

Effects of stream restoration by legacy sediment removal and floodplain reconnection on water quality and riparian vegetation

Vanessa B. Beauchamp and Joel Moore
Patrick McMahon, Patrick Baltzer, Ryan A. Casey, Christopher J. Salice, Kyle Bucher, and
Melinda Marsh

Towson University
vbeauchamp@towson.edu
moore@towson.edu

Final Report to the Chesapeake Bay Trust
Award #13974

December 2020

Executive Summary

Stream restoration is a common practice to reduce sediment and nutrient export. However, the effectiveness of many stream restorations in improving water quality is unmeasured and restoration approaches continue to be vigorously discussed. In the Mid-Atlantic region of the United States, activity by European settlers resulted in upland erosion and deposition of sediments 1–3 m in thickness in stream valleys. Subsequently, streams incised those legacy sediments creating steep, exposed banks, infrequent floodplain inundation, and water tables disconnected from floodplains. Legacy sediment removal (LSR) and floodplain reconnection (FR) proposes water quality improvement by restoration to a hydrological state closer to the pre-European. Reconnection of incised streams to their floodplains should result in decreased fluxes of nitrogen, phosphorus and sediment and in lower, wetter floodplains dominated by native hydrophytic plant species. We investigated water quality at nine sites, six restored with LSR/FR and three comparison sites, and surveyed riparian vegetation at the same six restored sites to determine how this method of stream restoration affects water chemistry and riparian plant communities.

Nitrogen baseflow concentrations and fluxes were elevated in urban and particularly agricultural watersheds with little apparent impact by restoration. Denitrification appeared to be constrained by carbon limitations. Ion concentrations were elevated in all watersheds compared to a forested reference, are not addressed by restoration, and represent a substantial ecological stressor. Storm event data from one site suggested small reductions in nutrient and sediment loads across the restored reach. High-frequency time series indicate that restoration effects are not observable at larger scales. The effects of restoration, particularly for denitrification, may not be observable for years and can be obscured by weather and climate-driven variability.

One major consequence of restoration on riparian vegetation was an 80% reduction in woody basal area. Loss of this carbon source may impede in-stream and floodplain denitrification. Woody plant communities became more hydrophytic, but species richness and diversity decreased with restoration. Woody beta diversity among sites decreased, indicating that active revegetation with a common suite of species may result in regional homogenization. Herbaceous beta diversity among sites increased, and many species that were significant indicators of restored sites had not been actively planted, suggesting in-situ seedbanks or upstream inputs contributed to revegetation and maintained a unique identity for each site. Restored herbaceous communities had more hydrophytic vegetation than reference reaches. On average, restored reaches had similar levels of herbaceous richness, diversity and floristic quality to unrestored reaches; however, while sites with lower quality herbaceous vegetation improved with restoration, sites with higher quality vegetation did not improve or saw a decrease in quality after restoration. The practice of legacy sediment removal and floodplain reconnection has the potential to create high quality riparian vegetation communities, but preservation of high-quality forest areas, even if they are atop legacy sediment terraces, should be considered, particularly if losses in tree canopy are not offset by gains in nutrient cycling.

Chesapeake Bay Trust – Restoration Research Award Program Specific Summary

What was/were your key restoration research question(s)?

FY 16 Q 6 Effectiveness of stream restoration to accomplish water quality and habitat goals – Effects of site condition on outcomes of stream restoration technique(s)

Specifically: What is the impact of site condition (such as land use, % impervious cover, watershed condition, existing habitat, and/or valley type) and/or watershed position (headwaters vs. downstream near the receiving waters) on the nutrient, sediment, habitat, and/or biological impacts of stream restoration approaches that aim for different function (e.g., floodplain reconnection, frequency of inundation, bank stabilization, etc.) or that use different techniques (e.g., RSC, NCD, stream valley restoration/legacy sediment removal)?

Our research questions were about the effect of site condition on nutrient and sediment removal and riparian vegetation diversity and quality as a result of LSR/FR projects. In our proposal, the site conditions that we expected might affect nutrient and sediment removal and riparian vegetation diversity and quality were impervious surface cover and restored stream length.

What are the results for your research question(s)?

Land use, including impervious surface cover, and restored stream length did not appear to have any effect on the riparian vegetation. Land use did substantially affect nutrient concentrations and export though agricultural land use and associated nitrogen inputs seemed to be more important than impervious surface cover. Restored stream length did not appear to have any effect on nutrient removal.

The findings of this study raise a number of questions for future research:

- What role does the in-situ seed bank play in the revegetation process?
- How does planting of different woody species (particular non-native willows) affect the successional trajectory of a site and the eventual composition of the forest overstory?
- Do the changes in vegetation community composition that we are seeing persist many years to decades post restoration?
- How long after restoration does it take for substantial denitrification to occur: either in-stream or in the adjacent riparian groundwater to an extent sufficient to influence stream concentrations and fluxes
- Related, measuring groundwater chemistry seems worthwhile. Doing so was initially proposed for this project but was cut to increase the focus on storm sampling.
- Are there ways to do restoration projects in a way that adds carbon to the system?
- Would high-frequency data, while expensive to collect, allow better characterization of storm loads?

Discuss the greatest challenges, including the lessons learned, and potential roadblocks to future progress.

A challenge on the front end was collecting data prior to stream restoration at three of sites, in part because of some administrative issues that delayed the start of data collection. Having the highest precipitation year on record occur during the middle of the study certainly influenced the results. High-frequency data, almost certainly as collected with sensors, seem like an important component of at least some future work. However, purchase, calibration, and maintenance of those sensors is time-intensive and therefore expensive.

Based on the results of the project, how would you refine and improve your project or approach in the future?

One of the strengths of the project was the inclusion of several sites so that the effects of land use and restoration length could be investigated.

What advice would you give to someone considering a similar project?

Try to get funded for a longer period of time so that the pre-restoration dataset is longer. Use high-frequency sensors for sediment transport. Establishing a rating curve for storm event sampling is an immense amount of work so try to leverage sites where discharge is already measured (though those rarely match with sites where restoration is happening). If storm events will be collected at a site(s), make sure they're close to a road or that motorized transportation is readily available. Collecting storm samples at sites that were 0.4–0.6 mi from the nearest paved surface certainly resulted in logistical challenges.

Introduction

Stream restoration has numerous mitigation and remediation goals. In the Chesapeake Bay, regulatory pressures motivate many restoration projects. Regulations focus on decreasing nitrogen, phosphorus, and total suspended sediment (TSS) loads to reduce eutrophication that has occurred in the Chesapeake Bay since the 1950s [2, 3]. Though billions of dollars are spent annually on restoration, the effectiveness of most restoration projects is unmeasured [4, 5]. Additionally, the geomorphological, hydrological, and biogeochemical state(s) to which streams are restored is vigorously discussed among scientists [6, 7].

Before European settlement, many Mid-Atlantic streams may have had multiple anastomosing channels with associated wetland complexes rather than a single, meandering channel [8, 9]. From the 1600s through early 1900s, increased soil erosion due to deforestation and agricultural activity along with widespread construction of impoundments in the Mid-Atlantic Piedmont such as milldams resulted in widespread deposition of fine-grained sediment in stream valleys [9, 10]. Over the last century, dam breaches and subsequent stream incision into these legacy sediments created streams with steep, exposed banks, often 1–3 m tall.

Stream entrenchment is associated with changes to riparian community composition and reduction of many critical ecosystem services performed by floodplains. Entrenchment creates elevated floodplains, resulting in lower water tables and decreased overbank flooding [11, 12], that disconnect the root systems of vegetation from the water table. In many areas, hydrophytic wetland and riparian plant species are replaced by upland species [13, 14]. Deep groundwater tables are associated with lower denitrification rates, decreasing the ability of these impaired systems to function as landscape-level nitrogen sinks [11, 15-18]. In addition to reductions in ecosystem services, legacy sediments are potentially a large source of suspended sediment [19-23], nitrate [17] and phosphorus entering streams and ultimately the Chesapeake Bay [24, 25].

Given this history, the restoration approach of legacy sediment removal (LSR) was developed to facilitate floodplain reconnection (FR). LSR restoration aims to lower the floodplain and expose buried relic wetland soils through a combination of legacy sediment removal and raising of stream bed elevation. Restored streams are expected to inundate floodplains more frequently, and water tables adjacent to these streams should be closer to the floodplain surface. Denitrification and sediment deposition may result from more frequent inundation. LSR/FR also should reduce bank erosion, a major sediment source in the Mid-Atlantic [19, 23].

Additionally, LSR restoration projects present an opportunity to reestablish hydrophytic species in floodplains and create diverse, high quality riparian and wetland plant communities [26, 27]. Conversely, priority effects of revegetating many project sites with the same few species [28-30], coupled with elimination of a site's unique seedbank through sediment removal, could cause regional homogenization of riparian habitats. Since riparian ecosystems are disturbance-adapted systems that receive seeds from over a wide area [31], they are particularly susceptible to colonization by invasive plant species [32-38]. This susceptibility may be increased by the extreme disturbance involved with floodplain restoration [39-42].

This project surveyed vegetation at six legacy sediment removal and floodplain reconnection sites near Baltimore, Maryland with the objective of characterizing changes to the vegetation community and as a result of restoration (Figure 1). Water quality data (including

nitrogen and phosphorus and TSS) were collected at the same sites, two regional reference sites, and an additional site used to investigate detectability of restoration effects at a larger scale. Land use at the restored sites spans a gradient from active agriculture to urban, and restored channel length varies by ~4.3 times. We investigated baseflow water quality across all sites, stormflow at one restored site, and stormflow plus fluxes at the site focused on scaling.

