an upstream plate of decreasing spacing, and removing a downstream plate. General trends suggested that longer weir lengths increased transit times, but that was not the case for many individual structures depending on specific plate-weir interactions. The specific structure with the greatest mean transit time was Weir 8x16 Plates 5/absent, followed closely by Weir 8x24 Plates 5/absent, followed by other combinations of the shortest weir height and Plates 5/absent and 10/absent.

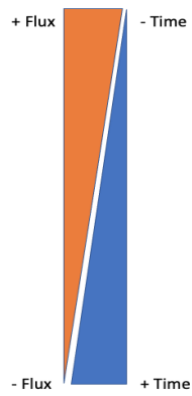
Maximizing either flux or transit time in isolation tended to reduce the other, but many design combinations created a balance of both outcomes. For example, the maximum flux structure (Weir 16x24 Plates 100/absent) was in the 75th percentile for transit time, and the maximum transit time structure (Weir 8x16 Plates 5/absent) was in the 19th percentile for flux. Another structure (Weir 16x24 Plates 20/absent) was in the 93rd percentile for flux and 88th for transit time. The optimization challenge may be thought of as maximizing the flux associated with the relevant timescales for a desired process or reaction. Surface-subsurface structure combinations can align fluxes and transit times, although the longer the target transit time, the lower the flux that can generally be produced.

Two design levers increased flux and transit times simultaneously: increasing weir length and adding an upstream plate. Longer length weirs actually caused slight declines in downwelling flux at the weir (Figure S1) due to the reduced $\Delta h/\Delta x$ gradient (i.e., the same Δh across the weir but longer Δx). However, decreased weir-driven flux reduced interference between the weir and upstream plate, leading to greater overall flux in the combined surface-subsurface structure. Weir-plate interference was greatest at the tallest weir heights and shortest upstream plate spacings, so longer weirs were especially productive in these scenarios. Further, longer weirs increased transit times principally by lengthening flowpaths rather than by reducing porewater velocities (and therefore flux). Lengthening weirs is consistent with other designs that extended hyporheic flowpath lengths to increase transit times, especially of the shortest, highest flux flowpaths at the center of a structure (e.g. ^{22,23}).

Upstream plates also increased flowpath lengths by causing complete exchange of hyporheic water from the upstream boundary condition, which drove deeper and longer downwelling flowpaths under the structure (e.g., Fig. 2 D vs. 2 B). Upstream plates were particularly important because they caused the coincidence of moderately high fluxes and longer transit time flowpaths, whereas high flux flowpaths near weirs typically had shorter lengths and transit times (Figure S1). We also note that increasing the plate spacing beyond the range studied here would eventually produce a separate hyporheic flow cell that could interfere with the weir-driven hyporheic flow cell. We refer to a “moderately spaced” upstream plate as the optimal spacing to reflect this limitation.

Weir height produced the clearest tradeoff between flux and transit time. Weir height can therefore be reduced when greater transit times may be required. However, such an increase in transit time comes from reduced porewater velocities rather than an increase in flowpath length. Thus, less flux is possible at the longer transit time when weir height is used as a design lever. As increasingly long transit times may be required (e.g., for a secondary reaction), the optimal structure approaches the maximum transit time structure with relatively low flux.

Table 3. Design decisions for single structures to maximize flux (Q) of various transit times (TT), or to maximize transit times.



	Weir Height	Weir Length	Upstream Plate	Downstream Plate
High Q Low TT	Taller	Shorter	Moderately far	Moderately far
High Q Moderate TT	Taller	Longer	Moderately far	Absent
Moderate Q High TT	Moderate	Longer	Moderately far	Absent
Low Q High TT	Shorter or absent	Longer or absent	Relatively close	Absent

How does the selection of a performance metric change design decisions?

To analyze the impact of each performance metric on design optimization, it is important to first discuss the sensitivities of each metric. Total mass removal is most sensitive to flux due to the nonlinearity of first-order decay, which assumes that short flowpaths are more efficient than longer flowpaths at removing pollutants. Using total mass removal as a performance metric favors high-flux, short transit time structures. In other words, designs that trade higher fluxes for shorter transit times would increase total mass removal (assuming first order decay), although designs that increase both flux and transit time are clearly better. Total mass removal for the secondary reaction also favors higher fluxes over longer transit times, as long as the transit times exceed the delayed onset threshold.

