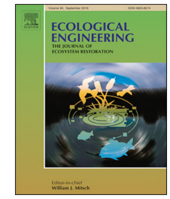
Ecological Engineering 124 (2018) 7–18

Contents lists available at ScienceDirect

Ecological Engineering

journal homepage: www.elsevier.com/locate/ecoleng



The multiscale effects of stream restoration on water quality

J. Thompson*, C.E. Pelc, W.R. Brogan III, T.E. Jordan

Smithsonian Environmental Research Center, Edgewater, MD, USA

ARTICLE INFO

Keywords:
Stream restoration
Nutrient pollution
Stormwater
Best management practice
Total mean daily loads
Chesapeake Bay

ABSTRACT

Stream restoration is often considered as an effective watershed management tool to reduce riverine loads of nitrogen, phosphorus, and suspended sediments, and meet government-mandated water quality goals. However, despite the billions of dollars which have been spent on stream restoration, questions remain about its effectiveness for improving water quality, as many studies report either mixed success or lack the adequate methodological framework to detect water quality improvements. In this study, we measured fluxes of nutrients and sediment in an eroded stream before and after restoration by filling the eroded channel with a mixture of sand, gravel, and woodchips stabilized with rock weirs at intervals along the channel. Our monitoring design used a before-after-control-impact (BACI) approach at two spatial scales, one at the reach-scale, and one farther downstream to detect whether reach-scale changes in nutrient and sediment loads propagated downstream. At the reach scale, we found that the restoration enhanced stream function, removing 44.8% of the phosphate, 45.8% of the total phosphorus, 48.3% of the ammonium, 25.7% of the nitrate, 49.7% of the total nitrogen, and 73.8% of the suspended sediment. However, due to hydrological variance, monitoring stations farther downstream suggested no detectable changes at the larger spatial scale relative to a reference stream, which highlights the challenges of detecting watershed-scale responses to small-scale stream restoration projects. This study provides a methodological framework for evaluating the effect of stream restoration on water quality at multiple scales and shows that reach-scale improvements may not be detectable at watershed-scales.

1. Introduction

Streams are unique as they are both receptors of watershed discharge, and chemically and biologically reactive conveyances that transport and transform water, nutrients, and particulate matter from terrestrial environments to larger water bodies (Cole et al., 2007, Gibson et al., 2015, Gomi et al., 2002). In the urban-suburban environment, increased development of impervious surfaces has disrupted the natural ability of streams and their floodplains to process nutrients and sequester sediment due to increased peak flows, reduced base flows, and enhanced channel erosion, which together limit water transit time and decrease habitat for organisms responsible for the biological retention of nutrients (Galster et al., 2006, Shuster et al., 2005, Walsh et al., 2005a).

Historically, urban storm water has been managed primarily via rapid transmission of storm water to streams to prevent flooding (Dunne and Leopold, 1978). The increasing recognition that urbanization and historical storm water management systems continue to cause negative impacts on the ecological health of freshwaters has led to an increased push for retrofitting urban and suburban landscapes with green infrastructure, such as storm water detention ponds, to ameliorate the negative impacts of urbanization on receiving waters (Weber et al., 2006, Walsh et al., 2005b). Some studies have reported the relative success of green infrastructure in reducing nutrient and sediment discharges to streams at the watershed scale (Pennino et al., 2016, Dietz and Clausen, 2008).

Data from a recent study encompassing the period 1945–2012 indicated that although nitrogen loading within the Chesapeake Bay is beginning to decline, the reductions still lag behind many comparable estuaries undergoing intense management (Harding et al., 2016). As part of a push to improve water quality, the Chesapeake watershed Total Maximum Daily Load (TMDLs), adopted in 2010, have dictated pollutant reduction requirements of 25% for total nitrogen (TN), 24% for total phosphorus (TP), and 20% for total suspended sediment (TSS). The U.S. Environmental Protection Agency (EPA) has set a tight timeline requiring the implementation of all necessary pollution control measures to achieve these levels by 2025 (https://www.epa.gov/chesapeake-bay-tmdl).

Within the mid-Atlantic region of the U.S., and particularly the coastal plain of the Chesapeake Bay watershed, stream restoration has become an increasingly used tool to improve water quality and meet water quality goals such as TMDLs by enhancing the natural pollutant-

* Corresponding author.
E-mail address: thompsonj@si.edu (J. Thompson).

https://doi.org/10.1016/j.ecoleng.2018.09.016
Received 12 February 2018; Received in revised form 7 September 2018; Accepted 12 September 2018

attenuating functions of streams. Meeting these goals through stream restoration has come at a large financial cost globally, with a total of over 9 billion dollars invested in stream restoration projects in the contiguous U.S. (Bernhardt et al., 2005), and costs per project averaging 3.21 million euros in Europe (Pander and Geist, 2013).

Recently, stream restoration has evolved from stream stabilization techniques to dramatic geomorphological alterations. One contemporary restoration approach backfills deeply incised channels using a mixture of sand, gravel, and woodchips, and places large rock weirs across the channel to restore pool and riffle sequences and enhance stream-floodplain connectivity (Brown et al., 2010). These newer, more invasive techniques, termed regenerative stormwater conveyance, have come under scrutiny with suggestions that philosophically, stream restoration has moved stormwater management structures into the stream, thereby shifting the onus of responsibility away from the watershed and onto the channel (Palmer et al., 2014). Recent studies have indicated the variable effect that stream restoration has on nitrogen and sediment retention and observed a limited capacity for pollutant removal during high flows (Filoso et al., 2015, Filoso and Palmer, 2011).

Despite large public and private investment, there have been relatively few published studies evaluating the effect that these newer stream restorations have on water quality and hydrology. A recent review by Newcomer Johnson et al. (2016) found 27 peer reviewed studies of nutrient retention within streams that were restored by raising the stream bed to near bank-level, as in our study. Yet 19 of these studies were assessed using short term nutrient addition experiments to determine nutrient uptake and the potential for denitrification, with a further four assessing the effect by differences in nutrient concentrations between treatment and control reaches. Only four studies out of 27 in the review by Newcomer Johnson et al. (2016) used a mass balance approach to determine the nutrient retention effect of this type of restoration, with none of the synthesized studies assessing the effect of restoration using a before-after-control-impact design.

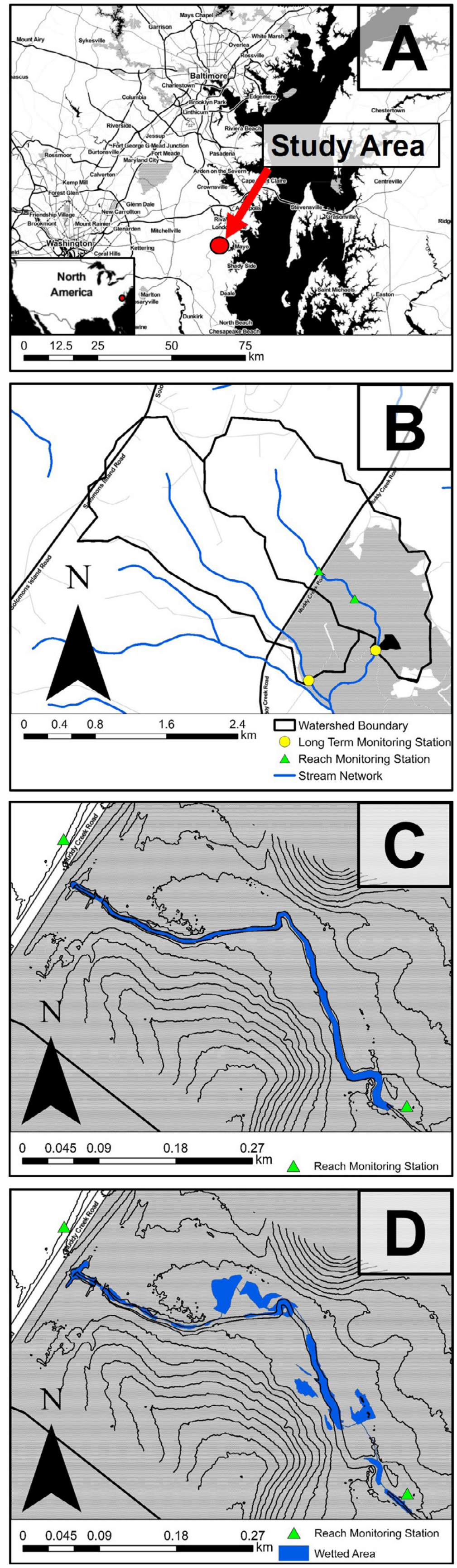
Inadequate monitoring of restoration projects is not unique to the U.S. and is an issue in Europe. Pander and Geist (2013) found that 87% of restoration projects in Bavaria, Germany from 1994 to 2011 had no monitoring data, with only 4% of restoration having any form of pre-restoration characterization. Monitoring the success of restoration is critical, as it is only with adequate data that success can be judged (Pander and Geist, 2013; Geist, 2015; Geist and Hawkins, 2016).

Many existing studies have been limited to highly urban environments (Filoso et al., 2015, Filoso and Palmer, 2011, Williams et al., 2017, Cizek et al., 2017), yet stream restoration and stormwater management are not unique to urban environments. Further, while isolating the water quality benefits of stream restoration at the reach scale provides useful information on their effectiveness, there is also a need to place these reach-scale water quality changes within a larger spatial perspective to see if local water quality changes propagate farther downstream. In this study we quantified the effects of a stream restoration on hydrology and on net removal of nutrients and suspended solids within the restored reach and downstream of the reach within a larger watershed. We used a before-after-control-impact design to test our hypotheses that stream restoration reduces nutrient and sediment fluxes at both the reach and watershed scales.