If legacy sediment removal and floodplain reconnection is a successful restoration approach, we anticipated that fluxes of nitrogen, phosphorus and TSS would decrease in the restored reaches. We also anticipated that a shallower water table depth and increase in the frequency of overbank flows would result in a shift from upland and facultative to more hydrophytic plant species [27, 43]. If these projects result in biotic homogenization and invasion by exotic species [28, 30, 44], we expected to see decreases in alpha and beta diversity and an increase in the importance of exotic and invasive plant species in restored reaches when compared to unrestored reaches.

Methods

Site characteristics

Six restoration sites with watershed areas of 0.33–8.18 km² were studied (Figure 1, Table 1). All sites were located in the Piedmont physiographic province, Maryland, USA and underlain by silica-rich, metamorphic bedrock. Watershed areas and land use characteristics were determined using USGS StreamStats with land use data primarily coming from the 2011 National Land Cover Database [45]. Six restored sites were studied: three restorations were completed 1–3 years before the beginning of water quality sampling and three restorations occurred during the study period with samples collected pre- and post-restoration (Table 1). All sites were restored by Ecotone, Inc. Restoration involved removing 1–2.5 meters of sediment to substantially lower the floodplain, which at some sites involved removal of mature riparian forest 10s of meters in width (more detailed site descriptions in Supplementary Information including land use maps and pictures, Figure S1, S2). Following removal of sediment, restored sites were planted with grasses and small trees. Three restored sites have substantial active cultivated crop agriculture in the watersheds (FMB, CBR, and BTR, Table 1) with active agriculture within 10–100 m of the restored reaches at two sites (FMB and CBR). A fourth restored site (NSR) is primarily active pasture with a small riparian buffer. The fifth restored site (BCB) is currently suburban with ~9% impervious surface cover (ISC) and active, ongoing development; previous land use at the restoration site was agricultural, and some agriculture remains in the upper reaches of the watershed. The sixth restored site (PTR) is highly urban (49% ISC).

Three additional sites for water quality monitoring were located at US Geological Survey (USGS) gaging stations. Two served as regional reference sites: a forested watershed with no ISC or agriculture (POBR), which is a sub-watershed of a low-density suburban watershed containing ~72% forest cover (BARN). The third (PTRG) was used to investigate the detectability of restoration effects at a larger scale; the highly urban restored site PTR comprised a drainage area that was 14.8% of the PTRG watershed and was located ~4.4 km upstream of PTRG.

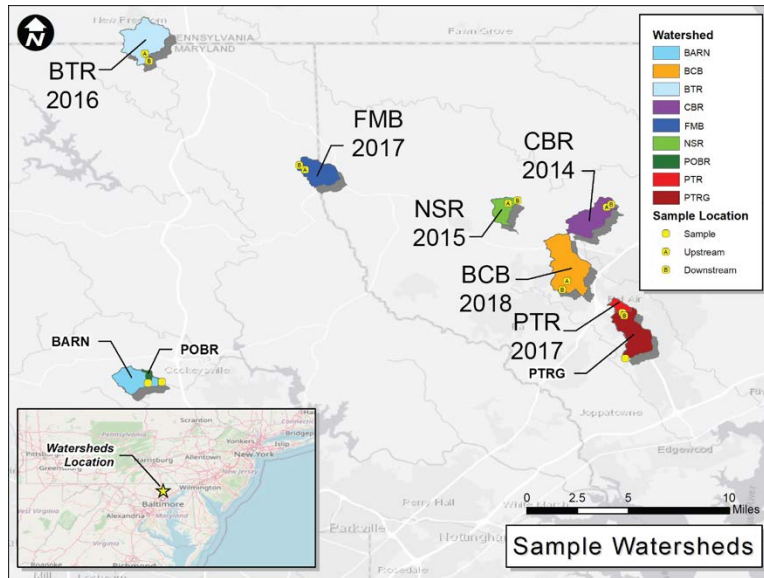


Figure 1: Map of the six restored watersheds with the upstream and downstream locations indicated (BTR = Beetree Run, FMR = First Mine Run, RIG = Rigdon/North Stirrup Run, EDW = Edwards/Cabbage Run, BCB = Bear Cabin Branch, PTR = Plumtree Run). The two regional reference sites at USGS gages that are forested and mostly forested (POBR = Pond Branch, BARN = Baisman Run) site at a USGS gage is downstream of the highly urban PTR site (PTRG = Plumtree Run Gage).

Table 1: Watershed characteristics (restored sites sorted from most to least agriculture)

Site name	Latitude	Longitude	Restoration length & completion date	Watershed area	Developed	Imperv. Surface Cover	Forest	Agricultural (est.) ^a
			<i>m & yr</i>	<i>sq km</i>	%	%	%	%
<i>Restored sites</i>								
First Mine Branch (FMB)			732					
Up (FMBU)	39.63037	-76.57542	Jun. 2017	2.93	4.25	0.53	22.7	73.1
Down (FMBD)	39.63423	-76.57990		3.88	4.86	0.47	26.4	68.7
Cabbage Run (CBR)			408					
Up (CBRU)	39.60432	-76.36004	2014	4.40	31.5	2.19	10.7	57.8
Down (CBRD)	39.60579	-76.35730		4.97	28.1	1.97	11.0	60.9
Beetree Run (BTR)			1621					
Up (BTRU)	39.71404	-76.69006	2016	6.03	33.0	4.24	14.3	52.7
Down (BTRD)	39.70938	-76.68813		6.55	30.4	3.90	14.5	55.1
North Stirrup Run (NSR)			792					
Up (NSRU)	39.60696	-76.43075	2015	1.83	7.44	0.78	43.8	48.8
Down (NSRD)	39.60967	-76.42274		2.25	7.00	0.65	37.7	55.3
Bear Cabin Branch (BCB)			1120					
Up (BCBU)	39.55047	-76.38959	Apr. 2018	7.07	52.5	10.2	21.9	25.6
Down (BCBD)	39.54515	-76.39227		8.18	49.4	8.96	21.6	29.0
Plumtree Run (PTR)			378					
Up (PTRU)	39.52910	-76.34932	Sep. 2017	0.88	97.1	48.8	3.29	0.00
Down (PTRD)	39.52687	-76.34811		0.96	95.5	48.6	5.01	0.00
<i>Reference sites</i> ^b								
Pond Branch (POBR)	39.48029	-76.68755		0.33	0.00	0.00	100	0.00
Baisman Run (BARN)	39.47958	-76.67798		3.81	25.2	1.09	71.7	3.10
Plum. Run Gage (PTRG)	39.49624	-76.34745		6.48	82.3	27.0	14.3	3.40

^a Agricultural is estimated as non-developed and non-forest. If 100% - developed - forested was <0, it was set to 0. Land cover data cover data are from the USGS StreamStats site with Developed and Impervious surface cover based on the 2011 National Land Cover Database.

^b Note the watershed areas used are from the USGS gaging station web pages. USGS gaging station numbers: POBR = 01583570, BARN = 01583580, Plumtree Run Gage = 01581752

Water Chemistry - Discharge & sample collection regime

At the six restored sites, area-velocity (Ott MF Pro) or salt dilution methods were used to measure discharge when baseflow water samples were collected. At FMB, stage-discharge data were measured to establish a rating curve, which combined with barometric and water pressure data collected every five minutes to create a high-frequency discharge record from April 2019 – May 2020. At the two regional reference sites (POBR, BARN) and the site downstream of the highly-urban restored site (PTRG), USGS discharge data were used.

At the three sites restored prior to this study (Table 1), baseflow samples were collected approximately monthly from fall 2017 to summer 2019 with sampling increased to every two weeks during summer 2018. For the three sites restored during this study, baseflow samples were collected pre- and post-restoration. Baseflow samples were collected every two weeks for several months preceding restoration. Baseflow sampling ceased during restoration, resumed every two weeks for at least three months following completion of restoration, and then continued on a monthly basis. At the six restored sites, samples were collected upstream and downstream of the restored reach. Samples were collected during several storm events at FMB in summer of 2019 and winter of 2020.

We also collected baseflow samples at the regional reference sites. Discrete and high-frequency USGS data from PTRG were used to study storm events and calculate daily and cumulative annual loads for nitrate (NO_3^-), total dissolved phosphorus (TDP), and TSS for 2014–2019. Regression models based on discrete samples were used to calculate high-frequency time series (for more details see Supplementary Information).

Water Chemistry - Measurements and analyses

Temperature, pH, specific conductance (SC), and dissolved oxygen (DO) were measured *in situ* using a handheld meter calibrated <24 hours previously. Baseflow samples were field-filtered using a pump and Geotech 0.45 μm Dispos-A-Pore filters. Storm samples were collected with an Isco autosampler, retrieved shortly after storms ended, and filtered in the laboratory using syringe filters (0.45 μM). Samples collected for analysis of alkalinity and of TDN and dissolved organic carbon (DOC) were collected with no headspace. Most samples were stored at 4°C until analysis with some samples frozen until shortly before analysis for N and P speciation.

TSS measurements followed Standard Methods [46]. Alkalinity is reported as HCO_3^- concentrations and was determined by the Gran titration method with titrations performed with a Mettler Toledo G20 with a Rondolino autosampler; data quality were checked with each run using a gravimetrically prepared standard. TDN and DOC (non-purgeable organic carbon) were measured with a Shimadzu TOC-CSN. Major ions and NO_3^- were measured using a Dionex ICS 5000 ion chromatograph with two columns. Silica, $\text{NO}_3^- + \text{NO}_2^-$, NH_4^+ , PO_4^{3-} (orthophosphate), and TDP were measured colorimetrically with a Seal AQ1 Discrete Analyzer. NIST-traceable check standards were used during analytical runs for TDN/DOC, major ions, and nutrients plus silica. If concentrations differed by >10% from expected, then results were double-checked via spot-checking with samples re-run as necessary.