Percent mass removal and RSF present interesting contrasts to total mass removal and to each other. Neither percent mass removal nor RSF reward flux-for-transit time tradeoffs, but for different reasons. Percent mass removal does not incorporate any flux weighting and was exclusively sensitive to transit time. RSF is flux-weighted, but does not incorporate the nonlinearity of first-order decay. Instead, RSF assigns equal weight to each second of the entire flowpath transit time and is thus sensitive to the full spectrum of transit times. The present study featured relatively short transit timescales and reaction-limited flowpaths (e.g., percent removal < 100%), especially for the secondary reaction. However, transport-limited flowpaths

could readily be generated in other systems with faster reaction rates, larger structures, or lower hydraulic conductivities. Care should be taken with RSF to ensure that optimization balances fluxes and transit times, rather than biasing designs toward extremely long, transport-limited flowpaths that do not actually sustain the reaction of interest.

Importantly, the model results presented here provide illustrative trends that are relevant to a variety of hyporheic structures. Adjustment of homogeneous model parameters (e.g., hydraulic conductivity, reaction rates) would shift the absolute values of our results, but the underlying patterns would remain. Spatial heterogeneities could easily shift both the patterns and absolute values of our results, but there are insufficient data in the literature to parameterize spatial heterogeneity in a meaningful way. One possible nuance we did not study is the case where the initial, shallow HZ sediments have higher reaction rates than deeper sediments (e.g., ^{36,37}). If this were the case in our system, that would skew all metrics toward higher flux, shorter transit time flowpaths associated with weirs. Alternatively, reaction rates may be faster along flowpaths with low porewater velocity compared to high porewater velocity flowpaths, due to greater contact time and, possibly, more favorable reaction conditions ^{8,38}. If that were the case in our system, it would skew results toward deeper, slower flowpaths originating near upstream plates. Overall, our approach demonstrates that the hyporheic hydraulics of combined surface-subsurface structures can be evaluated using a forward modeling approach to inform design, which could readily be extended to other permeability or reactivity fields, structure geometries, boundary conditions, design objectives, or performance metrics.

How should these results be applied to stream restoration projects?

In this study we assumed homogeneous, isotropic conditions, no streambed slope, constant depth to an impermeable layer, no ambient groundwater interactions, and no neighboring or multi-scale features. Such assumptions may be valid in engineered channels, but are not necessarily appropriate for natural channels and can have dramatic impacts on hyporheic exchange around in-stream structures¹³. For example, restoration structures placed in a pre-existing hyporheic exchange cell may actually decrease fluxes and mass removal relative to pre-restoration conditions^{12,39}. Likewise, lateral or vertical upwelling from regional aquifers, meanders, or floodplain ponds and wetlands can seasonally shift hyporheic exchange patterns²² or reduce their function⁴⁰. Such construction siting, land use, and hydrogeologic factors must be considered in any restoration project^{41–43}, but are beyond the scope of this numerical study.

The short, rapid flowpaths at the center of a hyporheic flow cell tend to be more stable across time than the long, slow flowpaths at the fringe of a hyporheic cell^{13,32,44}. For example, flowpaths near the center of a structure (e.g., “patch-scale” exchange¹³) are relatively resistant to forcings not considered in this study, such as annual changes in stream discharge and focused groundwater upwelling³¹, changes in surrounding streambed topography³², or sediment heterogeneities⁴⁵. Total mass removal mostly depends on these high-flux, relatively short timescale flowpaths and thus is expected to be relatively consistent through time. In contrast, RSF and percent mass removal are more dependent on the longest flowpaths at the

fringe of a hyporheic flow cell. These distal flowpaths are more weakly associated with each hyporheic structure and are more easily converted from local- to intermediate-scale flowpaths by the addition of plates. Prior studies have shown that distal flowpaths at the edges of a structure are also more sensitive to changes in neighboring features or groundwater upwelling^{31,32}. Thus, the presence and length of intermediate-scale flowpaths depend on structure design and the surrounding hydraulic context. In this study, intermediate-scale flowpaths were up to 4m long (i.e., to the downstream boundary condition), which was reasonable given the common 2-5 channel width (i.e., 2-5 m in the modeled flume) spacing of geomorphic features in natural streams^{13,46}. Subsurface plates, in particular, can both generate intermediate-scale flowpaths and shield them from the surrounding hydraulic context. In heavily engineered channels, plates (or other impermeable material such as logs, boulders, or compacted clay) may even be extended around and below the structure to shield engineered hyporheic flowpaths from lateral and vertical interference, thereby promoting consistent intermediate-scale flows.

The design parameters in this study were selected to demonstrate transferable concepts that are relevant to multiple stream restoration structure types and sizes. Indeed, specific design decisions (e.g., hydraulic head gradient, weir/cap length) may be as important as, or more important than, what kind of structure is built (e.g., weir, cross-vane, plunge pool). For example, a stream restoration modeling study in Sweden found that adding weirs (relatively tall height but short length) would enhance primary reactions but decrease the secondary reaction of interest (i.e., denitrification)¹². To improve denitrification, the authors raise the possibility of removing weirs and reducing streambed roughness, thereby decreasing fluxes to increase transit times¹². Importantly, our results show that longer weir lengths, especially in combination with subsurface plates, could improve both primary and secondary reactions simultaneously. In other words, these techniques use longer flowpaths (e.g. via upstream plate and longer weirs) rather than reduced flux (e.g., by reducing weir heights) to increase transit time.