2. Methods

2.1. Study system and design

We studied the restoration of a 452 meter (m) reach of the North Branch of Muddy Creek in Edgewater, Maryland, USA (Fig. 1), which was constructed between late December 2015 and February 2016. Before restoration, the reach had become deeply incised (up to 2.9 m below bankfull height) downstream of the culvert under Muddy Creek Road (MD 468). The reach was restored using a regenerative storm water conveyance design (Brown et al., 2010), which involved filling



(caption on next page)

Fig. 1. a) Location of study site within the Chesapeake Bay region, b) map of the two paired watersheds showing the two long-term monitoring treatment and reference watersheds (yellow circles) and the two reach monitoring watersheds at the inlet and outlet (green triangles), c) map of the wetted extent within the monitored reach prior to restoration, and d) map of the wetted extent after restoration. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the incised channel with a layer of gravel topped with a mixture of sand and woodchips. The fill was stabilized by a series of rock weirs, and berms were added in places to deflect stream flow out onto the floodplain. This design aimed to increase the residence time of water and encourage frequent flooding of the floodplain to increase sediment deposition, and nutrient removal via biological uptake. This design modified the shape and area of the stream channel as shown in Fig. 1.

Our study examined two spatial scales. We measured the effect of the restoration at the reach scale by comparing flows of water, nutrients, and suspended solids at the inlet and outlet of the restored reach before and after the restoration. We also measured the effect at a larger scale by comparing fluxes 609 m farther downstream of the restored reach, herein referred to as the treatment watershed, to fluxes from an adjacent watershed, herein referred to as the reference watershed, before and after restoration (Fig. 1). These two watersheds have similar areas and land cover composition (Table S1; *supplementary information*) and are among several sub-watersheds of the Rhode River estuary that have been monitored by the Smithsonian Environmental Research Center since the mid-1970s for discharges of water, suspended solids, nitrogen, phosphorus, and organic carbon (Jordan et al., 1997, Correll et al., 1992). Comparisons of these watersheds have previously used a before-after-control-impact design to assess nutrient and sediment retention within a beaver pond established in the treatment watershed, historically called watershed 101 (Correll et al., 2000). The inlet of the restored reach receives water from 59% of the treatment watershed while the outlet receives water from 66% of the treatment watershed. Thus, the treatment watershed receives discharges of water from an additional 34% of its area that does not flow through the restored reach. This area of the treatment watershed that is downstream of the restoration includes the previously studied beaver pond.

2.2. Discharge measurements and water sampling

We measured stream depth at each sampling station using a float and counterweight within a stilling well. The float and counterweight were connected to a potentiometer with the analogue signal converted to stream depth using a pre-determined constant. For the long-term monitoring treatment and reference stations, discharge was approximated using water depth within a 120^0 v-notch weir. For the temporary stations at the restored reach inlet and outlet, we developed rating curves for the period before the restoration. After the restoration was completed, we used a v-notch weir at the outlet and developed a new rating curve for the inlet.

To ensure representative sampling, we conducted flow-paced sampling at all stations using an ISCO 3700 autosampler controlled by a Campbell Scientific CR10X datalogger. Flow-paced sampling was achieved by programing the datalogger to trigger the autosampler upon reaching a set threshold of total cumulative discharge, which automatically reset after each sample. Each discrete sample taken by the autosampler was split and composited into two carboys using a T-shaped splitter, with one carboy containing 25 ml of 12 N H_2SO_4 for immediate sample preservation, and the other carboy unpreserved. The acidified sample was used for nutrient analysis, with the unpreserved sample used for total suspended sediment (TSS). Typically, each composite sample represented an integration of 20–60 samples taken over seven days, with the sampling frequency increasing with increasing flow rate (e.g., Fig. S1; *supplementary information*). Therefore, our sampling method yielded flow-weighted mean concentrations.

We modified the sampling protocol after the restoration was complete to include additional composited storm samples. To achieve this, we added two electromagnetic solenoid valves so that when each flow-paced sample was taken, the first solenoid valve was activated, and the sample was distributed to the original two carboys. When stage increased above baseflow the autosampler was triggered twice, with the first solenoid valve distributing the first sample to the original two carboys, and the second solenoid valve distributing the second sample to two storm-only carboys. We set the storm activation threshold prior to each storm.

2.3. Laboratory analysis

We used analytical methods consistent with previous studies at the monitoring sites to ensure that the results of this study were comparable to the historical data (Jordan et al., 1991). Because our method of sample collection preserved samples in H_2SO_4 immediately at the time of sampling, particulate phosphate (PO_4) and ammonium (NH_4) were converted to dissolved PO_4 and NH_4 by the acid preservative, and therefore, both PO_4 and NH_4 should be considered as Total PO_4 and Total NH_4, which henceforth, we refer to as PO_4 and NH_4, respectively.

Samples for inorganic nutrient analysis were collected from the acid-preserved sample and filtered prior to analysis using a prewashed 0.45 µm Millipore filter. On the undigested filtered samples, PO_4 was measured colorimetrically after reaction with stannous chloride and ammonium molybdate (American Public Health Association, 2005). NH_4 was oxidized to nitrite (NO_2) using alkaline hypochlorite (Strickland and Parsons, 1977), and reacted with sulfanilamide before being determined colorimetrically (American Public Health Association, 2005). Nitrate (NO_3) was reduced to NO_2 by passing samples through a cadmium column, before being reacted with sulfanilamide and measured colorimetrically (American Public Health Association, 2005). We report the total concentration of NO_3 plus NO_2 because acid preservation may have altered the partitioning between NO_3 and NO_2. For simplicity, we refer to this total as NO_3. Total Kjeldahl N (TKN) was determined by digestion of samples to NH_4 with sulfuric acid, Hengar granules, and hydrogen peroxide (Martin, 1972). The NH_4 in the digestate was steam distilled and analyzed using an Astoria Pacific International (API) 300 micro-segmented flow through analyzer with digital detector (API method A303-S021, APHA, 1995). Total nitrogen (TN) was calculated as the sum of TKN and NO_3. We analyzed Total P on unfiltered samples from the acid preserved sample, which was digested to PO_4 with perchloric acid (King, 1932), and subsequently measured PO_4 colorimetrically after reaction with stannous chloride and ammonium molybdate (American Public Health Association, 2005). Total suspended solids (TSS) were measured by filtering the unpreserved composite sample through prewashed and preweighed 0.4 µm Nuclepore filters and reweighed after being dried to a consistent weight (American Public Health Association, 2005).

2.4. Load, flow-weighted mean concentrations, and retention calculations

Measured nutrient and TSS concentrations within the composite samples represented the flow-weighted mean concentrations for each sampling interval. Unlike many interpolation-based load estimation methods which approximate flow-weighted mean concentrations (FWMCs) based on a limited number of annual samples, our sampling method typically combined 20–60 individual sampling occasions per week, giving us confidence in our sample's representativeness of the true FWMC. Accordingly, we estimated total loads by summing the total discharge for each sampling interval and multiplying it by the corresponding FWMC.

At the long-term stations, sampling intervals prior to restoration were typically one month in duration, whereas after restoration, intervals were reduced to seven days. We aggregated loads to monthly totals to ensure comparability between the periods before and after restoration. For the temporary stations directly above and below the treatment reach, we kept our intervals at seven days and did not aggregate to monthly totals. However, when plotting data, we computed daily loads for visualization clarity.

Total load retention within the treatment reach was estimated via mass balance by subtracting the load at the outlet from the load at the inlet. The net change was then divided by the load at the inlet to express retention as a percentage. Thus, all positive net loads represent retention within the treatment reach and all negative net loads represent increasing loads through the reach. Sampling intervals without corresponding inflow-outflow samples (due to no flow events or mechanical failure) were omitted from analysis.

2.5. Flux calculations from integrated storm samples

Hydrograph separation of discrete events is inherently uncertain and has long been a focus of research within hydrological science (Hooper and Shoemaker, 1986). In this respect, due to the uncertainties with distinguishing event flow from pre-event flow using numerical hydrograph separation, we did not conduct a mass balance for storm loads. Instead, we chose to estimate the mean flux for each integrated storm sample by averaging the flow rate ($m^3 s^{-1}$) for an entire storm and multiplying this mean flow rate by the concentration measured in the storm composite sample. This approach minimized the uncertainties in distinguishing between event and pre-event water whilst also allowing the assessment of the restoration performance in higher flows.

2.6. Hydrological analysis

We conducted our hydrological analysis for the four monitoring stations in three parts. First, we computed flow duration curves (FDCs) for each site on an annual basis, before and after restoration. Second, we characterized hydrological flashiness by computing Q5:Q95 ratios. Hydrograph Q5:Q95 ratios numerically describe hydrological flashiness by dividing the flow value ($m^3 s^{-1}$) which is exceeded 5% of the time against the flow value which is exceeded 95% of the time over a predefined observation period. Thus, greater discrepancies between base flow and storm flow result in a larger Flashiness Index (FI). Finally, we used a digital recursive filter to compute base flow and quick flow following the method of Lyne and Hollick (1979), with an alpha level of 0.975. The baseflow index (BFI) was then computed as the ratio of modeled baseflow to total stream flow over a pre-defined observation period.