All plotting and statistical tests were done in R [47–52]. Water chemistry data (concentrations and fluxes) were generally not normally distributed. Non-parametric Wilcoxon tests were used to determine if differences were statistically significant ($p < 0.05$) with unpaired

tests run for comparisons between sites or years and paired tests runs for upstream–downstream comparisons at a single site.

Vegetation – Field Sampling

The three sites restored 1–3 years prior to this project each included an unrestored and a restored reach, while the three sites restored in 2017-2018, during the project period, included a project reach sampled pre-restoration and post-restoration and an additional unrestored “control” reach sampled concomitantly with the project reach, providing a Before-After-Control-Impact design. At all sites except Beetree Run, reaches within a site were immediately adjacent. At Beetree Run, the unrestored and restored reaches were separated by 4.8 km. Reaches were sampled with 3 to 18 transects spaced at 25 to 70 m intervals. Transects spanned the floodplain to a maximum of 50 m on each bank. Cover of all herbaceous species, regardless of height, and of woody individuals under 1 m tall was estimated to the nearest 10% in 1 m² plots along each transect at regular intervals (5-10 m depending on transect length). Herbaceous layer vegetation was sampled in late spring and early fall to capture seasonal changes. Due to the timing of construction, fall sampling of the pre-restoration reaches at First Mine Branch and Plumtree Run was conducted in 2016. All other reaches were sampled in spring and fall 2017 and 2018. Unknown species were identified using the Flora of Virginia [53].

Woody vegetation >1 m tall was quantified in 400 m² plots (20 x 20 m) placed randomly along each transect on each side of the stream wherever possible. Plot dimensions and area were modified when needed due to topography. All individuals >1 m tall and >0.5 cm stem diameter at base were identified placed into 0.5-2.0 cm or 2.1-5.0 cm diameter size classes. Diameter at breast height was measured on individuals larger than 5.0 cm diameter. Woody vegetation was sampled once on unrestored and restored reaches and before and after restoration on project reaches. Woody vegetation was also sampled once on control reaches to allow for future comparisons of community change, but those data are not reported for this study.

Vegetation - Community profiles

For each year of herbaceous layer sampling, the two seasonal data sets were combined using the maximum cover values for species recorded in both seasons. Plant communities in each reach were described by functional group, wetland indicator status, conservatism coefficient, and nativity. Functional group and wetland indicator status were assigned according to the classifications in the USDA PLANTS database [1]. Plot level cover-weighted wetland indicator scores (WIS) were calculated by multiplying by the relative cover of each species in each plot by the numeric WIS for that species and summing all the scores in the plot [54]. Plot scores were averaged to obtain a reach-level WIS. Coefficients of conservatism (CC) for all native species [55] were taken from the Mid-Atlantic Piedmont database within The Universal Floristic Quality Assessment Calculator [56]. When CC values were unavailable for the Mid-Atlantic Piedmont region, the Mid-Atlantic Coastal Plain and Mid-Atlantic Ridge & Valley databases were used as supplements. A Floristic Quality Index (FQI) was calculated for each site using CC values [55].

Since the FQI treats all non-native species equally, we created a companion index sensitive to the abundance of exotic and invasive species. Non-native species were classified as

exotic or invasive by consulting a compilation of regional invasive plant species lists published by the University of Maryland Extension [57], which combines information from 28 regional invasive plant species lists. Invasive species were defined as those present on three separate state-level lists within the UME List or listed under the Maryland Noxious Weed Seed Law. When the Mid-Atlantic National Parks Service list was used to determine invasiveness, only listings rated higher than “low” or “insignificant” were counted. Species nativity scores (native = 3, exotic = 2, invasive = 1) were multiplied by the relative cover of each species in each plot to derive a plot-level weighted nativity score (NS). Plot scores were averaged to generate a nativity score for each reach.

The woody community profiles were calculated in a similar fashion to herbaceous communities with basal area ($\text{cm}^2 \text{ m}^{-2}$) taking the place of species cover for wetland indicator and nativity calculations. Importance values (IV) were also calculated for woody species at each reach by taking the average of relative frequency, density and basal area for each species.

Vegetation - Biodiversity

Rarified richness and Shannon diversity values for each site were calculated for the herbaceous layer using incidence frequency data in iNEXTOnline [58, 59] to account for the different numbers of plots sampled among sites. For the herbaceous layer at restored and unrestored reaches, rarified richness and Shannon diversity values were averaged over two years. For the sites restored during the project period we compared pre-restoration values to post-restoration values in the project reach. Comparisons between years 1 and 2 in the control reaches at these sites illustrate the background level of change between the two years, but these data were not included in the statistical analyses. Within-site beta diversity was calculated by dividing reach-level rarefied richness (gamma diversity) by average plot richness (alpha diversity). Beta diversity (Horn similarity) among sites was calculated using SPADE_R [60, 61].

Biodiversity descriptors for woody vegetation were calculated using the same process as for the herbaceous vegetation, with the exception of rarefied richness and diversity, which could not be calculated as topography preventing us from making the plots a consistent size. Woody richness is the number of woody species encountered in each reach. We used iNEXT to calculate Shannon diversity as the effective number of equally common species, as with the herbaceous layer data.

All dependent variables for each reach and the change in these variables with restoration were checked for normality with the Shapiro-Wilk test and for homogeneity of variances with Levene’s test. Comparisons were conducted between reach type using paired t-tests, except for basal area, herbaceous WIS and woody within-site beta diversity which did not meet the assumptions required for parametric statistics and were compared between reaches using Wilcoxon signed-rank test. All analyses were conducted with SPSS version 25 [62]. Significant differences in beta diversity among pre-restoration/unrestored and post-restoration/restored reaches were determined by using the standard errors calculated from 100 bootstrap replications to derive a t-value and corresponding p value [61]. The relationship between the change in each variable and the pre-restoration or unrestored value for each reach was examined with Pearson correlations. We also used Pearson correlations to determine

if changes in any dependent variables were related to project length or watershed-level ISC, or if change in herbaceous WIS was related to change in basal area.

Vegetation - Community Composition

Non-metric multidimensional scaling (NMDS) ordination was used to visualize differences in plant community composition among reaches and years using cover data for the herbaceous layer and importance values for the woody layer. The herbaceous dataset consisted of 12 reaches (three sites with restored and unrestored reaches and three sites with project and control reaches) sampled in two consecutive years for a total of 24 reaches. The woody dataset consisted of 12 reaches (three sites with restored and unrestored reaches sampled once and three sites with project reaches sampled twice). Herbaceous layer species present in fewer than four reaches throughout the study and woody layer species present in fewer than two reaches were removed from further analysis. Each NMDS analysis was conducted with PC-ORD version 5 using the slow and thorough autopilot setting with a Sorensen (Bray-Curtis) distance measure and a random starting configuration [63, 64].

Differences in community composition between restored and reference reaches were examined by conducting an indicator species analysis in PC-ORD version 5 [64] with the same data used for the NMDS. We took the top 30 ranked herbaceous species and top 10 ranked woody species by indicator value in each reach type and compared these groups by wetland indicator score, nativity, life form and life span proportions between restored and unrestored reaches with Fisher's Exact Test. We then used Monte Carlo tests on the same indicator species data to identify significant herbaceous and woody layer indicators of each reach type.

Results

Baseflow nitrogen concentrations and loads

Baseflow nitrogen concentrations and daily loads were controlled primarily by land use (Figure 2a, b) with restoration showing minimal effects. From the upstream to the downstream ends of the six post-restoration reaches, median baseflow nitrogen concentrations differed by <0.5 mg/L following restoration, and baseflow daily loads decreased significantly along the restored reaches only at the two sites where discharge also decreased. At the three sites with pre- and post-restoration data, baseflow nitrogen concentrations did not significantly change following restoration (Figure 3a, b), and baseflow nitrogen daily loads increased significantly at two sites (Figure 3c, d) due to record high precipitation in 2018. These baseflow results indicate little to no change in rates of denitrification or nitrogen removal by other in-stream processes for 1–5 years following restoration.

Land use was the primary driver of baseflow nitrogen concentrations and daily loads, which were lowest at the forested reference site and highest at the agricultural sites (Figure 2a, b). The sites can be classified into four groups based on nitrogen concentrations. First, the lowest