Our results could also be applied to larger-scale engineered hyporheic plunge pool structures (e.g., ^{22,23}), where the weir height would be replaced by the waterfall height across the structure as the main source of hydraulic pressure gradient, and hyporheic flowpath length could be optimized by lengthening the impermeable clay cap along the sediment-water interface (e.g., from 2.5m to 5m). Subsurface plates could be installed to bracket the structure, or the same effect may be achieved by minimizing the hyporheic zone depth at discrete points upstream and downstream from the structure (e.g., by not “over-excavating” brackets around the structure, thereby leaving “plates” of the relatively impermeable native material intact). Other engineered hyporheic zone studies have also shown that hyporheic flux can be increased without reducing transit times by expanding the hyporheic cross-sectional area^{22,23}, and by increasing both the hydraulic conductivity and porosity (in proportion, so that higher porosity mitigates any increase in porewater velocity) of sediment materials². Taken together, our results and the existing literature point to the potential of high flux, long transit time hyporheic flowpaths in various hyporheic structures with some combination of upstream buried plates, longer length weirs (or caps), and deeper or wider streambeds of more permeable, porous material.

Full characterization of hyporheic flux and transit time distributions may not be possible in all field studies. Each metric in our study was calculated from hundreds of particle traces, but we also compared those results to simpler proxies: each structure's total flux, average transit time, and their product (average flux-weighted transit time). Total mass removal for the primary and secondary reactions, and primary RSF, were most strongly correlated with average FW-TT ($R^2 = 0.64, 0.44, \text{ and } 0.63$, respectively; Table S1). Primary and secondary percent mass removal, and secondary RSF, were most correlated with average transit time $R^2 = 0.48, 0.46, \text{ and } 0.40$, respectively; Table S1). Notably, total flux alone did not provide the best correlation with any metric, likely due to the predominantly reaction-limited conditions in the study. Due to the relatively small structures and high permeability sediments, less than 2% of scenarios had an average Damköhler number greater than 1, and the maximum average Damköhler number for any structure was 1.73.

What is the optimal design of surface-subsurface structures for hyporheic contaminant attenuation?

Three design features consistently improved performance across all metrics. First, the presence of an upstream plate was critical to increase fluxes while also increasing transit times via deeper, longer flowpaths. Scenarios without upstream plates always performed worse than structures with plates (Figures 3-5). However, the benefits of a plate were reduced by a very close upstream spacing (i.e., 5 cm spacing), especially for taller weirs that otherwise would have driven more extensive downwelling.

Second, longer length weirs performed better than shorter length weirs in almost all cases. Lengthening a weir slightly decreased fluxes at the weir by decreasing the average hydraulic gradient, yet the overall structure flux increased due to reduced weir-plate interference. As plate-driven fluxes tend to have longer transit times than weir-driven fluxes, shifting even slightly from weir-driven toward plate-driven hyporheic exchange also increases transit times and FW-TT. Our results agree with past work¹³ showing that weir height was far more important than flowpath length in determining hyporheic velocity and flux (in ¹³, flowpath length was controlled by hydrogeologic setting rather than weir length). The main mechanism of increasing transit times in longer weirs was thus to lengthen flowpath geometries rather than reducing velocities and fluxes, which was another reason that increasing weir length benefitted most metrics.

Third, increasing downstream plate spacing also increased performance by all metrics. Downstream plates tended to constrain hyporheic fluxes, geometries, and transit times by compressing the hyporheic capacity through a structure. The greater the compression, the lower the performance (e.g., Figure 5), analogous to, and consistent with, the reduced hyporheic function associated with the constraint of shallow bedrock (e.g., ¹³). Our results showed substantial reductions in all performance metrics for downstream plates at even moderate spacings (e.g., 50-100 cm). However, a downstream plate would be less impactful in

systems with lower hydraulic conductivity, faster reaction rates, or larger structure scales. Downstream plates may be inevitable for multiple structures in series, as one structure's upstream plate would serve as the downstream plate of the preceding structure. A moderately distant downstream plate spacing (e.g., 50- to 100-cm) would allow many structures to be placed close together (i.e., greater density per stream length) and can also make flowpath TTDs more uniform (i.e., approaching plug flow), which may be useful for targeting a narrow transit time band or to simplify the task of quantifying hyporheic contaminant attenuation (e.g., all downwelling and upwelling locations in a structure have similar contaminant concentrations).