2.7. Data analysis

Randomized Intervention Analysis (RIA) was used to assess the changes in loads and FWMCs in response to the restoration (Carpenter et al., 1989). RIA is a widely used statistical method to assess ecosystem level changes within a BACI design and is regularly applied to assess ecological and water quality changes in response to experimental manipulation (Hood et al., 2018, Ukonmaanaho et al., 2016, Wallace et al., 2015). In our study, we applied RIA to paired observations within the reference and treatment sites before and after the restoration, with the test statistic derived as the difference between the mean differences in paired observations before and after restoration. We generated 10,000 permutations of the time series of differences without regard to time (i.e. before or after restoration) to assess whether a nonrandom change occurred in loads and FWMC since restoration. RIA was applied to both spatial scales; at the reach scale using the inlet as a reference and the outlet as a treatment, and at a larger scale using the reference and treatment watersheds. P-values were derived as the proportion of randomized time series with greater mean differences than the RIA test statistic. All observations were $\log 10\,(x + 1)$ transformed for the analysis.

Because RIA evaluates the relative differences between a treatment and reference system before and after restoration, there is no requirement for system similarity prior to manipulation. However, one assumption, as with BACI designs in general, is that the relative differences between the treatment and reference systems are similar through time, and would remain similar if no manipulation took place (Stewart-Oaten and Bence, 2001). Streams are particularly vulnerable to nonsteady differences through time, as water quality is directly influenced by changes in weather and land use in their watersheds. Our paired study watersheds were rural, and land use did not change throughout our observation period giving us confidence in our application of BACI and RIA. All analyses were conducted in R version 3.2.3 (R Development Core Team, 2015).

3. Results

3.1. The effect of stream restoration on hydrology

The FDCs of the inlet and outlet of the reach were broadly similar prior to restoration at discharges $> 0.02\,m^3 s^{-1}$ but diverged at discharges $< 0.02\,m^3 s^{-1}$, indicating sustained flows at the outlet of the reach relative to the inlet (year 2015 in Fig. 2). This behavior changed after restoration; at low flows ($< 0.01\,m^3 s^{-1}$), FDCs for 2016 and 2017 indicated that the inlet had sustained flows compared to the outlet. The observed change in low flow behavior between the inlet and outlet appeared to be unrelated to changes in the total discharge, as the patterns in the annual water yield remained the same for both monitoring locations after restoration, and did not indicate diverging water yields between the inlet and outlet (Fig. S2, *supplementary information*).

As with the reach-scale monitoring locations, the FDCs of the reference and treatment watershed monitoring stations were similar at discharges $> 0.02\,m^3 s^{-1}$ both before and after the restoration. However, in contrast to the inlet and outlet monitoring stations, differences in FDCs between the reference and treatment watersheds were apparent at discharges $< 0.01\,m^3 s^{-1}$ after the restoration, with the treatment watershed exhibiting sustained flows for longer periods compared to the reference watershed.

Despite changes in FDCs for all four monitoring stations, the modeled BFIs were similar between the years before and after the restoration (Fig. 3a). Inter-annual variability in the BFI for all stations were similar, suggesting that the ratio of groundwater inputs to total stream flow were unaffected by the restoration.

The magnitude of change in the FI before and after the restoration differed among the monitoring locations (Fig. 3b). The change in the FI between the reach inlet and outlet was surprisingly minor and, in contrast to our expectations, the FI at the inlet decreased after the restoration whereas the FI at the outlet remained similar. This suggests that the restoration dampened the hydrological response of the water flowing into the restoration, while having a minor effect on the hydrological response of water flowing out of the restoration. Yet these changes were small, particularly in comparison to the magnitude of change observed at the treatment watershed monitoring location, which had a FI that was 8.6 ($\pm$ 1.9 S.D.) points higher than the reference watershed prior to restoration but 7 ($\pm$ 10.1 S.D.) points lower than the reference the first year after restoration. These differences suggest that although restoration did not produce any meaningful reduction in hydrological flashiness within the restored reach, the effect was observable farther downstream. The mechanism driving this change in the treatment watershed is unclear as most of the watershed is comprised of forests, agriculture, and low-density residences which remained constant throughout our study, precluding land use changes downstream of the restored reach from having an impact. We did not have a long enough time series for statistical analysis, so more research would be required to determine whether a statistical change was apparent.

We assessed the change in the annual water yields at all stations and

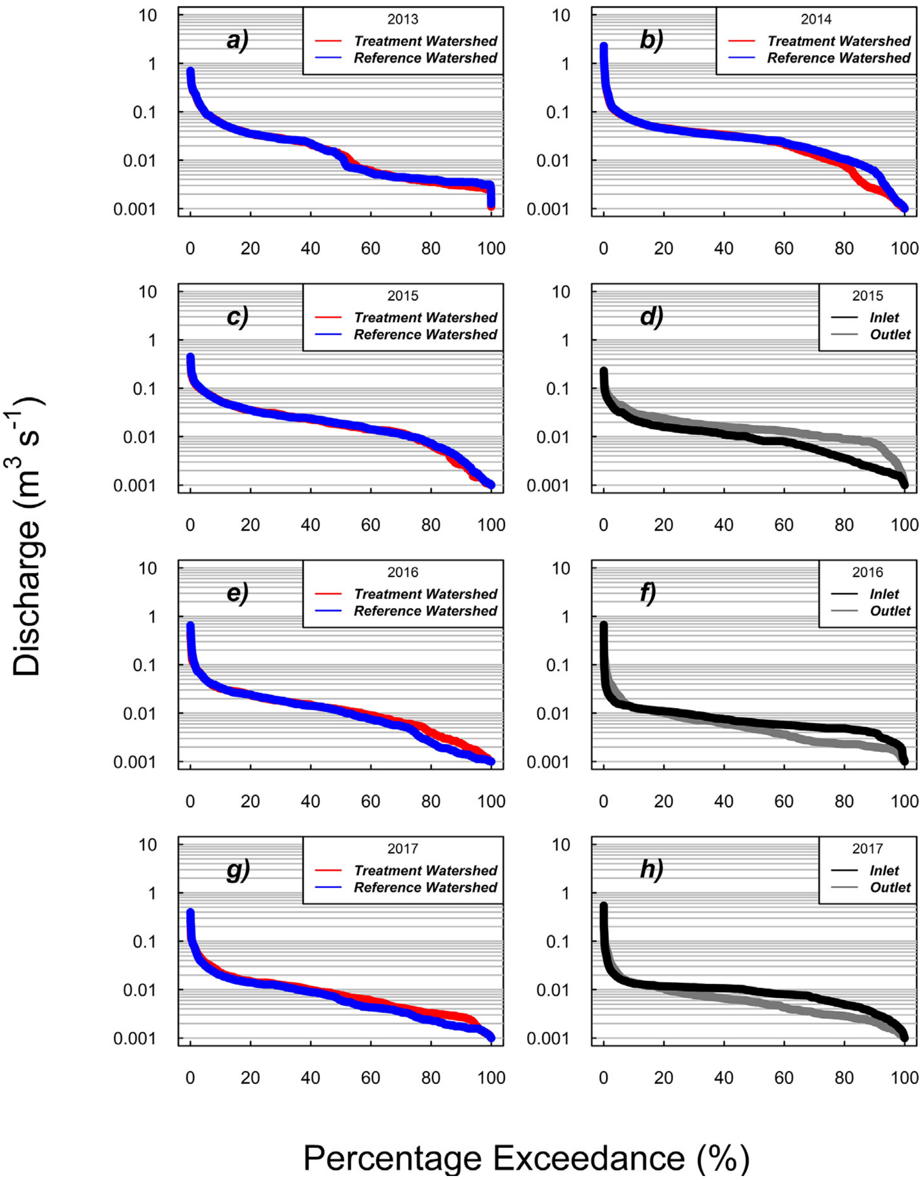


Fig. 2. Flow-duration curves for the long-term monitoring stations (blue and red) from 2013 to 2017 (a–c, e, and g), and for the reach monitoring stations (gray and black) from 2015–2017 (d, f, and h). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

found that changes between the years prior to and post restoration were much larger than the differences between the stations (Fig. S2 *supplementary information*). Differences in the annual water yield of the reach outlet and inlet were consistent before and after restoration, with the water yield at the outlet being 21.7% greater than the inlet before the restoration and 19.4% greater after restoration. Differences between water yields of the reference and treatment watersheds were also consistent before and after the restoration, with the water yield of the reference being 8.5% greater than the treatment prior to restoration, and 9.3% greater after restoration. Despite the difference in water yields between each monitoring station being similar before and after restoration, the absolute water yields decreased after restoration by 39.1%, 40.3%, 39%, and 38.5% for monitoring stations at the inlet, outlet, treatment watershed, and reference watershed, indicating less stream runoff was observed during the period after restoration.

3.2. The effect of the stream restoration on water quality

3.2.1. Reach changes in nutrient concentrations and loads

Our stream sampling method was set up to take flow-paced samples for the entire duration of the monitoring period without interruption. On occasion, mechanical failure prevented paired observations at both the inlet and outlet, and reference and treatment watershed stations. Further, the two streams we monitored as part of this study are not perennial and often experienced cessation of flow from late summer to early winter. Consequently, our data completeness at the inlet and outlet was 65.6% of the pre-restoration period, and 64.9% of the post-restoration period. Our samples represented 66.3% of all stream flow for the monitoring period prior to restoration and 59.2% of all stream flow for the monitoring period after restoration. Data completeness for the treatment and reference watersheds was 83.7% of the pre-restoration period, and 50% of the post-restoration period, which resulted in sampling 92.6% of the observed flow prior to restoration and 53.1% of the observed flow after restoration. Despite periodic mechanical problems and episodes of no-flow, our sampling regime can be considered more representative of the annual loads and concentrations than other studies that do not use a flow-integrated sampling approach.