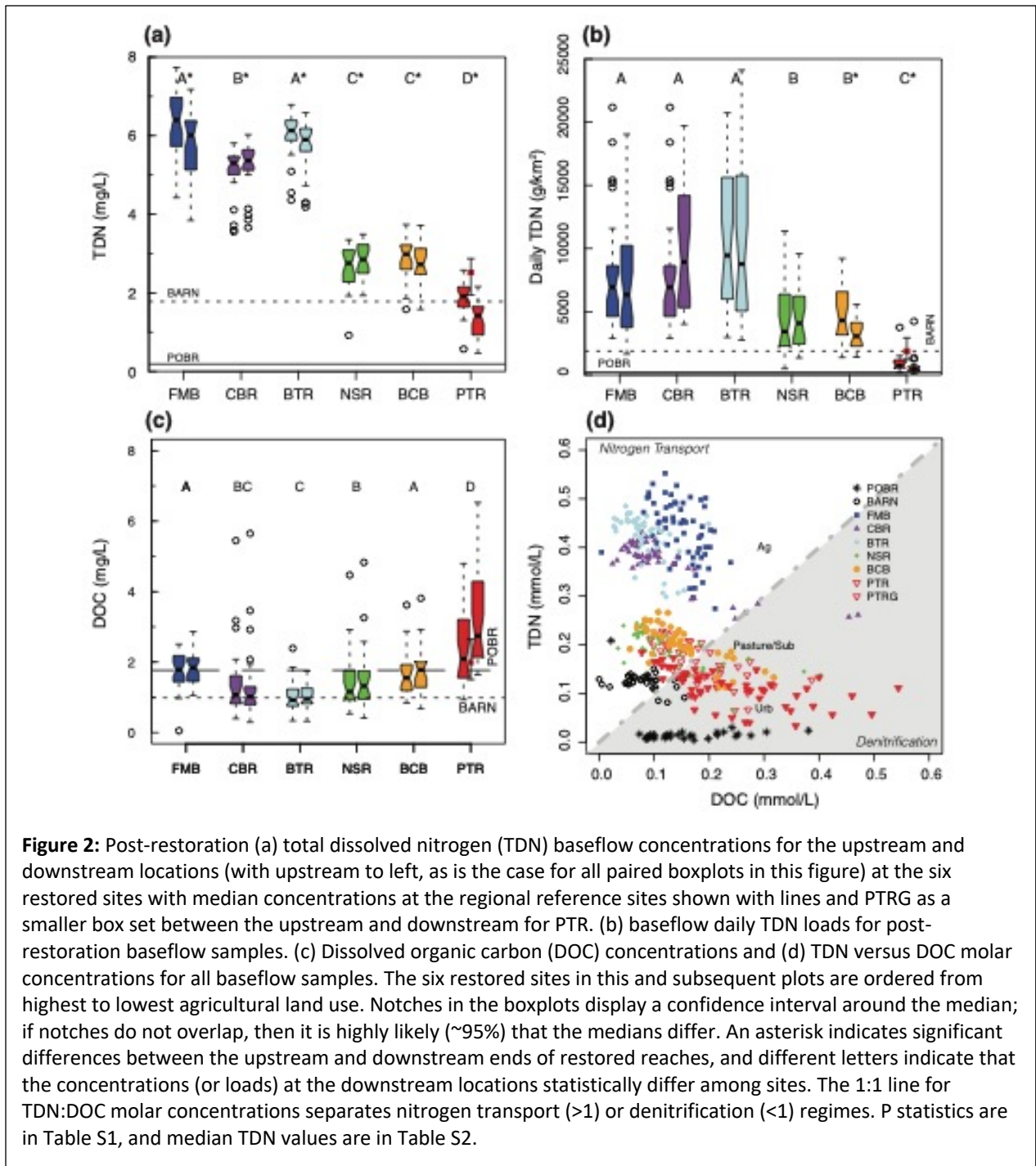
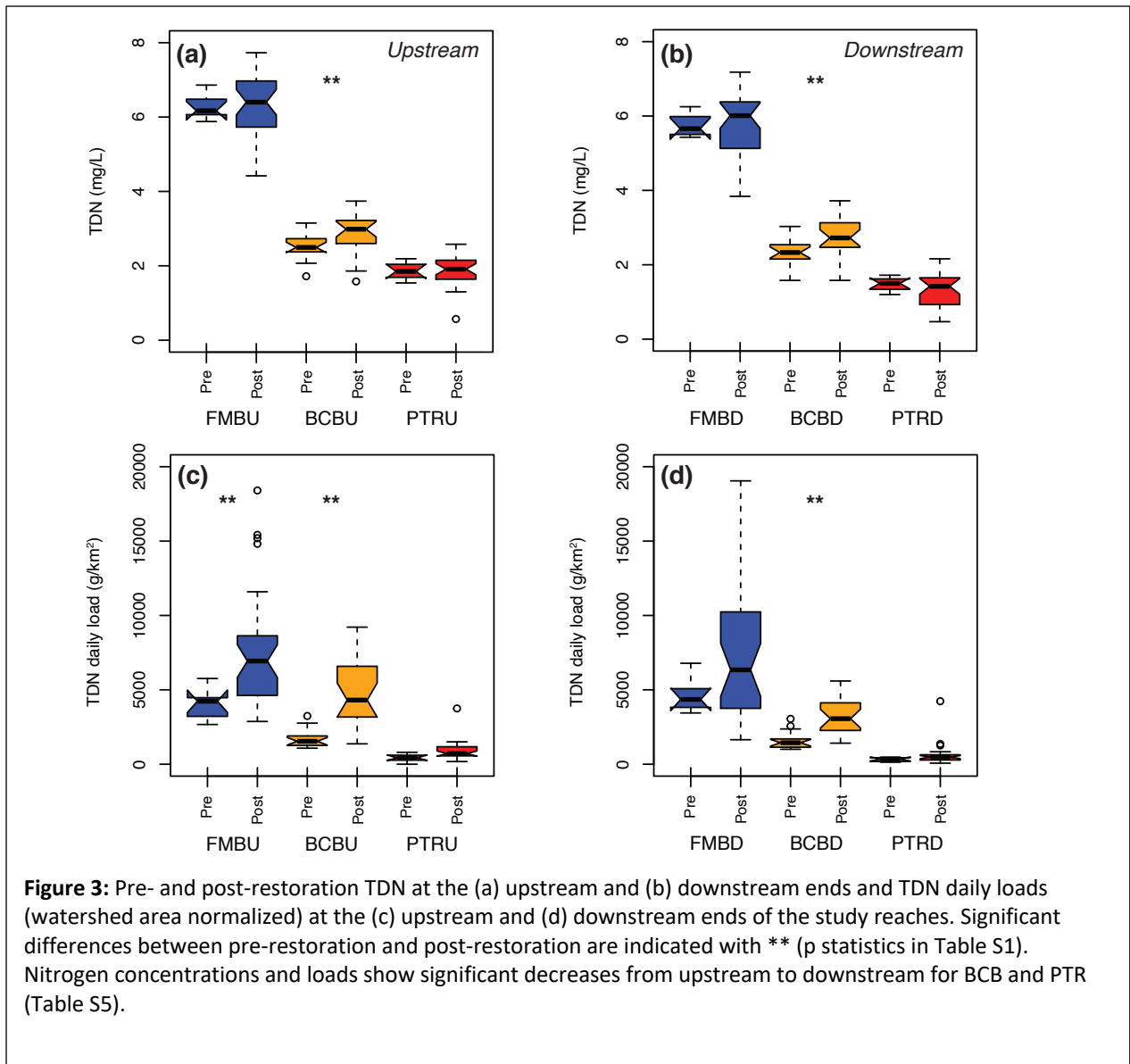


Figure 2: Post-restoration (a) total dissolved nitrogen (TDN) baseflow concentrations for the upstream and downstream locations (with upstream to left, as is the case for all paired boxplots in this figure) at the six restored sites with median concentrations at the regional reference sites shown with lines and PTRG as a smaller box set between the upstream and downstream for PTR. (b) baseflow daily TDN loads for post-restoration baseflow samples. (c) Dissolved organic carbon (DOC) concentrations and (d) TDN versus DOC molar concentrations for all baseflow samples. The six restored sites in this and subsequent plots are ordered from highest to lowest agricultural land use. Notches in the boxplots display a confidence interval around the median; if notches do not overlap, then it is highly likely (~95%) that the medians differ. An asterisk indicates significant differences between the upstream and downstream ends of restored reaches, and different letters indicate that the concentrations (or loads) at the downstream locations statistically differ among sites. The 1:1 line for TDN:DOC molar concentrations separates nitrogen transport (>1) or denitrification (<1) regimes. P statistics are in Table S1, and median TDN values are in Table S2.

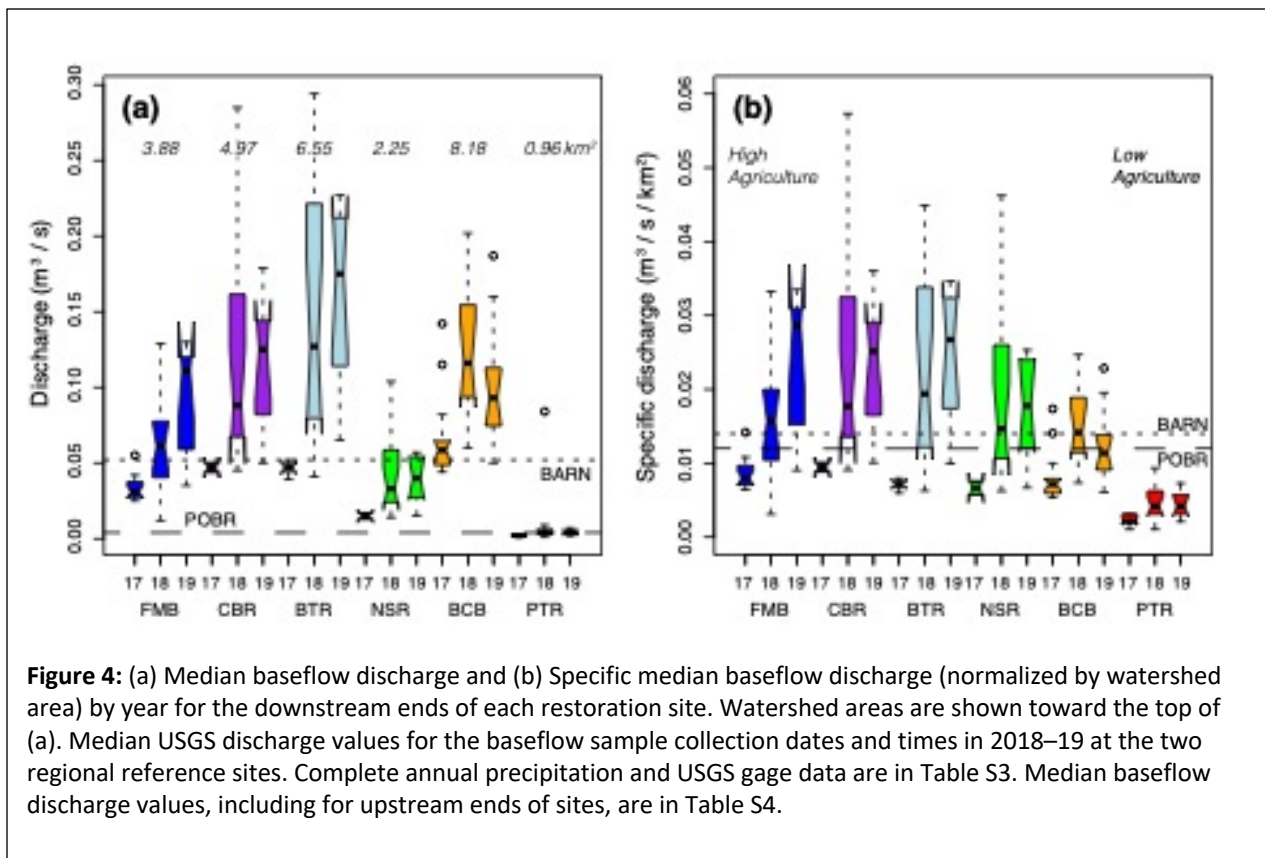
concentrations occurred in the forested watershed (POBR): <0.25 mg/L for TDN (Figure 2a; Table S2). Second, the highly urban (PTR) and mostly forested, suburban reference (BARN) sites had median concentrations of 1.5–2 mg/L with nitrogen likely coming predominantly from sewer (PTR) or septic (BARN) systems and atmospheric deposition. Third, the active pasture (NSR), suburban/partially agricultural (BCB), and USGS urban reference (PTRG) sites had median TDN of 2.5–3.0 mg/L with likely nitrogen sources including agricultural activity such as fertilizer