Weir height was the only design decision that would change based on the metric of interest. Our simulations showed that increasing weir height was essential to maximize fluxes and total mass removal. Primary RSF also increased with weir height. In contrast, decreasing weir heights and removing weirs altogether improved primary percent mass removal and all secondary reaction metrics. In other words, the optimal weir height moved in opposite directions depending on primary vs secondary reactions and between total mass removal and percent mass removal. Several structures (e.g., Weir absent Plates 100/absent and Weir 8-16/24 Plates 100/absent) performed well for all metrics, but clear tradeoffs still existed, especially between total and percent mass removal.

It is rare to see specific water quality goals for stream restoration that define priority contaminants and evaluation metrics. Water quality is often mentioned in discussions with stakeholders and restoration professionals, but typically in generic terms and as a secondary project goal or ancillary benefit of stream restoration⁴⁷. Fortunately, three out of four design variables have clear and consistent guidance regardless of reaction types or performance metrics. However, in the case of weir height where different metrics suggest opposing designs, the question must be asked: "What is the goal of this structure?" Practitioners will be challenged to decide between multiple positive but mutually-exclusive alternatives, such as optimizing for high flux flowpaths for pollutant load reductions, versus lower flux flowpaths for habitat refugia in well-filtered hyporheic upwelling zones, or trying to balance multiple, competing benefits. Such choices must be made, because ubiquitous degradation throughout a watershed precludes a straightforward restoration to pre-human influence⁴⁸.

To treat orders of magnitude higher pollutant loads with piecemeal restoration projects, restoration structures may need to be more efficient and more densely packed than their natural analogs. Reach- and network-scale analyses have shown that many kilometers of restoration may be needed to substantially improve water quality, especially for secondary reactions like denitrification^{10,49,50}. However, the modeled structures in the above studies, like the actual structures they represented, did not simultaneously optimize fluxes and transit times. Design principles from the present study could dramatically increase the reach-scale water quality benefits of each structure while also increasing the number of restoration structures per meter of stream restoration. In our simulations, the most effective structure for primary reactions was 6-fold more effective than a weir optimized only by height (i.e., standard practice, as in ^{12,14}). The secondary reaction was even more sensitive to design, as weirs without plates did not produce any secondary mass removal. Of course, secondary removal could be

nonzero in the same structures if the reaction rates were faster or hydraulic conductivity were lower, but many surface hyporheic structures are unable to generate the necessary conditions for secondary reactions like denitrification (e.g., ^{10,11,51}). However, our results show that otherwise ineffective surface structures can be coupled with subsurface structures to stimulate secondary reactions. More efficient structures placed more densely in the streambed could improve water quality in hundreds of meters rather than kilometers. Of course, the most efficient design depends on individual project goals and site-specific factors. Here we present some considerations for which metrics align with various applications, along with the relevant design variables (Table 4).

Table 4. Translating hyporheic water quality goals into appropriate metrics and designs.

Goal	Mass attenuation - primary reaction	Mass attenuation - secondary reaction	Percent removal - primary or secondary reaction	Individual flowpath monitoring
Example Scenario	Total maximum daily load	Total maximum daily load	Low pollutant concentrations in upwelling zone	Quantifying hyporheic contaminant attenuation
Example pollutants	Ammonia, Metformin	Nitrate, Guanyldurea	Any	Any
Direct Metric	Total mass removal [M/T]	Total mass removal [M/T]	Percent mass removal	Any
Proxy Metric	Average flux-weighted transit time	Average flux-weighted transit time	Average transit time	Any
Weir Dimensions	Taller height, longer length	Moderate height, longer length	Shorter height, longer length	Any height, longer length
Upstream Plate	Moderately close to maximize structure density	Moderately close to maximize structure density	Relatively close to reduce porewater velocities	Moderately close to isolate feature-scale flowpaths
Downstream Plate	Moderately close to maximize structure density	Moderately close to maximize structure density	Distant or absent to maximize flowpath length	Moderately close to promote plug flow

Conclusions

1. Optimized surface-subsurface structures can integrate the relatively high fluxes from surface structures with the relatively long transit times of subsurface structures.
2. When used to design hyporheic structures, performance metrics based on total mass removal select for higher hyporheic fluxes, whereas metrics based on reaction completeness (i.e., percent removal and RSF) select for longer transit times.

3. The optimal weir height depends entirely on the choice of performance metric. In contrast, all metrics were improved by longer weir lengths, the presence of an upstream plate, and the absence of a downstream plate. These three factors improved all metrics by generating high-flux flowpaths with relatively long transit times.
4. Combined surface and subsurface structures that are optimized by weir height, weir length, and plate spacing can deliver many times the water quality performance of surface or subsurface structures in isolation.
5. The selection of a performance metric should align with specific restoration goals. The chosen metric should be used to optimize structure designs and monitoring efforts.