We observed large changes in FWMC between the periods before and after restoration. Prior to the restoration, the FWMC at the inlet vs. the outlet typically fell along the 1:1 line indicating negligible changes

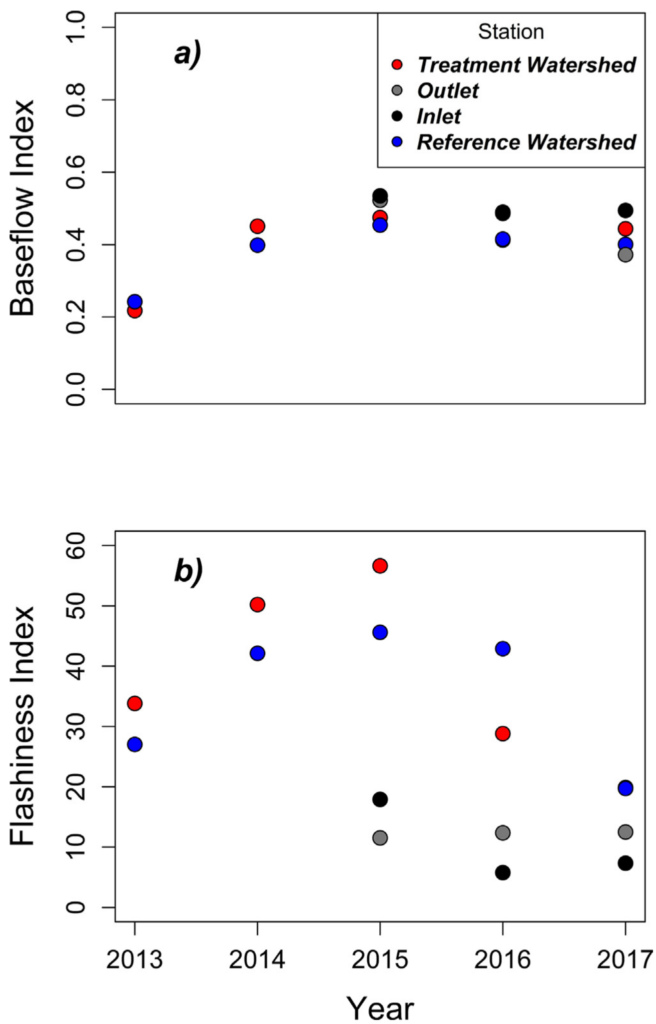


Fig. 3. a) Modeled baseflow indices for the long-term monitoring stations (blue and red) from 2013 to 2017, and for the reach monitoring stations (gray and black) from 2015–2017, and b) Q5:Q95 flashiness indices for the long-term monitoring stations (blue and red) from 2013 to 2017, and for the reach monitoring stations (gray and black) from 2015–2017. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

in concentrations along the reach (Fig. 4). Conversely, after restoration PO_4, Total P, TSS, NH_4, NO_3, and Total N typically showed repeatedly higher concentrations at the inlet than the outlet, indicating a reduction in nutrient concentrations. At the outlet, concentrations of most water quality parameters decreased since restoration by over 10%, with average NO_3 concentrations decreasing by 71.4% after restoration. Only NH_4 was found to increase at the outlet after the restoration, with the mean concentrations 3.4 times higher since the restoration (Table 1 and Fig. 4).

FWMCs of NH_4, Total N, and TSS increased at the inlet after restoration. Mean concentrations of TSS more than doubled, with NH_4 exhibiting an 8.5-fold increase after restoration. These changes may reflect interannual variability or perhaps pooling upstream of the inlet that occurred after the restoration. After periods of low flow, we measured high FWMCs of NH_4 and TSS at the inlet, particularly just after restoration construction (Fig. S3, *supplementary information*). Therefore, increases in the upstream FWMCs could also be due to lower flows observed in the period after restoration, due to coincidental decrease in precipitation.

Although the cause for increases in the inlet concentrations remain unclear, comparing the relative differences between the FWMCs of the stream water flowing into and out of the reach indicated net reductions of concentrations for all parameters following restoration. We assessed the relative differences between the inlet and outlet FWMCs using RIA, which indicated that the restoration significantly reduced FWMCs of Total P ($P = 0.0124$), NO_3 ($P = 0.008$), and Total N ($P = 0.0132$) at an alpha level of 0.05. Interestingly, RIA only indicated a reduction of PO_4 and TSS at an alpha level of 0.1 ($P = 0.077$ and $P = 0.085$; Table 1) and showed no significant change in NH_4 concentrations ($P = 0.1227$; Table 1). TSS, PO_4, and NH_4 concentrations had particularly large variability over the study period, with standard deviations typically close to, or larger than mean concentrations. RIA is best applied where ecosystem change is unidirectional. As particulate transport is inherently stochastic, it is unsurprising that RIA analysis did not detect significant changes in the FWMC of TSS, PO_4, and NH_4 (which contained the particulate fraction due to the acid preservative) after restoration, despite large reductions in mean concentrations within the reach overall.

As with FWMCs, we also observed large changes in loads after restoration. Prior to the restoration, weekly loads at the inlet and outlet were similar, suggesting little retention within the reach (Fig. 5). After the restoration, loads of PO_4, Total P, TSS, NH_4, NO_3, and Total N were generally lower at the outlet than at the inlet, indicating enhanced net retention (Table 1). At the outlet, most mean weekly loads decreased by 24%, while average NO_3 loads decreased by 67.3%. Only NH_4 load was found to increase at the outlet compared to loads observed before the restoration, with the mean load more than doubling after restoration (Table 1). Changes in the loads at the inlet were also variable, with mean Total P, NH_4, Total N, and TSS increasing after restoration.

RIA indicated statistically significant net reduction in weekly loads after restoration for PO_4 ($P = 0.018$) and Total P ($P = 0.012$) at the 0.05 alpha level, with NH_4 ($P = 0.082$), and TN ($P = 0.0534$) significant at the 0.1 level. Surprisingly, despite large reductions of NO_3 and TSS loads after restoration, RIA did not indicate that those reductions were statistically significant ($P > 0.01$).

We also estimated total mass removal (i.e., total inflowing mass minus total outflowing mass) for the pre-restoration and post-restoration periods (Fig. 6). This mass balance for the entire monitoring period provided a clear assessment of restoration efficiency. Prior to restoration, except for NO_3 and Total N, total loads generally increased within the reach, with PO_4, Total P, NH_4, and TSS increasing by 21.8%, 29.5%, 16.4%, and 18.4%, respectively. After restoration, total mass removal for all parameters was positive, decreasing the load by as much as 73.8% (TSS) and as little as 25.7% (NO_3).

3.2.2. Net retention during storms

In addition to our flow-paced composite sampling, which included samples across all flow conditions, we also conducted flow-paced composite sampling to exclusively isolate storm flows. We monitored eight storm events after the restoration at both the inlet and outlet out of a total of 24 storms. We calculated net flux ($kg\,m^3\,s^{-1}$) for all monitored storms and plotted these net fluxes against total event precipitation (a surrogate for storm size). In contrast to our expectations, we found few instances where net flux varied with storm size and only Total P and PO_4 indicated decreases in retention efficiency with larger storms (Fig. S4, *supplementary information*), although net flux and storm rainfall were not significantly correlated. We did not find NO_3 retention to have any clear patterns with total precipitation. However, the storm sizes we monitored were relatively small (10–40 mm of rainfall per storm) for this region, and thus future work should address retention efficiency across a greater range of rainfall.

3.3. The effect of reach scale restorations on downstream nutrient export

For both watersheds, FWMCs increased and loads decreased after the restoration, except for NO3 (Table 2 and Fig. 7 and Fig. 8). Those changes likely reflect the coincidental decrease in water yield after

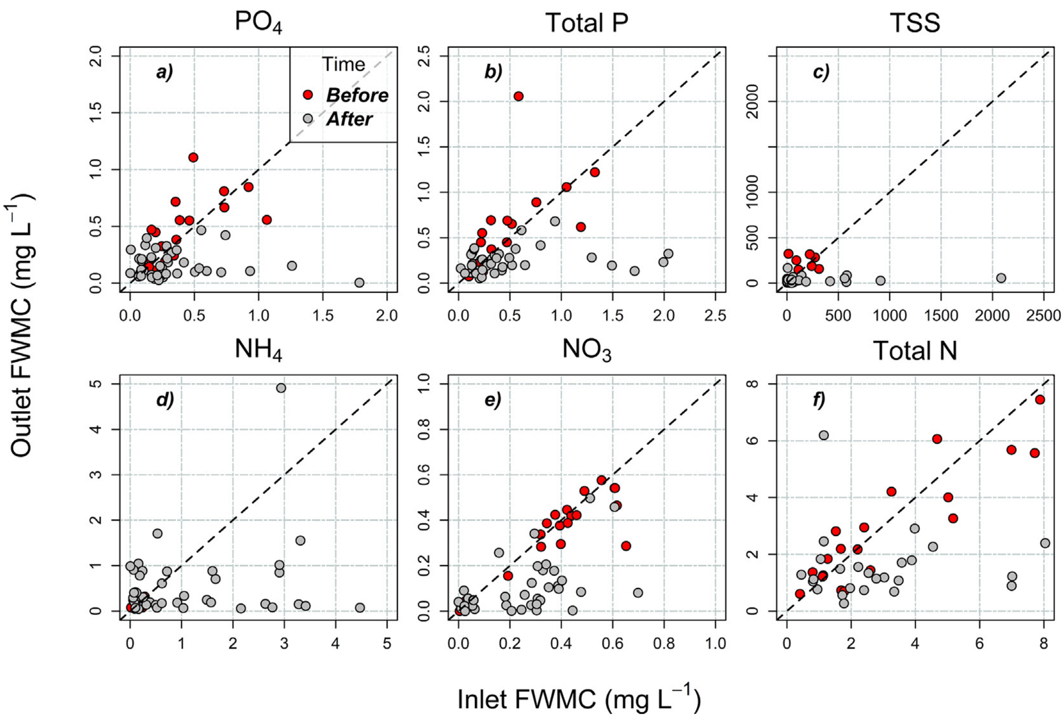


Fig. 4. Relationship between the inflowing and outflowing flow-weighted mean concentrations of a) PO_4, b) Total P, c) TSS, d) NH_4, e) NO_3, and f) Total N prior to the restoration (red points), and after the restoration (gray points). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Mean weekly FWMCs and loads at the inlet and outlet before and after restoration, with standard deviations in parenthesis. P-values derived from RIA analysis, where *indicates $P < 0.1$, **indicates $P < 0.05$, and ***indicates $P < 0.01$.