application (NSR, BCB) and septic (BCB) or sewer (PTRG) systems. Fourth, the three active cultivated crop agriculture sites (FMB, CBR, BTR) had median TDN of 5.4–6.0 mg/L with fertilizer application representing the primary nitrogen source.

Differences in baseflow discharge across the restored sites were minimal in most cases with little effect on nitrogen concentrations or loads. While larger watersheds had higher baseflow discharge (Figure 4a), specific baseflow discharge (normalized to watershed area) was similar across restored watersheds with the exception of lower values at the highly urban PTR watershed (Figure 4b). Baseflow differences did play an important role in controlling baseflow nitrogen loads between the upstream and downstream ends of restored sites and also from 2017 to 2018–19 at the three sites with pre- and post-restoration data (Figure 4; Tables S3, S4).

Baseflow differences between the upstream and downstream ends within sites were a primary driver of nitrogen loads. The four agricultural sites (FMB, CBR, BTR, NSR) had higher baseflow discharge at the downstream ends of restored reaches (though only significantly



higher for FMB, $p < 0.001$; Table S4). For the FMB and BTR sites, higher discharge at the downstream end resulted in similar nitrogen loads at the upstream and downstream ends despite small but significant decreases in nitrogen concentrations across the restored reaches (Figure 2a, b). The CBR and NSR sites also had similar nitrogen daily loads at the upstream and downstream ends due to similar or higher discharge and small but statistically significant increases in nitrogen concentrations from upstream to downstream. At the (sub)urban BCB and PTR sites, discharge decreases from the upstream to downstream ends (Table S4), which was not statistically significant, did contribute to a statistically significant decrease in daily baseflow nitrogen loads (Figure 2b).

The three sites with pre- and post-restoration data strongly indicate that restoration did little to change baseflow nitrogen since upstream to downstream differences, or lack of differences, remained the same post-restoration (Figure 3). Baseflow nitrogen concentrations did not change following restoration at the upstream and downstream ends of the FMB and PTR sites and increased at BCB (Figure 3a, 3b). Post-restoration baseflow nitrogen loads increased compared to pre-restoration loads at both the upstream and downstream ends of BCB and at the upstream end of FMB (Figure 3c, d).

Changes in baseflow nitrogen concentrations and loads along the study reaches at these three sites, *i.e.*, between the upstream and downstream ends, were similar before and after restoration. At all three sites, baseflow nitrogen concentrations were slightly lower (0.12–0.51 mg/L) at the downstream ends both pre- and post-restoration (Figure 3a, b; Table S5). At the agricultural FMB site, nitrogen loads were significantly different between upstream and downstream before but not significantly different restoration. These differences between

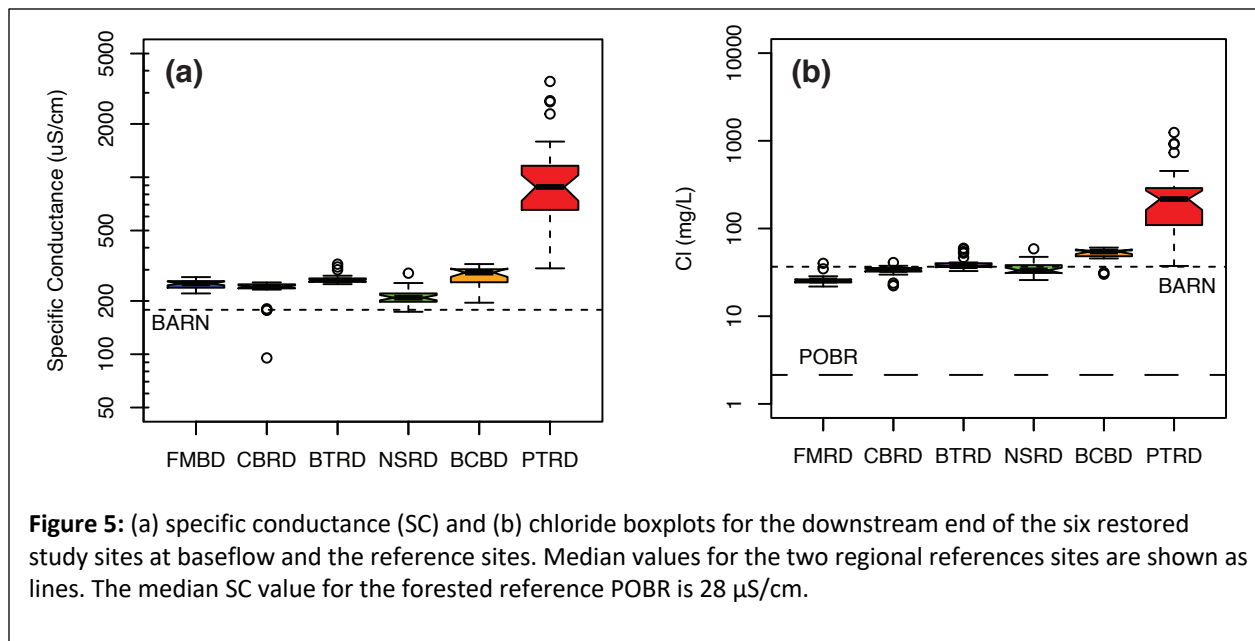
upstream and downstream did not change with the higher 2018 and 2019 discharge (Figure 4) and similar or higher nitrogen concentrations and loads (Figure 3). Both before and after restoration, daily loads for baseflow nitrogen were significantly lower at the downstream end of the study reaches for the (sub)urban BCB and PTR sites (Table S5). Thus it appears that restoration did little to alter in-stream nitrogen cycling at these sites, which differs from observations at other restored sites [65]. While some denitrification may occur in riparian groundwater, it is not detectable. Increases in nitrogen concentrations and especially loads due to high precipitation and baseflow have been previously observed in the region [66, 67]. These increases may arise from several sources: (1) an elevated water table flushing nitrogen from normally hydrologically-unsaturated portions of the shallow subsurface [68], (2) restoration-associated changes such as exposure of nitrogen-rich soils or other material, (3) or elevated nitrogen in riparian groundwater following tree removal [69].

Baseflow dissolved organic carbon, nitrogen speciation & cycling, and phosphorus

DOC, DOC:TDN ratios, and nitrogen speciation yield insights about why little denitrification or in-stream nitrogen removal occurs. Median DOC concentrations range between ~1 and 2.7 mg/L and do not closely correspond to land use characteristics or N concentrations (Figure 2c). Median DOC at the highly urban site is >2 mg/L; at the other sites, DOC is 0.9–1.8 mg/L. Median baseflow DOC:TDN ratios are lowest at the three cultivated crop sites, are ~1 at the pasture and suburban restored sites, and >1 at the highly urban restored site (Figure 2d). At baseflow, nitrate represents ~100% of dissolved nitrogen at all sites except the forested reference site (POBR) where NO_3^- is undetectable (Figure S3a). Colorimetric analyses of NO_3^- (or $\text{NO}_3^- + \text{NO}_2^-$) are equivalent to ion chromatography (Figure S3b).

Carbon availability can limit denitrification, especially in agricultural or urban sites, *e.g.*, when DOC is <2 mg/L or DOC to N ratios are <1 [5, 65, 70-73]. The extensive tree removal that is typically part of LSR/FR may further reduce carbon availability at these sites. Thus in-stream denitrification at restored cultivated crop sites is almost certainly limited by carbon availability and likely limited much of the time at the pasture and suburban restored sites. The pasture and (sub)urban sites are where in-stream denitrification may occur with some frequency based on DOC:TDN ratios. At the BCB and PTR site, similar discharge at the upstream and downstream ends may indicate some groundwater exchange and interaction. By comparison, the four agricultural sites (FMB, CBR, BTR, NSR) appear to be gaining streams with higher discharge at the downstream ends (Table S4). At the highly urban site (PTR), denitrification and/or nitrate removal may be indicated by the lowest dissolved oxygen observed at any site (Figure S4a).