Corresponding Author

*Skuyler P. Herzog

Email: herzogs@oregonstate.edu

Present Addresses

†If an author's address is different than the one given in the affiliation line, this information may be included here.

Jason Galloway - Agricultural Catchments Programme, Department of Environment, Soils and Landuse, Teagasc, Johnstown Castle Environment Research Centre, Wexford, Co. Wexford, Ireland

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. ‡These authors contributed equally. (match statement to author names with a symbol)

Funding Sources

Any funds used to support the research of the manuscript should be placed here (per journal style).

Ward would like to add some here when the time comes.

Acknowledgements

We thank the German Federal Environment Agency Umweltbundesamt for their encouragement and support with this project, including access to the flumes in the artificial stream and pond system.

Abbreviations

FW-TT, average Flux-Weighted Transit Time; FW-TTD, Flux-Weighted Transit Time Distribution; HZ, Hyporheic Zone; RSF, Reaction Significance Factor; TMDL, Total Maximum Daily Load; TT, Transit Time.

References

- (1) Vaux, W. G. Intragravel Flow and Interchange of Water in a Streambed. *Fishery Bulletin* **1968**, 66 (3), 479–489.
- (2) Herzog, S. P. P.; Higgins, C. P. P.; McCray, J. E. E. Engineered Streambeds for Induced Hyporheic Flow: Enhanced Removal of Nutrients, Pathogens, and Metals from Urban Streams. *Journal of Environmental Engineering* **2016**, 142 (1), 04015053. [https://doi.org/10.1061/\(ASCE\)EE.1943-7870.0001012](https://doi.org/10.1061/(ASCE)EE.1943-7870.0001012).
- (3) Hester, E. T.; Gooseff, M. N. Moving beyond the Banks: Hyporheic Restoration Is Fundamental to Restoring Ecological Services and Functions of Streams. *Environmental Science and Technology* **2010**, 44 (5), 1521–1525. <https://doi.org/10.1021/es902988n>.
- (4) Lewandowski, J.; Arnon, S.; Banks, E.; Batelaan, O.; Betterle, A.; Broecker, T.; Coll, C.; Drummond, J. D. J. D.; Garcia, J. G. J. G.; Galloway, J.; Gomez-Velez, J.; Grabowski, R. C. R. C.; Herzog, S. P. S. P.; Hinkelmann, R.; Höhne, A.; Hollender, J.; Horn, M. A. M. A.; Jaeger, A.; Krause, S.; Prats, A. L. A. L.; Magliozzi, C.; Meinikmann, K.; Mojarrad, B. B.; Mueller, B. M. B. M.; Peralta-Maraver, I.; Popp, A. L. A. L.; Posselt, M.; Putschew, A.; Radke, M.; Raza, M.; Riml, J.; Robertson, A.; Rutere, C.; Schaper, J. L. J. L.; Schirmer, M.; Schulz, H.; Shanafield, M.; Singh, T.; Ward, A. S. A. S.; Wolke, P.; Wörman, A.; Wu, L. Is the Hyporheic Zone Relevant beyond the Scientific Community? *Water (Switzerland)* **2019**, 11 (11). <https://doi.org/10.3390/w11112230>.
- (5) Ward, A. S.; Gooseff, M. N.; Johnson, P. A. How Can Subsurface Modifications to Hydraulic Conductivity Be Designed as Stream Restoration Structures? Analysis of Vaux's Conceptual Models to Enhance Hyporheic Exchange. *Water Resources Research* **2011**, 47 (8). <https://doi.org/10.1029/2010WR010028>.
- (6) Valett, H. M.; Morrice, J. A.; Dahm, C. N.; Campana, M. E. Parent Lithology, Surface-Groundwater Exchange, and Nitrate Retention in Headwater Streams. *Limnology and Oceanography* **1996**, 41 (2), 333–345. <https://doi.org/10.4319/lo.1996.41.2.0333>.
- (7) Gu, C.; Hornberger, G. M.; Herman, J. S.; Mills, A. L. Effect of Freshets on the Flux of Groundwater Nitrate through Streambed Sediments. *Water Resources Research* **2008**. <https://doi.org/10.1029/2007WR006488>.
- (8) Harvey, J. W.; Böhlke, J. K.; Voytek, M. A.; Scott, D.; Tobias, C. R. Hyporheic Zone Denitrification: Controls on Effective Reaction Depth and Contribution to Whole-Stream Mass Balance. *Water Resources Research* **2013**, 49 (10), 6298–6316. <https://doi.org/10.1002/wrcr.20492>.
- (9) Harvey, J. W.; Fuller, C. C. Effect of Enhanced Manganese Oxidation in the Hyporheic Zone on Basin-Scale Geochemical Mass Balance. *Water Resources Research* **1998**, 34 (4), 623–636. <https://doi.org/10.1029/97WR03606>.
- (10) Azinheira, D. L.; Scott, D. T.; Hession, W.; Hester, E. T. Comparison of Effects of Inset Floodplains and Hyporheic Exchange Induced by In-Stream Structures on Solute Retention. *Water Resources Research* **2014**, 50 (7), 2108–2123. <https://doi.org/10.1002/2012WR013085>.Received.
- (11) Clary, J.; Jones, J.; Leisenring, M.; Strecker, E.; Bledsoe, B.; Lammers, R. *Stream Restoration BMP Database: Version 1.0 Summary Report*; Water Environment & Reuse Foundation.
- (12) Morén, I.; Wörman, A.; Riml, J. Design of Remediation Actions for Nutrient Mitigation in