	Inlet		Outlet		P
	Before	After	Before	After	
FWMC (mg L)					
PO_4	0.38 (0.28)	0.42 (0.63)	0.43 (0.29)	0.17 (0.12)	0.0766*
Total P.	0.44 (0.37)	0.51 (0.53)	0.55 (0.48)	0.24 (0.13)	0.0124**
NH_4	0.15 (0.08)	1.27 (1.5)	0.16 (0.08)	0.55 (0.81)	0.1227
NO_3	0.38 (0.2)	0.25 (0.18)	0.35 (0.17)	0.1 (0.12)	0.0084***
Total N.	3.12 (2.4)	5.64 (9.41)	2.92 (2.05)	1.88 (1.91)	0.0132**
TSS	90.59 (103.2)	191.32 (404.64)	112.24 (118.5)	33.11 (32.56)	0.0852*
Load (kg)					
PO_4	1.13 (0.88)	1.13 (1.62)	1.37 (1.18)	0.62 (1.2)	0.0176**
Total P.	1.39 (1.16)	1.51 (2.38)	1.79 (2)	0.82 (1.17)	0.0122**
NH_4	0.56 (0.42)	3.01 (3.7)	0.65 (0.67)	1.56 (2.5)	0.0815*
NO_3	1.86 (2.08)	0.69 (0.76)	1.59 (1.65)	0.52 (1.08)	0.2807
Total N.	12.22 (12.35)	12.81 (18.03)	10.84 (9.89)	6.45 (9.14)	0.0534*
TSS	291.83 (338.98)	471.22 (865.62)	345.43 (382.85)	123.46 (240.44)	0.1243

restoration (Fig. S2, *supplementary information*).

Using RIA, we assessed the effects of the restoration on FWMCs and loads leaving the treatment watershed by comparing them to FWMCs and loads leaving the reference watershed before and after the restoration. Surprisingly, despite large changes at the outlet of the restored reach, only loads of PO_4 were found to significantly decrease ($P = 0.059$) at an alpha level of 0.1, in water exiting the treatment watershed about 609 m downstream of the outlet of the restored reach. No detectable changes were observed in the loads of other parameters exiting the treatment watershed. Likewise, we were not able to detect any change in FWMCs leaving the treatment watershed resulting from the restoration. Thus, for our study watershed, we cannot reject our null hypothesis that reach scale stream restoration does not reduce watershed nutrient and sediment fluxes, despite reductions at the reach scale being large and statistically significant.

4. Discussion

4.1. The effect of restoration on hydrology

In contrast to the dramatic geomorphological changes resulting from the restoration, FDCs before and after restoration were broadly similar. However, notable divergence between the inflowing and outflowing FDCs were observed at low flows (0.001–0.01 $m^3 s^{-1}$). Prior to restoration, the FDCs indicated that flows were sustained for longer periods at the outlet than at the inlet of the restored reach (Fig. 2). After restoration, this pattern reversed with flows being sustained for longer periods at the inlet than at the outlet (Fig. 2). Annual Q5:Q95 flashiness ratios indicated that flashiness remained broadly consistent at the outlet over the three-year monitoring period. At the inlet, treatment, and reference stations flashiness indices had comparable inter-annual variability (Fig. 3b). For the year prior to restoration, the inlet had a higher FI than the outlet. After the restoration, the outlet had a higher FI than the inlet. As FI is influenced by the ratio of high to low flows, these subtle changes are likely reflective of the changes in FDCs.

Although suppression of low flows by the restoration may run counter to the aims of the restoration, this finding is not unique. Hammersmark et al. (2008) investigated the hydrological effect of stream restoration in a Montane Meadow in Northern California, USA, and used hydrological modeling to determine changes in hydrological processes before and after restoration. Within their restoration, Hammersmark et al. (2008) found significant increases in surface and subsurface storage, reduced low flows, and a reduction in total flow, half of which was attributed to evapotranspiration. Similar findings were also reported by Cizek et al. (2017) from a restoration in North Carolina, USA, who found significant amounts of inflowing water converted to shallow interflow. In our study, we saw increases in

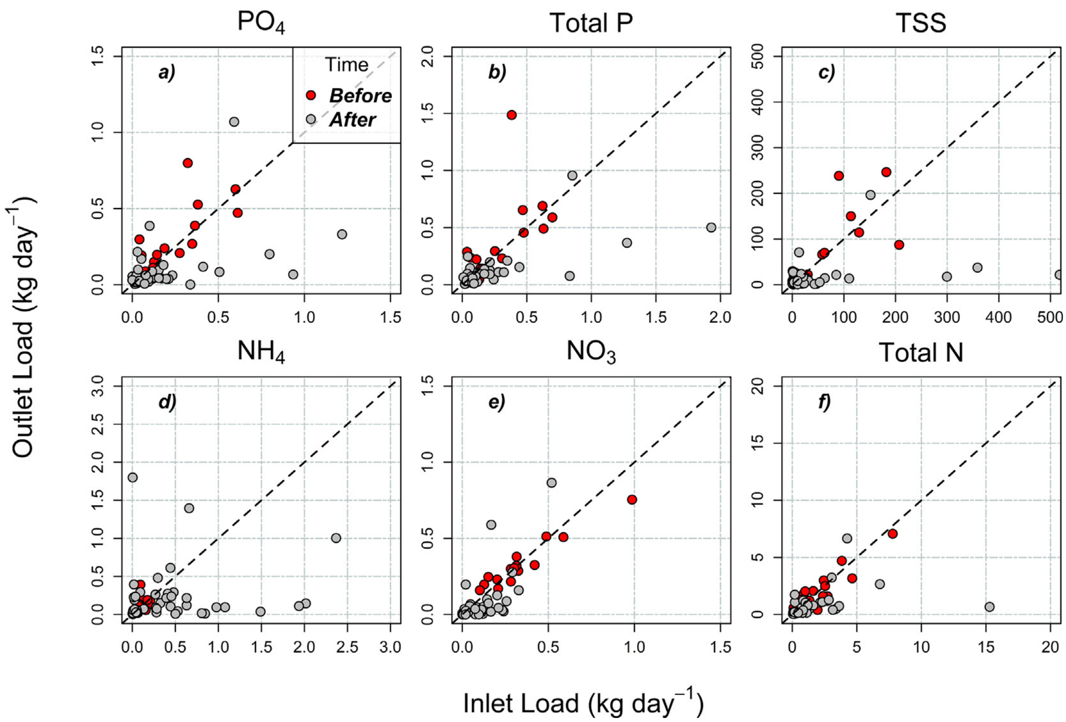


Fig. 5. Relationship between the inflowing and outflowing daily loads of a) PO_4, b) Total P, c) TSS, d) NH_4, e) NO_3, and f) Total N prior to the restoration (red points) and after the restoration (gray points). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

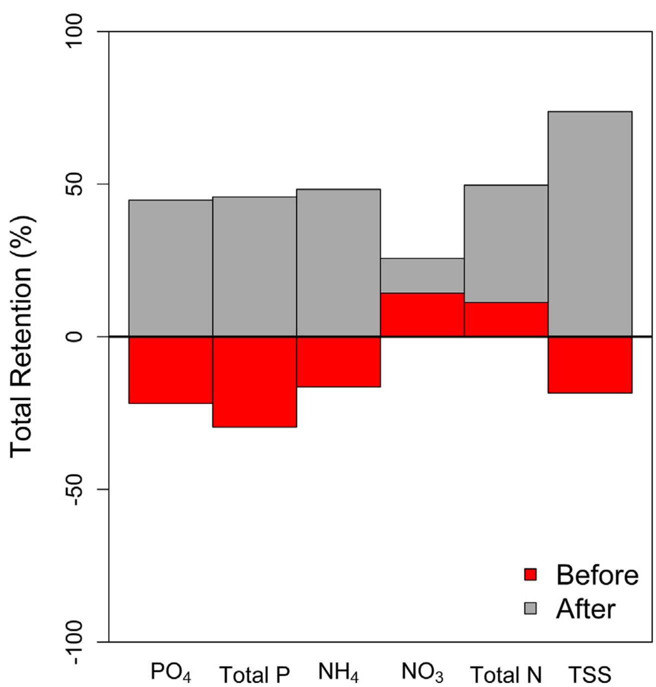


Fig. 6. Total mass retention of PO_4, Total P, TSS, NH_4, NO_3, and Total N before the restoration (red bars), and after the restoration (gray bars). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Mean weekly FWMCs and loads of at reference and treatment before and after restoration, with standard deviations in parenthesis. P-values derived from RIA analysis, where *indicates $P < 0.1$, **indicates $P < 0.05$, and ***indicates $P < 0.01$.