Baseflow orthophosphate (oP) concentrations were low with the restored sites and forested/mostly forested regional references having median concentrations at or below the detection limit of 0.005 mg/L following restoration (Figure S5; Table S6). A few sites (FMB, NSR, BCB, PTR) occasionally had concentrations of 0.005–0.020 mg/L. The post-restoration median concentration at the upstream end of the agricultural FMB site was 0.007 mg/L, which decreased to a median below detection at the downstream end. By contrast, median stormflow oP concentrations at FMB are much higher at 0.461 and 0.254 mg/L for upstream and downstream, respectively (see more below; Table S6) At the three sites with pre-restoration data, the median oP concentration at FMB significantly decreased following restoration with a



larger decrease at the downstream end (0.0305 to 0.0025 mg/L; Table S6); concentrations showed no significant post-restoration change at BCB or PTR.

Baseflow non-nutrient chemistry

Non-nutrient chemistry reflects land use with the reference forested site (POBR) differing substantially with other sites. Differences for most non-nutrient parameters typically were larger among sites than within sites. Median SC values and chloride concentrations illustrate cross-site differences (Figure 5). SC and chloride median concentrations at the agricultural, suburban, and mostly forested reference sites were about an order of magnitude higher than the forested reference site (Figure 5). Median SC and chloride were approximately another order of magnitude higher at the highly urban PTR site (see Supplementary Information for details on other parameters).

Non-nutrient chemistry is relevant since ecological lift is often a stream restoration goal. Elevated SC reflects elevated ion concentrations, a phenomenon sometimes called *freshwater salinization syndrome* [74]. Low ion concentrations at the forested reference site (POBR) are represent regional background conditions to which pre-development aquatic organisms were adapted [75-78]. Agricultural and urban watersheds also have differing ion proportions, e.g., anions shift from being bicarbonate-dominated at the forested reference site to chloride-dominated at all other sites except FMB (Figure S4e). Elevated, differing, and variable ion concentrations driven by land use represent a significant ecological stressor for aquatic organisms and thus affects the post-restoration aquatic community [79-85].

Post-restoration storm events at First Mine Branch (FMB)

Storm event concentrations of DOC, ammonia (NH_4^+), orthophosphate, TDP, and TSS at post-restoration FMB were substantially elevated compared to baseflow concentrations (Figure 6; Table S7). Export of dissolved nitrogen and phosphorus along with TSS were lower at the downstream end than the upstream end, suggesting that the restored reach may successfully

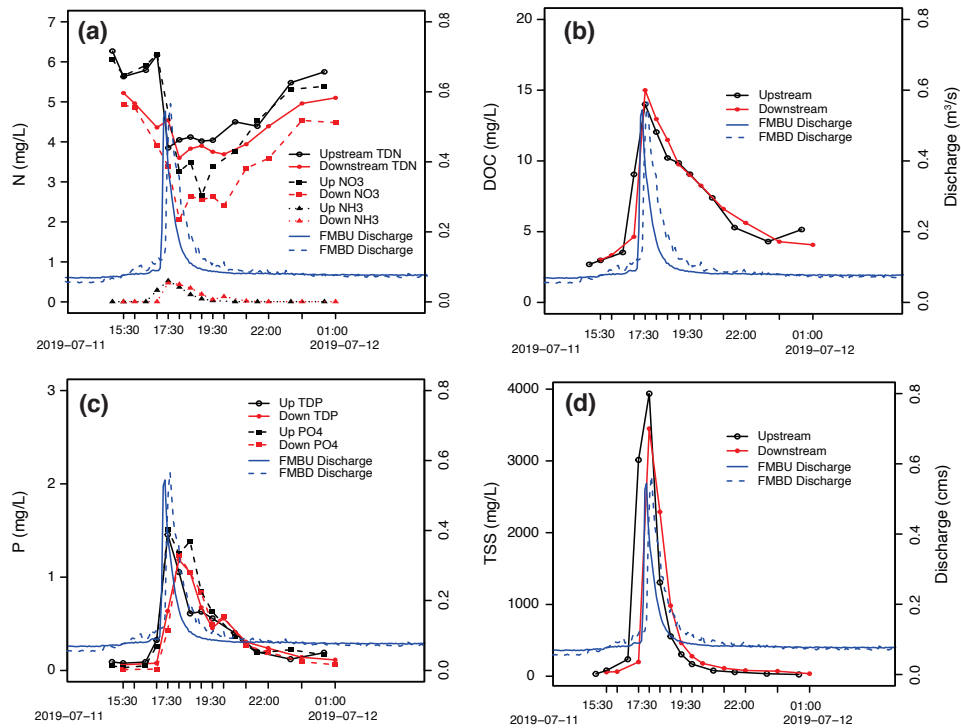


Figure 6: Discharge and concentrations for a July 11, 2019 storm at First Mine Branch (active row crop) site for (a) total dissolved nitrogen, (b) dissolved organic carbon, (d) dissolved phosphorus, and (c) total suspended sediment.

reduce area-normalized loads (Figure 7). Storm event concentrations differ substantially from baseflow, including substantial increases in (1) ammonia from below detection to >1 mg/L (Figure 6a), (2) DOC from ~ 2.7 to 15 mg/L (Figure 6b), (3) orthophosphate and TDP from close to the detection limit to 1.2–1.35 mg/L (Figure 6c), and (4) TSS from near the detection limit (<10 mg/L; Table S7) to median and peak concentrations of 161 mg/L and 3450 mg/L, respectively (Figure 6d; Table S7).

TDN and NO_3^- concentrations were diluted during storms, and so storm loads were similar to the high end of daily baseflow load. During storms, TDN and NO_3^- showed hysteresis with maximum dilution at, or just after peak discharge, and then a slow increase on the falling limb (Figure 6a). Ammonia concentrations showed less hysteresis and were mostly closely associated with peak, and thus overland, flow. The TDN concentration decrease was smaller than NO_3^- , which is partially explainable by increased NH_4^+ concentrations (Figure 6a). The mixture of nitrogen species during storms indicates an important role for surface/near-surface event water contributions of NH_4^+ and organic N in storm export in contrast to all nitrogen being present as nitrate in baseflow (Fig S3a).

Dissolved carbon and phosphorus appeared to be contributed both from overland flow and quickflow subsurface pathways (*e.g.*, interflow) while TSS was likely associated predominantly with overland flow. DOC and phosphorus concentrations displayed hysteresis with an abrupt increase until peak discharge or just after and then subsequent slow decrease (Figure 6b, c). The slow decrease on the falling limb suggests contributions from interflow or other (shallow) subsurface pathways that reach the stream more slowly than overland flow. In contrast, TSS

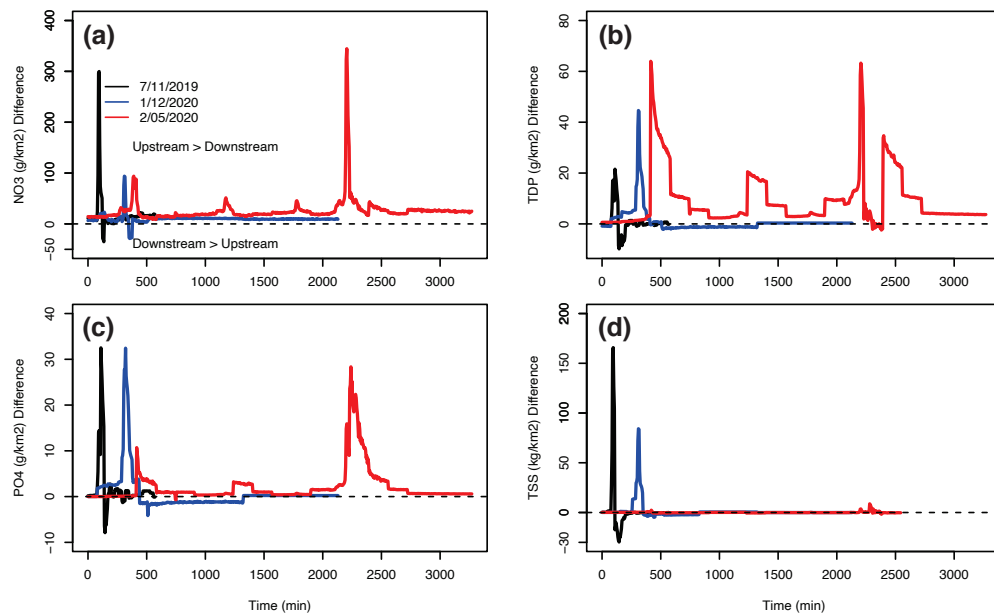


Figure 7: First Mine Branch (FMB) storms, differences between upstream and downstream fluxes for (a) nitrate, (b) total dissolved phosphorus, (c) orthophosphate, and (d) total suspended sediments. Values greater than 0 indicate that the area-normalized flux is higher at the upstream end of the FMB site than the downstream end.

showed little hysteresis with the highest concentrations, and thus fluxes, associated with peak, or near peak discharge (Figure 6d).

Paired comparisons of area-normalized export for nitrogen, phosphorus, and TSS at the upstream and downstream end of FMB across three storm events indicate small load reductions along the restored reach (Figure 7). The peak reduction in nitrogen load was ~5% (300 g/km² versus daily baseflow load of 6343 g/km² from downstream at FMB). Similarly, peak reductions for dissolved phosphorus were ≤5% and for TSS were 6–19%. Integrated across the storms, the reductions ranged from 10–37% between the upstream and downstream ends, which was similar to the observed reductions in specific (or area-normalized) discharge. These results are positive but only suggestive without pre-restoration storm data or additional post-restoration data.

A few factors may contribute to smaller loads at the downstream end. Peak phosphorus and TSS are higher at the upstream end for the three storm events while peak discharge is similar between the upstream and downstream ends (Figure 6, Figure S6). With similar peak discharge between upstream and downstream but a 32% larger watershed area at the downstream end (Table 1), area-normalized exports would be expected to decrease.