- the Hyporheic Zone. *Water Resources Research* **2017**, 53 (11), 8872–8899.
<https://doi.org/10.1002/2016WR020127>.
- (13) Hester, E. T.; Doyle, M. W. In-Stream Geomorphic Structures as Drivers of Hyporheic Exchange. *Water Resources Research* **2008**, 44 (3).
<https://doi.org/10.1029/2006WR005810>.
- (14) Liu, S.; Chui, T. F. M. Optimal In-Stream Structure Design through Considering Nitrogen Removal in Hyporheic Zone. *Water* **2020**, 12 (5), 1399.
<https://doi.org/10.3390/w12051399>.
- (15) Sawyer, A. H.; Bayani Cardenas, M.; Buttles, J. Hyporheic Exchange Due to Channel-Spanning Logs. *Water Resources Research* **2011**, 47 (8).
<https://doi.org/10.1029/2011WR010484>.
- (16) Elliott, A. H.; Brooks, N. H. Transfer of Nonsorbing Solutes to a Streambed with Bed Forms: Theory. *Water Resources Research* **1997**, 33 (1), 123–136.
<https://doi.org/10.1029/96WR02784>.
- (17) Marzadri, A.; Tonina, D.; Bellin, A.; Vignoli, G.; Tubino, M. Semianalytical Analysis of Hyporheic Flow Induced by Alternate Bars. *Water Resources Research* **2010**, 46 (7).
<https://doi.org/10.1029/2009WR008285>.
- (18) Tonina, D.; Buffington, J. M. Effects of Stream Discharge, Alluvial Depth and Bar Amplitude on Hyporheic Flow in Pool-Riffle Channels. *Water Resources Research* **2011**, 47 (8). <https://doi.org/10.1029/2010WR009140>.
- (19) Ward, A. S.; Morgan, J. A.; White, J. R.; Royer, T. V. Streambed Restoration to Remove Fine Sediment Alters Reach-Scale Transient Storage in a Low-Gradient Fifth-Order River, Indiana, USA. *Hydrological Processes* **2018**, 32 (12), 1786–1800.
<https://doi.org/10.1002/hyp.11518>.
- (20) Zhou, T.; Endreny, T. A. Reshaping of the Hyporheic Zone beneath River Restoration Structures: Flume and Hydrodynamic Experiments. *Water Resources Research* **2013**, 49 (8), 5009–5020. <https://doi.org/10.1002/wrcr.20384>.
- (21) Crispell, J. K.; Endreny, T. A. Hyporheic Exchange Flow around Constructed In-Channel Structures and Implications for Restoration Design. *Hydrological Processes* **2009**, 23 (8), 1158–1168. <https://doi.org/10.1002/hyp.7230>.
- (22) Peter, K. T.; Herzog, S.; Tian, Z.; Wu, C.; McCray, J. E. J. E.; Lynch, K.; Kolodziej, E. P. E. P. Evaluating Emerging Organic Contaminant Removal in an Engineered Hyporheic Zone Using High Resolution Mass Spectrometry. *Water Research* **2019**, 150, 140–152.
<https://doi.org/10.1016/J.WATRES.2018.11.050>.
- (23) Bakke; Hrachovec; Lynch. Hyporheic Process Restoration: Design and Performance of an Engineered Streambed. *Water* **2020**, 12 (2), 425. <https://doi.org/10.3390/w12020425>.
- (24) Herzog, S.; Higgins, C. P.; Singha, K.; McCray, J. Performance of Engineered Streambeds for Inducing Hyporheic Transient Storage and Attenuation of Resazurin. *Environmental Science & Technology* **2018**, acs.est.8b01145. <https://doi.org/10.1021/acs.est.8b01145>.
- (25) Ibrahim, A.; Steffler, P.; She, Y. Estimating the Hyporheic Depth Beneath a Pool-Riffle Bedform Using the Rankine Body Analytical Method. *Journal of Hydrologic Engineering* **2020**, 25 (8), 04020034. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001944](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001944).
- (26) Krause, S.; Abbott, B. W.; Baranov, V.; Bernal, S.; Blaen, P.; Datry, T.; Drummond, J.; Fleckenstein, J. H.; Velez, J. G.; Hannah, D. M.; Knapp, J. L. A.; Kurz, M.; Lewandowski, J.; Martí, E.; Mendoza-Lera, C.; Milner, A.; Packman, A.; Pinay, G.; Ward, A. S.; Zarnetzke, J. P. Organizational Principles of Hyporheic Exchange Flow and