	Reference		Treatment		P
	Before	After	Before	After	
FWMC (mg L)					
PO_4	0.17 (0.17)	0.21 (0.24)	0.13 (0.1)	0.16 (0.23)	0.272
Total P.	0.24 (0.24)	0.27 (0.26)	0.19 (0.15)	0.22 (0.28)	0.539
NH_4	0.15 (0.17)	0.18 (0.22)	0.21 (0.24)	0.25 (0.23)	0.698
NO_3	0.31 (0.14)	0.25 (0.17)	0.13 (0.11)	0.11 (0.15)	0.816
Total N.	1.13 (0.85)	1.42 (0.77)	0.97 (0.45)	1.62 (1.26)	0.580
TSS	42.64 (53.89)	40.92 (54.54)	25.8 (23.54)	30.44 (47.93)	0.820
Load (kg)					
PO_4	8.5 (9.06)	4.05 (4.37)	7.58 (7.42)	2.9 (4.05)	0.059*
Total P.	12.13 (13.63)	5.74 (5.85)	10.92 (10.6)	4.41 (5.45)	0.195
NH_4	5.51 (6.41)	3.33 (2.17)	8.38 (8.07)	3.7 (2.06)	0.171
NO_3	23.29 (19.87)	6.42 (7.99)	12.5 (16.05)	2.79 (4.63)	0.751
Total N.	67.61 (54.71)	22.27 (18.16)	65.69 (47.29)	22.41 (22.37)	0.413
TSS	2351.16 (3712.01)	909.94 (1135.37)	1578.54 (1944.75)	560.84 (888.1)	0.194

groundwater levels and increased ponding within the treatment reach in response the restoration, which is consistent with the hydrological effects reported by Hammersmark et al. (2008) and Cizek et al. (2017).

We also observed that the FDCs of the reference and treatment watersheds diverged at low flows after restoration. However, the reduction of low flows at the outlet of the restored reach did not propagate 609 m downstream to the discharge point of the treatment watershed. Instead, after restoration we observed sustained flow in the treatment watershed relative to the reference (Fig. 2), confirmed by a reduction in the annual Q5:Q95 flashiness index (Fig. 3b). We hypothesize that, while increases in subsurface storage and surface ponding may have limited periods of prolonged flow at the outlet of the restored reach, this elevated groundwater exfiltrated back into the stream beyond the reach restoration, thus sustaining flow in the watershed as a whole. This conclusion is not unreasonable, as Cizek et al. (2017) found that conversion of surface water at the inlet to exfiltrating groundwater (referred to as 'seep-out') can be as much as 84% of the inflowing surface water.

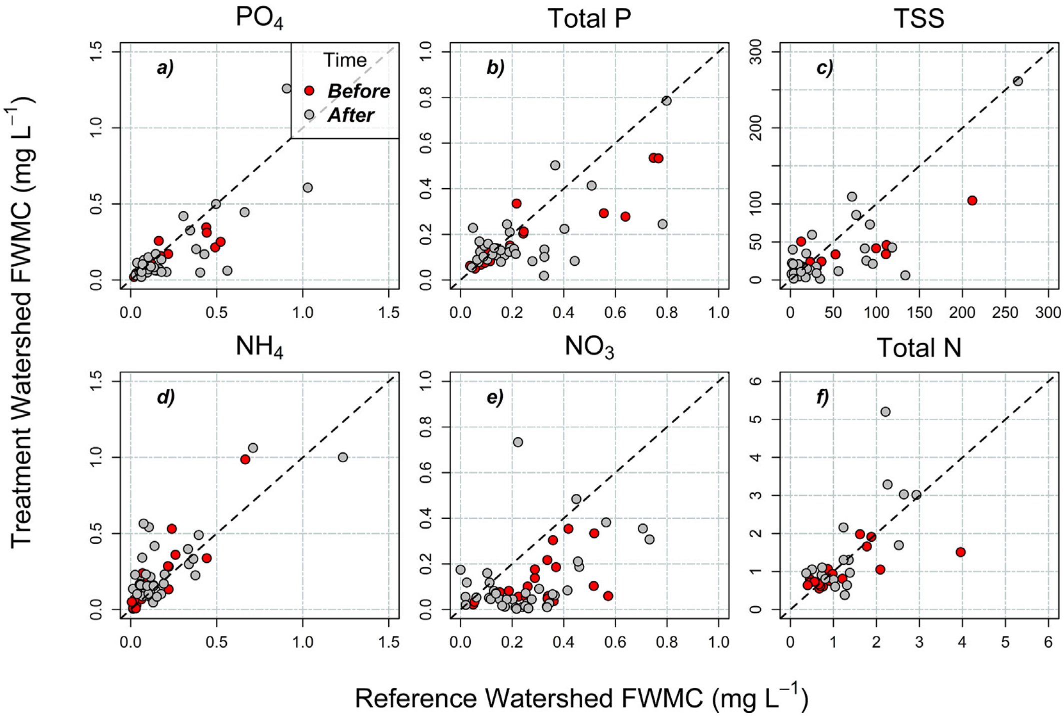


Fig. 7. Relationship between the treatment and reference flow-weighted mean concentrations of a) PO_4, b) Total P, c) TSS, d) NH_4, e) NO_3, and f) Total N before the restoration (red points), and after the restoration (gray points). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.2. The effect of the restoration on net nutrient removal

We quantified changes in FWMCs before and after restoration as inter-annual variances in water flux can have a large effect on nutrient and sediment loads. This is important when considering the effects of stream restoration, which can increase water storage and evapotranspiration, and may reduce loads solely by decreasing the water flux through the reach. Similarly, increases in the water flux can have a dilution effect on FWMCs, particularly when groundwater emerging in the restored reach has lower nutrient concentrations than the stream water. In our restoration we found that except for NH_4, FWMCs were reduced since the restoration. Our analysis of the effective

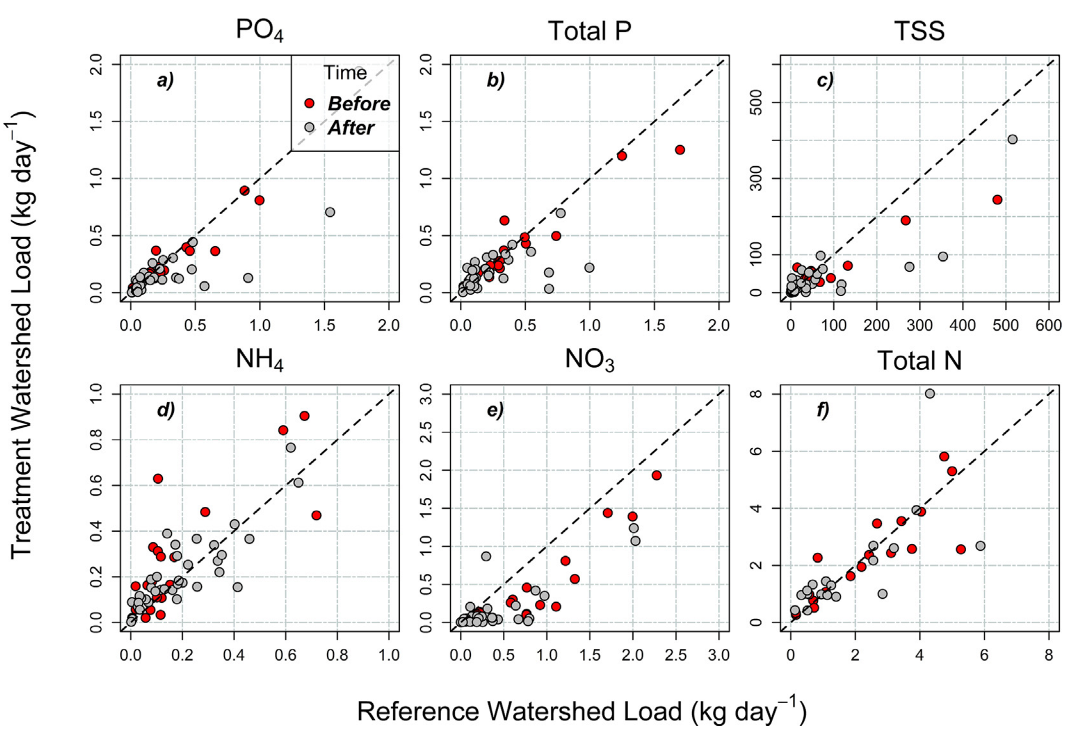


Fig. 8. Relationship between the treatment and reference daily loads of a) PO_4, b) Total P, c) TSS, d) NH_4, e) NO_3, and f) Total N before the restoration (red points), and after the restoration (gray points). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

change in water yields between the paired monitoring stations indicated minor changes in the differences before and after restoration, giving us confidence that our observed FWMC changes were mainly driven by functional changes rather than hydrological ones.

Recent studies of the water quality benefits of stream restoration have shown variable success. This is unsurprising considering the multitude of factors required for enhanced nutrient retention (Newcomer Johnson et al., 2016), and the low likelihood that all stream restorations and water management structures will improve all these necessary attributes. For example, Mueller Price et al. (2016) found inconsistent enhanced NO_3 uptake between restored and unrestored streams in Northern Colorado, USA. Likewise, Filoso and Palmer (2011) found variable effects of net N removal in two lowland restorations in Maryland, USA, with average N retention during storm flow fluxes found to vary from 1.16 g N to −0.77 g N per meter stream length per storm.