More substantial reductions at FMB might be achieved during larger storm events. The engineered bankful capacity at the restored site was 2.83 m³/s, and though the upper end of the rating curve was not well constrained, the high-frequency discharge data suggest bankful capacity was exceeded six times from late April 2019 to mid-June 2020, which is undoubtedly more frequent inundation than occurred pre-restoration. Unfortunately, none of these storms were sampled.

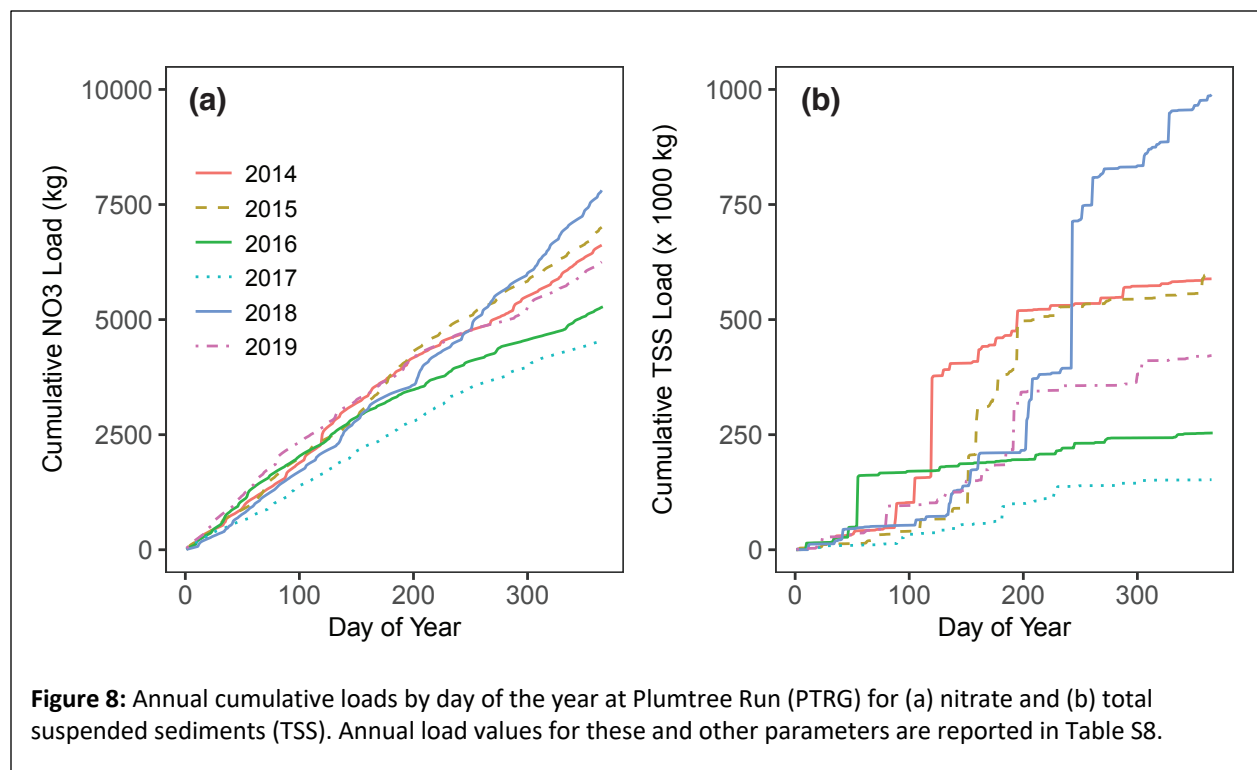


Figure 8: Annual cumulative loads by day of the year at Plumtree Run (PTRG) for (a) nitrate and (b) total suspended sediments (TSS). Annual load values for these and other parameters are reported in Table S8.

Water Quality - Upscaling – storm events downstream of the Plumtree (PTR) restoration

The USGS site (PTRG) downstream of the highly urban restored site (PTR) was used to investigate the larger-scale effects of the mid-2017 PTR restoration. Similar to FMB, NO₃⁻ is diluted as discharge increases while TSS and TDP show flushing behavior (Figure S7). At PTRG, TSS concentrations and discharge peak at about the same time. Based on high-frequency time series, annual loads for NO₃⁻ varied by <2x while annual loads for TSS varied by >7.5x (Figure 8; Table S8); TDP behaved similarly to TSS (Figure S8a). Annual load rankings for NO₃⁻, TDP, and TSS are generally the same and are highly correlated with median annual discharge ($R^2 = 0.94$, 0.84, and 0.82, respectively, Figure S8b, c, d). NO₃⁻ export increases relatively consistently through the year while TSS and TDP export shows substantial increases over brief periods indicating the importance of storm events (Figure 8). Predominantly baseflow-driven NO₃⁻ export and event-driven TSS and phosphorus export also is observed elsewhere in the region [3, 86-88].

Additionally, six paired pre- and post-restoration storm events were investigated based on similar precipitation intensities. Five of the six post-restoration events had higher loads for NO₃⁻, TSS, and TDP with three post-restoration events having TSS and TDP loads that were 3.8–17 times higher than similar pre-restoration events (Table S9). Higher post-restoration event loads are likely due to 2018 and 2019 being high precipitation and discharge years. It is not surprising that the effects of the PTR restoration were not seen downstream given the small reductions observed at FMB. Though the restored reach for the highly urban PTR site represented ~15% of the watershed draining to PTRG, high variability in annual export obscured the larger-scale effects of stream restoration.

Vegetation – Results

The clearest consequence of stream restoration was an 80% reduction in average basal area in restored reaches versus the unrestored reaches (Table 2). Herbaceous and woody layer vegetation at restored sites was significantly more hydrophytic than the unrestored reaches. Species richness decreased by 2-7 species per site for woody communities following restoration with a concomitant marginally significant drop in Shannon diversity (Table 2). Within the herbaceous layer, there was no consistent change in richness or diversity between restored and unrestored reaches, but the change in species richness and Shannon diversity between reach types was negatively correlated with the richness and diversity of the unrestored reach (Figure 9, c and d). Watershed-level ISC was positively correlated with change in woody layer floristic quality (Figure 10), but project length and watershed-level ISC showed no relationship with the amount of change measured for any of the other dependent variables (Table 3). Herbaceous vegetation in the control reaches that were measured over two years also trended more hydrophytic over this time, albeit less than the restored reaches. The changes in basal area and in WIS were not related ($r = 0.429$, $p > 0.5$). Vegetation in restored reaches was more heavily weighted toward native species, but this was only significant for the herbaceous community. Floristic quality did not change significantly for either layer, although the overall FQI for the woody layer decreased due to lower scores at five of six restored reaches (Table 2). In the herbaceous layer, site-level changes in floristic quality and nativity score were negatively related to unrestored reach values (Figure 10, a and b).

Beta diversity within sites increased in the woody layer with plots becoming more different within sites after restoration. Conversely, beta diversity of the woody layer among sites decreased following restoration, meaning that while areas within a site became more different, the sites themselves became more similar to each other. For the herbaceous layer, beta diversity among sites increased with restoration but there was no change in the similarity within sites (Table 2).

NMDS analysis of both the herbaceous and woody layers produced groups by site and restoration status. The herbaceous dataset produced a two-dimensional plot with a final stress of 9.98 and a final instability of 0.00 after 50 iterations (Figure 11). The two axes account for 85.9% of the variation in the dataset with 72.7% of that variation being explained by axis 1. Axis 1 separates the reaches by site, while Axis 2 separates the reaches by restoration status. Woody basal area and WIS were strongly correlated with Axis 1 and beta diversity within sites was strongly correlated with Axis 2. The woody layer dataset produced a one-dimensional solution that explains 63.7% of the variation in the data set, with a final stress of 26.5 and a final instability of 0.00 after 36 iterations (Figure 11). Reaches are grouped more closely by site than by status, but restoration shifts every reach to the right on Axis 1. Correlations of variables with Axis 1 reflect the overall decrease in basal area, species richness, Shannon diversity, WIS, FQI and basal area seen in the woody layer with restoration of these sites.

Indicator species for both communities indicate restoration significantly affected community structure (Figure 12). In the herbaceous layer, the percentage of wetland plants (WIS of OBL or FACW) in the top 30 most highly ranked species rose from 20% to 63% ($p = 0.006$), the percentage of native plants increased from 57% to 73% while invasive plants fell from 33% to 3% ($p = 0.006$), and the percentage of graminoids in the top 30 most important

Table 2. Comparison of unrestored and restored reaches.