- Biogeochemical Cycling in River Networks Across Scales. *Water Resources Research* **2022**, 58 (3), e2021WR029771. <https://doi.org/10.1029/2021WR029771>.
- (27) Gomez-Velez, J. D.; Harvey, J. W. A Hydrogeomorphic River Network Model Predicts Where and Why Hyporheic Exchange Is Important in Large Basins. *Geophysical Research Letters* **2014**, 41 (18), 6403–6412. <https://doi.org/10.1002/2014GL061099>.
- (28) 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. How Hydrologic Connectivity Regulates Water Quality in River Corridors. *Journal of the American Water Resources Association* **2019**. <https://doi.org/10.1111/1752-1688.12691>.
- (29) Posselt, M. Transformation of Micropollutants in the Hyporheic Zone, 2020.
- (30) Singha, K.; Loheide II, S. P. Linking Physical and Numerical Modelling in Hydrogeology Using Sand Tank Experiments and COMSOL Multiphysics. *International Journal of Science Education* **2011**, 33 (4), 547–571. <https://doi.org/10.1080/09500693.2010.490570>.
- (31) Schmadel, N. M.; Ward, A. S.; Wondzell, S. M. Hydrologic Controls on Hyporheic Exchange in a Headwater Mountain Stream. *Water Resources Research* **2017**, 53 (7), 6260–6278. <https://doi.org/10.1002/2017WR020576>.
- (32) Herzog, S. P.; Ward, A. S.; Wondzell, S. M. Multiscale Feature-Feature Interactions Control Patterns of Hyporheic Exchange in a Simulated Headwater Mountain Stream. *Water Resources Research* **2019**, 55 (12). <https://doi.org/10.1029/2019WR025763>.
- (33) Zarnetske, J. P.; Haggerty, R.; Wondzell, S. M.; Baker, M. A. Dynamics of Nitrate Production and Removal as a Function of Residence Time in the Hyporheic Zone. *Journal of Geophysical Research: Biogeosciences* **2011**, 116 (G1). <https://doi.org/10.1029/2010JG001356>.
- (34) Tian, Z.; Zhao, H.; Peter, K. T.; Gonzalez, M.; Wetzel, J.; Wu, C.; Hu, X.; Prat, J.; Mudrock, E.; Hettinger, R.; Cortina, A. E.; Biswas, R. G.; Kock, F. V. C.; Soong, R.; Jenne, A.; Du, B.; Hou, F.; He, H.; Lundeen, R.; Gilbreath, A.; Sutton, R.; Scholz, N. L.; Davis, J. W.; Dodd, M. C.; Simpson, A.; McIntyre, J. K.; Kolodziej, E. P. A Ubiquitous Tire Rubber-Derived Chemical Induces Acute Mortality in Coho Salmon. *Science* **2020**, eabd6951. <https://doi.org/10.1126/science.abd6951>.
- (35) McConville, M. B.; Cohen, N. M.; Nowicki, S. M.; Lantz, S. R.; Hixson, J. L.; Ward, A. S.; Remucal, C. K. A Field Analysis of Lampricide Photodegradation in Great Lakes Tributaries. *Environ. Sci.: Processes Impacts* **2017**, 19 (7), 891–900. <https://doi.org/10.1039/C7EM00173H>.
- (36) Harvey, B. N.; Johnson, M. L.; Kiernan, J. D.; Green, P. G. Net Dissolved Inorganic Nitrogen Production in Hyporheic Mesocosms with Contrasting Sediment Size Distributions. *Hydrobiologia* **2011**, 658 (1), 343–352. <https://doi.org/10.1007/s10750-010-0504-4>.
- (37) Stelzer, R. S.; Bartsch, L. A. Nitrate Removal in Deep Sediments of a Nitrogen-Rich River Network: A Test of a Conceptual Model. *Journal of Geophysical Research: Biogeosciences* **2012**, 117 (G2). <https://doi.org/10.1029/2012JG001990>.
- (38) Hampton, T. B.; Zarnetske, J. P.; Briggs, M. A.; Singha, K.; Harvey, J. W.; Day-Lewis, F. D.; MahmoodPoor Dehkordy, F.; Lane, J. W. Residence Time Controls on the Fate of Nitrogen in Flow-Through Lakebed Sediments. *Journal of Geophysical Research: Biogeosciences* **2019**, 124 (3), 689–707. <https://doi.org/10.1029/2018JG004741>.
- (39) Hester, E. T.; Brooks, K. E.; Scott, D. T. Comparing Reach Scale Hyporheic Exchange and