In our study, we did not observe a statistically significant change in the removal of NO_3 or TSS within the restored reach before and after restoration at alpha level of 0.1 (Table 1). However, a mass balance between the inlet and outlet for the periods before and after restoration indicated removal rates from −18.3% to 73.8% for TSS and from 14.3% to 25.7% for NO_3, before and after the restoration respectively. We attribute this lack of significance in weekly loads observed before and after restoration to variable retention rates before and after restoration. Using RIA to detect significant changes in annual loads, rather than weekly loads, before and after restoration might indicate statistically significant reductions for all parameters, because weekly variation would be removed. Unfortunately, our data were not of sufficient temporal length to perform RIA on annual loads.

The removal rates of nutrients and sediment observed in our study were high compared to the results of many other studies (Williams et al., 2017, Filoso et al., 2015, Filoso and Palmer, 2011). Although our use of v-notch weirs to measure stream discharge may have influenced nutrient and sediment retention by slowing the flow at the outlet of the restoration, particularly for TSS, PO_4 and TP, we do not consider this to be an important factor, as our v-notch weirs would have had a similar effect to the rock weirs installed as part of the restoration. The distance between the v-notch weir and the next upstream flow control was 38 m, which represents only 8.4% of the restored reach. The high retention rates in our study compared to others may be due to the relative lack of impervious surface in our watersheds (Williams et al., 2017). Many studies which document mixed success or failure are often located in heavily urbanized watersheds with limited potential for infiltration, where the ability for stream restoration to treat large pulses of storm flow is limited (Filoso et al., 2015, Filoso and Palmer, 2011). For example, a recent study of three restorations by Williams et al. (2017) in a watershed with a moderate percentage of impervious cover (45%) compared to other urban watersheds showed average percentage reductions of PO_4, TP, NH_4, NO_3, and TSS to be 7.3%, 35%, 48.7%, 62.5%, and 27.3% per ha, respectively. However, it should be considered that our sampling methods differed from Williams et al. (2017) as they did not use flow-paced sampling or instantaneous acid-preservation upon sampling. Thus, unlike Williams et al. (2017) our measurements of PO_4 and NH_4 retention include the particulate fraction that dissolved upon contact with the acid presevative. Nevertheless, our results were consistent with Williams et al. (2017) and indicate that stream restorations such as the one reported in this study show promise for enhancing the removal of nutrients and sediment.

Interestingly, our data indicate that retention efficiency did not change with the total flow for each sampling interval. It is possible that retention efficiency could increase with increasing flow because of increasing diversion of water onto the floodplain. However, it is also possible that retention efficiency could decrease with total flow due to concomitant increases in particle transport and decreases in residence time in the reach. Indeed, we observed some reductions in removal efficiency with increasing storm sizes for PO_4 and Total P, although our results were constrained to a small range of storm sizes (Fig. S4, *supplementary information*). Thus, we cannot discount increases in stream area as a factor for enhanced removal of Total P and NO_3. This is an important observation as many streams, particularly urban and suburban, have limitations on how much of the floodplain can be offset for stream restoration due to land ownership constraints. Thus, land use may not only dictate the success of stream restoration through hydrological variances, but also through the amount of land available for floodplain connection. We suggest future research quantify removal rates per stream area, as this not only has management applications, but would also help with comparability of mass balance studies of stream restorations, of which there are few, to nutrient spiraling studies, of which there are many (e.g., Mulholland et al. (2009)).

4.3. *The effect of reach scale restorations on downstream loads and implications for monitoring*

Results from RIA suggest that the reach level reductions in FWMC and loads did not propagate 609 m downstream to the discharge point of the treatment watershed. Considering the large retention observed at the reach scale, we found the absence of significant changes in watershed nutrient and sediment loads surprising. However, the area which the outlet of the restoration drains represents only 66% of the treatment watershed, with the other 34% containing a beaver pond; a natural sink of nutrients and sediment (Correll et al., 2000, Hill and Duval, 2009). As reported by Correll et al. (2000), the beaver pond in our treatment watershed has historically been observed to reduce Total N, Total P, and TSS loads by 18%, 21%, and 27%, respectively (although retention efficiency may have increased or decreased since). Thus, it is reasonable to assume that within our study watershed, the beaver pond may have dominated the water quality signal and masked any impact of the stream restoration upstream, with the outflowing nutrients and sediment at the lowest possible concentration both before and after restoration. Future studies should be conducted where treatment reaches are known to be large sources of nutrients and sediment to the wider watershed prior to restoration, to test the extent to which stream restoration can improve water quality at the watershed-scale.

This is an important finding, as many approaches to monitor the effect of stream restorations, particularly in urban watersheds, do not explicitly isolate and evaluate a restoration, but also include other storm water management structures as well as varying sources of nutrients and sediment within the watershed (Park et al., 1994, Aulenbach et al., 2017, Gold et al., 2017).

Our study demonstrates that evaluating load reduction resulting from stream restoration requires a multiscale approach to monitoring. Within our study, solely monitoring the inflow and outflow of the treatment reach before and after restoration would have indicated the striking benefits of restoration for meeting water quality improvements. Conversely, merely monitoring at the larger spatial scales in a BACI design may have indicated the failure of restoration to improve water quality. Considering that many decision-makers rely on the outcomes of studies like ours to guide management of water resources, there is an urgent requirement to understand the explicit performance of stream restoration, whilst also understanding the broader watershed response to localized restoration.

Although the BACI design provides a good basis for testing the response to environmental disturbances such as restoration, the approach has historically been criticized for a lack of replication (Hurlbert, 1984) and for being sensitive to divergent trends between reference and treatment sites, independent of a treatment effect, which can result in a misleading assessment of a treatment effect (Smith et al., 1993). However, we suggest that the ability to replicate stream restoration is both financially impractical and physically impossible. Stream restorations with extensive geomorphological manipulation often cost upwards of one million dollars, and while the broad concepts of restoration design are agreed upon by stream restoration professionals, in practice, many stream restorations are constructed differently to suit each individual site and cannot easily be replicated. Thus, as is often the case with large-scale experiments (Likens et al., 1977, 1970), treatment replication within our BACI design was not possible. However, our multiscale approach may be a possible solution to the problems of replication, whilst also providing information on the aggregate effect of reach-scale restoration.

5. Conclusions

We found that the restoration of a coastal intermittent stream increased removal rates of nutrients and sediment from the water column. Using a BACI monitoring design at the reach scale and at the larger watershed scale, our study has shown that:

- At the reach scale, Flow-Weighted Mean Concentrations (FWMCs) of Total P, NO_3, and Total N were reduced after restoration ($p < 0.05$) as were FWMCs for PO_4 and TSS ($p < 0.1$). FWMCs for NH_4 did not show a significant reduction. Weekly loads for PO_4 and Total P ($p < 0.05$) and for NH_4 and Total N ($p < 0.1$) showed significant reductions.
- Total net retention after restoration increased to 44.8%, 45.8%, 48.3%, 25.7%, 49.7%, and 73.8% for PO_4, Total P, NH_4, NO_3, Total N, and TSS, respectively, when assessed at the reach scale.
- Although large reductions in loads were observed at the reach scale, except for weekly loads of PO_4 ($P < 0.1$), no significant reductions resulting from the restoration were apparent at the watershed scale when comparing the treatment and reference watersheds.
- The lack of significant load reduction has important implications for how stream restorations are monitored and highlights the need to put reach scale restorations in the perspective of wider watershed processes.

This study revealed that although stream restoration can enhance water quality by removing nutrients from the water column, these changes were not of a magnitude large enough to allow detectable changes due to the restoration 0.6 km farther downstream. It is likely that in a watershed with other nutrient removal systems, such as the beaver ponds in our study, much larger nutrient and sediment reductions in restored reaches are necessary to observe watershed-scale reductions in nutrient and sediment export. This has implications for future monitoring designs of other studies, and we suggest that a multiscale approach be used to quantify both the local and wider effects of stream restoration.

Acknowledgements

We thank Donald Weller at the Smithsonian Environmental Research Center and Jim Hood at Ohio State University for helpful discussion of statistical analysis and Randomized Intervention Analysis. We thank Max Ruehrmund and interns Julianne Rolf, Brendan Player, and Lauren Mosesso for field work assistance who were funded by the National Science Foundation Research Experience for Undergraduates Program. Funding for research was provided by award from the Chesapeake Bay Trust (award 13138), Maryland Department of Natural Resources (award 14-13-1717 TRF 09), Chesapeake Rivers Association, Inc, and Underwood and Associates, LLC. Declarations of interest: none.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecoleng.2018.09.016.