- Denitrification Induced by Instream Restoration Structures and Natural Streambed Morphology. *Ecological Engineering* **2018**, *115*, 105–121. <https://doi.org/10.1016/J.ECOLENG.2018.01.011>.
- (40) Boano, F.; Revelli, R.; Ridolfi, L. Reduction of the Hyporheic Zone Volume Due to the Stream-Aquifer Interaction. *Geophysical Research Letters* **2008**, *35* (9), L09401. <https://doi.org/10.1029/2008GL033554>.
- (41) Magliozzi, C.; Coro, G.; Grabowski, R.; Packman, A. I.; Krause, S. A Multiscale Statistical Method to Identify Potential Areas of Hyporheic Exchange for River Restoration Planning. *Environmental Modelling & Software* **2018**. <https://doi.org/10.1016/J.ENVSOFT.2018.09.006>.
- (42) Craig, L. S.; Palmer, M. A.; Richardson, D. C.; Filoso, S.; Bernhardt, E. S.; Bledsoe, B. P.; Doyle, M. W.; Groffman, P. M.; Hassett, B. A.; Kaushal, S. S.; Mayer, P. M.; Smith, S. M.; Wilcock, P. R. Stream Restoration Strategies for Reducing River Nitrogen Loads. *Frontiers in Ecology and the Environment* **2008**, *6* (10), 529–538. <https://doi.org/10.1890/070080>.
- (43) Wu, L.; Singh, T.; Gomez-Velez, J.; Nützmann, G.; Wörman, A.; Krause, S.; Lewandowski, J. Impact of Dynamically Changing Discharge on Hyporheic Exchange Processes Under Gaining and Losing Groundwater Conditions. *Water Resources Research* **2018**. <https://doi.org/10.1029/2018WR023185>.
- (44) Ward, A. S.; Schmadel, N. M.; Wondzell, S. M. Time-Variable Transit Time Distributions in the Hyporheic Zone of a Headwater Mountain Stream. *Water Resources Research* **2018**, *54* (3), 2017–2036. <https://doi.org/10.1002/2017WR021502>.
- (45) Gomez-Velez, J. D.; Krause, S.; Wilson, J. L. Effect of Low-Permeability Layers on Spatial Patterns of Hyporheic Exchange and Groundwater Upwelling. *Water Resources Research* **2014**, *50* (6), 5196–5215. <https://doi.org/10.1002/2013WR015054>.
- (46) Montgomery, D. R.; Buffington, J. M.; Smith, R. D.; Schmidt, K. M.; Pess, G. Pool Spacing in Forest Channels. *Water Resources Research* **1995**, *31* (4), 1097–1105. <https://doi.org/10.1029/94WR03285>.
- (47) Herzog, S. P.; Eisenstein, W. A.; Halpin, B. N.; Portmann, A. C.; Fitzgerald, N. J. M.; Ward, A. S.; Higgins, C. P.; McCray, J. E. Co-Design of Engineered Hyporheic Zones to Improve in-Stream Stormwater Treatment and Facilitate Regulatory Approval. *Water (Switzerland)* **2019**, *11* (12). <https://doi.org/10.3390/w11122543>.
- (48) Merrill, L.; Tonjes, D. J. A Review of the Hyporheic Zone, Stream Restoration, and Means to Enhance Denitrification. *Critical Reviews in Environmental Science and Technology* **2014**, *44* (21), 2337–2379. <https://doi.org/10.1080/10643389.2013.829769>.
- (49) Hester, E. T.; Hammond, B.; Scott, D. T. Effects of Inset Floodplains and Hyporheic Exchange Induced by In-Stream Structures on Nitrate Removal in a Headwater Stream. *Ecological Engineering* **2016**, *97*, 452–464. <https://doi.org/10.1016/j.ecoleng.2016.10.036>.
- (50) Calfe, M. L.; Scott, D. T.; Hester, E. T. Nitrate Removal by Watershed-Scale Hyporheic Stream Restoration: Modeling Approach to Estimate Effects and Patterns at the Stream Network Scale. *Ecological Engineering* **2022**, *175*, 106498. <https://doi.org/10.1016/J.ECOLENG.2021.106498>.
- (51) Gordon, R. P.; Lautz, L. K.; Daniluk, T. L. Spatial Patterns of Hyporheic Exchange and Biogeochemical Cycling around Cross-Vane Restoration Structures: Implications for Stream Restoration Design. *Water Resources Research* **2013**, *49* (4), 2040–2055. <https://doi.org/10.1002/wrcr.20185>.