References

American Public Health Association, 2005. Standard Methods for the Examination of Water and Wastewater. American Public Health Association (APHA), Washington, DC, USA.
Aulenbach, B.T., Landers, M.N., Musser, J.W., Painter, J.A., 2017. Effects of impervious area and BMP implementation and design on storm runoff and water quality in eight small watersheds. JAWRA J. Am. Water Res. Assoc. 53 (2), 382–399.
Bernhardt, E.S., Palmer, M.A., Allan, J., Alexander, G., Barnas, K., Brooks, S., Carr, J., Clayton, S., Dahm, C., Follstad-Shah, J., 2005. Synthesizing US river restoration efforts. Science 308 (5722), 636–637.
Brown, T., Berg, J., Underwood, K. (2010) Low Impact Development 2010: Redefining Water in the City, pp. 207–217.
Carpenter, S.R., Frost, T.M., Heisey, D., Kratz, T.K., 1989. Randomized intervention analysis and the interpretation of whole-ecosystem experiments. Ecology 70 (4), 1142–1152.
Cizek, A.R., Hunt, W.F., Winston, R.J., Lauffer, M.S., 2017. Hydrologic performance of regenerative stormwater conveyance in the North Carolina Coastal Plain. J. Environ. Eng. 143 (9), 05017003.
Cole, J.J., Prairie, Y.T., Caraco, N.F., McDowell, W.H., Tranvik, L.J., Striegl, R.G., Duarte, C.M., Kortelainen, P., Downing, J.A., Middelburg, J.J., Melack, J., 2007. Plumbing the global carbon cycle: integrating inland waters into the terrestrial carbon budget. Ecosystems 10 (1), 172–185.
Correll, D.L., Jordan, T.E., Weller, D.E., 1992. Nutrient flux in a landscape: effects of coastal land use and terrestrial community mosaic on nutrient transport to coastal waters. Estuaries Coasts 15 (4), 431–442.
Correll, D.L., Jordan, T.E., Weller, D.E., 2000. Beaver pond biogeochemical effects in the Maryland Coastal Plain. Biogeochemistry 49 (3), 217–239.
Dietz, M.E., Clausen, J.C., 2008. Stormwater runoff and export changes with development in a traditional and low impact subdivision. J. Environ. Manage. 87 (4), 560–566.
Dunne, T., Leopold, L.B., 1978. Water in Environmental Planning. Macmillan.
Galster, J.C., Pazzaglia, F.J., Hargreaves, B.R., Morris, D.P., Peters, S.C., Weisman, R.N., 2006. Effects of urbanization on watershed hydrology: The scaling of discharge with drainage area. Geology 34 (9), 713–716.
Filoso, S., Smith, S.M., Williams, M.R., Palmer, M.A., 2015. The efficacy of constructed stream–wetland complexes at reducing the flux of suspended solids to Chesapeake Bay. Environ. Sci. Technol. 49 (15), 8986–8994.
Filoso, S., Palmer, M.A., 2011. Assessing stream restoration effectiveness at reducing nitrogen export to downstream waters. Ecol. Appl. 21 (6), 1989–2006.
Geist, J., 2015. Seven steps towards improving freshwater conservation. Aquat. Conserv. Mar. Freshwater Ecosyst. 25 (4), 447–453.
Geist, J., Hawkins, S.J., 2016. Habitat recovery and restoration in aquatic ecosystems: current progress and future challenges. Aquat. Conserv. Mar. Freshwater Ecosyst. 26 (5), 942–962.
Gibson, C.A., O'Reilly, C.M., Conine, A.L., Lipshutz, S.M., 2015. Nutrient uptake dynamics across a gradient of nutrient concentrations and ratios at the landscape scale. J. Geophys. Res. Biogeosci. 120 (2), 326–340.
Gold, A., Thompson, S., Piehler, M., 2017. Water quality before and after watershed-scale implementation of stormwater wet ponds in the coastal plain. Ecol. Eng. 105, 240–251.
Gomi, T., Sidle, R.C., Richardson, J.S., 2002. Understanding processes and downstream linkages of headwater systems: headwaters differ from downstream reaches by their close coupling to hillslope processes, more temporal and spatial variation, and their need for different means of protection from land use. Bioscience 52 (10), 905–916.
Hammersmark, C.T., Rains, M.C., Mount, J.F., 2008. Quantifying the hydrological effects of stream restoration in a montane meadow, northern California, USA. River Res. Appl. 24 (6), 735–753.
Harding, L., Gallegos, C., Perry, E., Miller, W., Adolf, J., Mallonee, M., Paerl, H., 2016. Long-term trends of nutrients and phytoplankton in Chesapeake Bay. Estuaries Coasts 39 (3), 664–681.
Hill, A.R., Duval, T.P., 2009. Beaver dams along an agricultural stream in southern Ontario, Canada: their impact on riparian zone hydrology and nitrogen chemistry. Hydrol. Processes 23 (9), 1324–1336.
Hood, J.M., Benstead, J.P., Cross, W.F., Huryn, A.D., Johnson, P.W., Gíslason, G.M., Junker, J.R., Nelson, D., Ólafsson, J.S., Tran, C., 2018. Increased resource use efficiency amplifies positive response of aquatic primary production to experimental warming. Global Change Biol. 24 (3), 1069–1084.
Hooper, R.P., Shoemaker, C.A., 1986. A comparison of chemical and isotopic hydrograph separation. Water Resour. Res. 22 (10), 1444–1454.
Hurlbert, S.H., 1984. Pseudoreplication and the design of ecological field experiments. Ecol. Monogr. 54 (2), 187–211.
Jordan, T.E., Correll, D.L., Weller, D.E., 1997. Relating nutrient discharges from watersheds to land use and streamflow variability. Water Resour. Res. 33 (11), 2579–2590.
Jordan, T.E., Correll, D.L., Miklas, J., Weller, D.E., 1991. Long-term trends in estuarine nutrients and chlorophyll, and short-term effects of variation in watershed discharge. Mar. Ecol. Prog. Ser. 121–132.
King, E.J., 1932. The colorimetric determination of phosphorus. Biochem. J 26 (2), 292.
Lyne, V., Hollick, M., 1979. Stochastic time-variable rainfall-runoff modelling, pp. 89–93.
Likens, G., Bormann, F., Pierce, R., Eaton, J., Johnson, N., 1977. Biogeochemistry: Of a Forested Ecosystem. Springer-Verlag.
Likens, G.E., Bormann, F.H., Johnson, N.M., Fisher, D., Pierce, R.S., 1970. Effects of forest cutting and herbicide treatment on nutrient budgets in the Hubbard Brook watershed-ecosystem. Ecol. Monogr. 40 (1), 23–47.
Martin, D.F., 1972. Marine Chemistry, vol. 1 Marcel Dekker, New York.
Mueller Price, J., Baker, D., Bledsoe, B., 2016. Effects of passive and structural stream restoration approaches on transient storage and nitrate uptake. River Res. Appl. 32 (7), 1542–1554.
Mulholland, P.J., Hall, R.O., Sobota, D.J., Dodds, W.K., Findlay, S.E., Grimm, N.B., Hamilton, S.K., McDowell, W.H., O'Brien, J.M., Tank, J.L., Ashkenas, L.R., 2009.

Nitrate removal in stream ecosystems measured by ^{15}N addition experiments: denitrification. Limnol. Oceanogr. 54 (3), 666–680.
Newcomer Johnson, T.A., Kaushal, S.S., Mayer, P.M., Smith, R.M., Sivirichi, G.M., 2016. Nutrient retention in restored streams and rivers: a global review and synthesis. Water 8 (4), 116.
Palmer, M.A., Filoso, S., Fanelli, R.M., 2014. From ecosystems to ecosystem services: stream restoration as ecological engineering. Ecol. Eng. 65, 62–70.
Pander, J., Geist, J., 2013. Ecological indicators for stream restoration success. Ecol. Indic. 30, 106–118.
Park, S.W., Mostaghimi, S., Cooke, R.A., McClellan, P.W., 1994. BMP impacts on watershed runoff, sediment, and nutrient yields. JAWRA J. Am. Water Resour. Assoc. 30 (6), 1011–1023.
Pennino, M.J., McDonald, R.I., Jaffe, P.R., 2016. Watershed-scale impacts of stormwater green infrastructure on hydrology, nutrient fluxes, and combined sewer overflows in the mid-Atlantic region. Sci. Total Environ. 565, 1044–1053.
R Development Core Team, 2015. R: A Language and Environment for Statistical Computing, Vienna, Austria.
Shuster, W.D., Bonta, J., Thurston, H., Warnemuende, E., Smith, D., 2005. Impacts of impervious surface on watershed hydrology: a review. Urban Water J. 2 (4), 263–275.
Smith, E.P., Orvos, D.R., Cairns Jr, J., 1993. Impact assessment using the before-after-control-impact (BACI) model: concerns and comments. Can. J. Fish. Aquat. Sci. 50 (3), 627–637.
Stewart-Oaten, A., Bence, J.R., 2001. Temporal and spatial variation in environmental impact assessment. Ecol. Monogr. 71 (2), 305–339.
Strickland, J., Parsons, T., 1977. Fisheries Research Board Bulletin. Fisheries Research Board.
Ukonmaanaho, L., Starr, M., Kantola, M., Laurén, A., Piispanen, J., Pietilä, H., Perämäki, P., Merilä, P., Fritze, H., Tuomivirta, T., Heikkinen, J., 2016. Impacts of forest harvesting on mobilization of Hg and MeHg in drained peatland forests on black schist or felsic bedrock. Environ. Monit. Assess. 188 (4), 228.
Wallace, J.B., Eggert, S.L., Meyer, J.L., Webster, J.R., 2015. Stream invertebrate productivity linked to forest subsidies: 37 stream-years of reference and experimental data. Ecology 96 (5), 1213–1228.
Walsh, C.J., Roy, A.H., Feminella, J.W., Cottingham, P.D., Groffman, P.M., Morgan II, R.P., 2005a. The urban stream syndrome: current knowledge and the search for a cure. J. North Am. Benthol. Soc. 24 (3), 706–723.
Walsh, C.J., Fletcher, T.D., Ladson, A.R., 2005b. Stream restoration in urban catchments through redesigning stormwater systems: looking to the catchment to save the stream. J. North Am. Benthol. Soc. 24 (3), 690–705.
Weber, T., Sloan, A., Wolf, J., 2006. Maryland's Green Infrastructure Assessment: development of a comprehensive approach to land conservation. Landscape Urban Plann. 77 (1), 94–110.
Williams, M.R., Bhatt, G., Filoso, S., Yactayo, G., 2017. Stream restoration performance and its contribution to the Chesapeake Bay TMDL: challenges posed by climate change in urban areas. Estuaries Coasts 1–20.