Strata	Variable	Dataset 1	Dataset 2	Change	t	df	p
Woody	BA (cm²/m²)	27.88 +/- 6.75	5.99 +/- 2.65	-21.89 +/- 4.49	-2.201 (z)	NA	0.028
	WIS	3.46 +/- 0.08	2.95 +/- 0.18	-0.51 +/- 0.17	3.018	5	0.029
	FQI	16.83 +/- 1.85	14.59 +/- 2.32	-2.24 +/- 0.95	2.360	5	0.065
	NS	2.34 +/- 0.07	2.59 +/- 0.08	0.25 +/- 0.12	2.091	5	0.091
	Richness	27.00 +/- 4.97	21.50 +/- 5.17	-5.50 +/- 1.84	2.990	5	0.030
	Shannon	21.09 +/- 3.88	16.96 +/- 4.38	-4.13 +/- 1.64	2.524	5	0.053
	Beta Within	3.19 +/- 0.45	5.16 +/- 0.94	1.96 +/- 0.70	-2.201 (z)	NA	0.028
	Beta Among	0.63 +/- 0.01	0.55 +/- 0.03	-0.08 +/- 0.03	2.441	10	0.035
Herbaceous	WIS	3.01 +/- 0.16	2.61 +/- 2.61	-0.40 +/- 0.13	-2.201 (z)	NA	0.028
	FQI	28.61 +/- 2.75	29.82 +/- 1.81	1.20 +/- 1.25	-0.964	5	0.379
	NS	1.95 +/- 0.07	2.16 +/- 0.03	0.21 +/- 0.08	2.765	5	0.040
	Richness	67.47 +/- 3.64	81.93 +/- 4.59	14.47 +/- 7.13	-2.028	5	0.098
	Shannon	49.54 +/- 2.65	59.33 +/- 4.26	9.79 +/- 6.16	-1.589	5	0.173
	Beta Within	4.68 +/- 0.29	4.78 +/- 0.25	0.10 +/- 0.34	-0.280	5	0.791
	Beta Among	0.64 +/- 0.004	0.74 +/- 0.001	0.10 +/- 0.005	20.791	10	<0.001
	Herbaceous Control	WIS	3.09 +/- 0.20	2.91 +/- 0.24	-0.18 +/- 0.13		
	FQI	32.49 +/- 2.12	29.53 +/- 2.84	-2.96 +/- 1.06			
	NS	2.16 +/- 0.16	2.14 +/- 0.17	-0.02 +/- 0.03			
	Richness	66.87 +/- 1.01	55.87 +/- 5.01	-11.00 +/- 4.13			
	Shannon	50.67 +/- 1.60	42.26 +/- 3.50	-8.41 +/- 2.68			
	Beta Within	4.30 +/- 0.69	4.41 +/- 0.42	0.12 +/- 0.31			
	Beta Among	0.61 +/- 0.037	0.69 +/- 0.027	0.08 +/- 0.046			

Bold values are comparisons with $p < 0.05$.

Dataset 1 (pre-disturbance) = values for unrestored reaches averaged over two years, values for project reaches pre-restoration and the first year of data for control reaches.

Dataset 2 (post-disturbance) = values for restored reaches averaged over two years, values for project reaches post-restoration and the second year of data for control reaches.

Comparisons between control reaches over time (Dataset 1 vs. Dataset2) are provided to show background interannual variation but were not included in statistical analyses.

WIS = Wetland indicator status weighted by relative cover. Lower scores indicate higher importance of wetland species as WIS run from 1 (obligate wetland) to 5 (upland).

FQI = Floristic quality index calculated from coefficient of conservatism scores [55]. Generally, 1–19 is low quality, 20–35 is high quality, and above 35 is exceptional. Non-native species were not included in these calculations.

NS = Nativity score weighted by relative cover. Native = 3, exotic = 2, invasive = 1

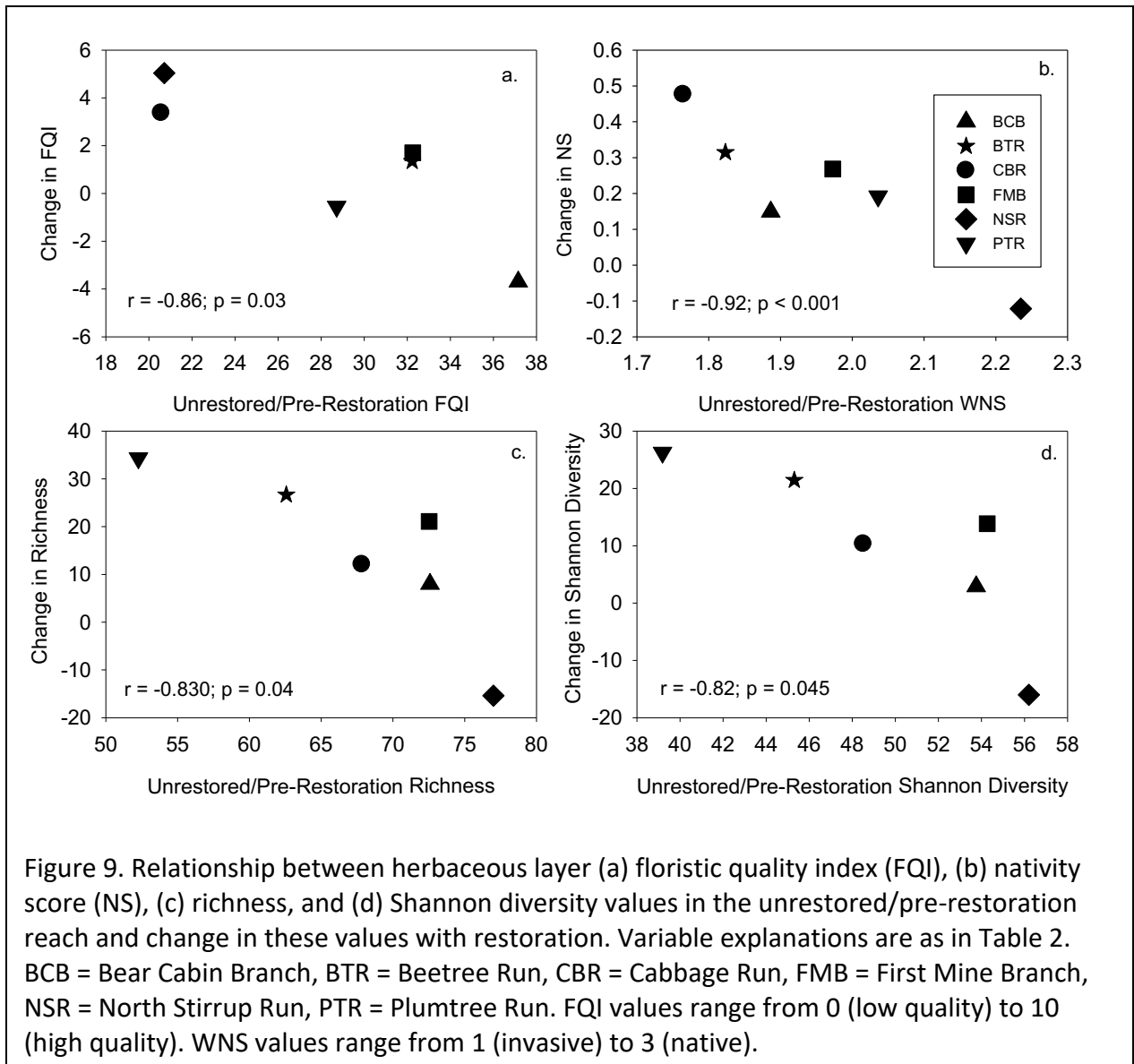
Richness = Data are number of species at each site. Site-level rarefied values for herbaceous layer, absolute site-level values for woody layer.

Shannon diversity = Data are expressed as effective number of equally common species [89].

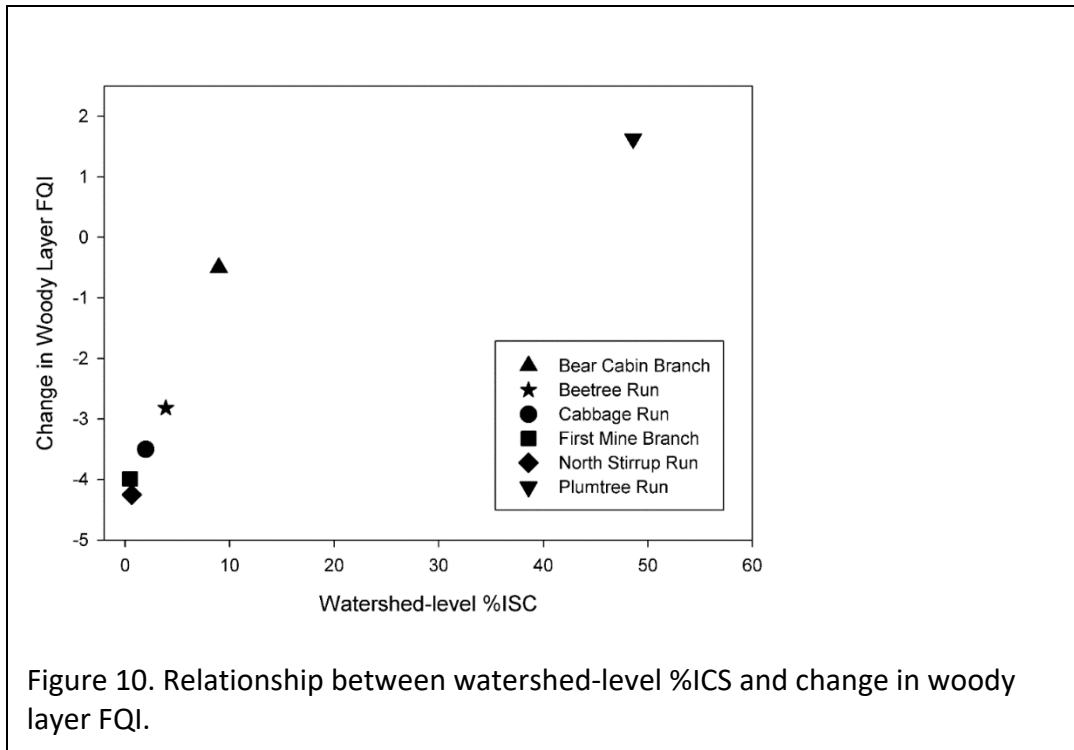
Beta within = Values are average plot richness divided by rarefied (herbaceous layer) or by absolute (woody layer) site level richness

Beta among = Values are Horn index compositional resemblance and run from 0 (communities are identical) to 1 (communities are totally distinct) [60, 61]. Values were compared between pre and post-restoration using the standard errors calculated from 100 bootstrap replications to derive a t-value and corresponding p value [61].

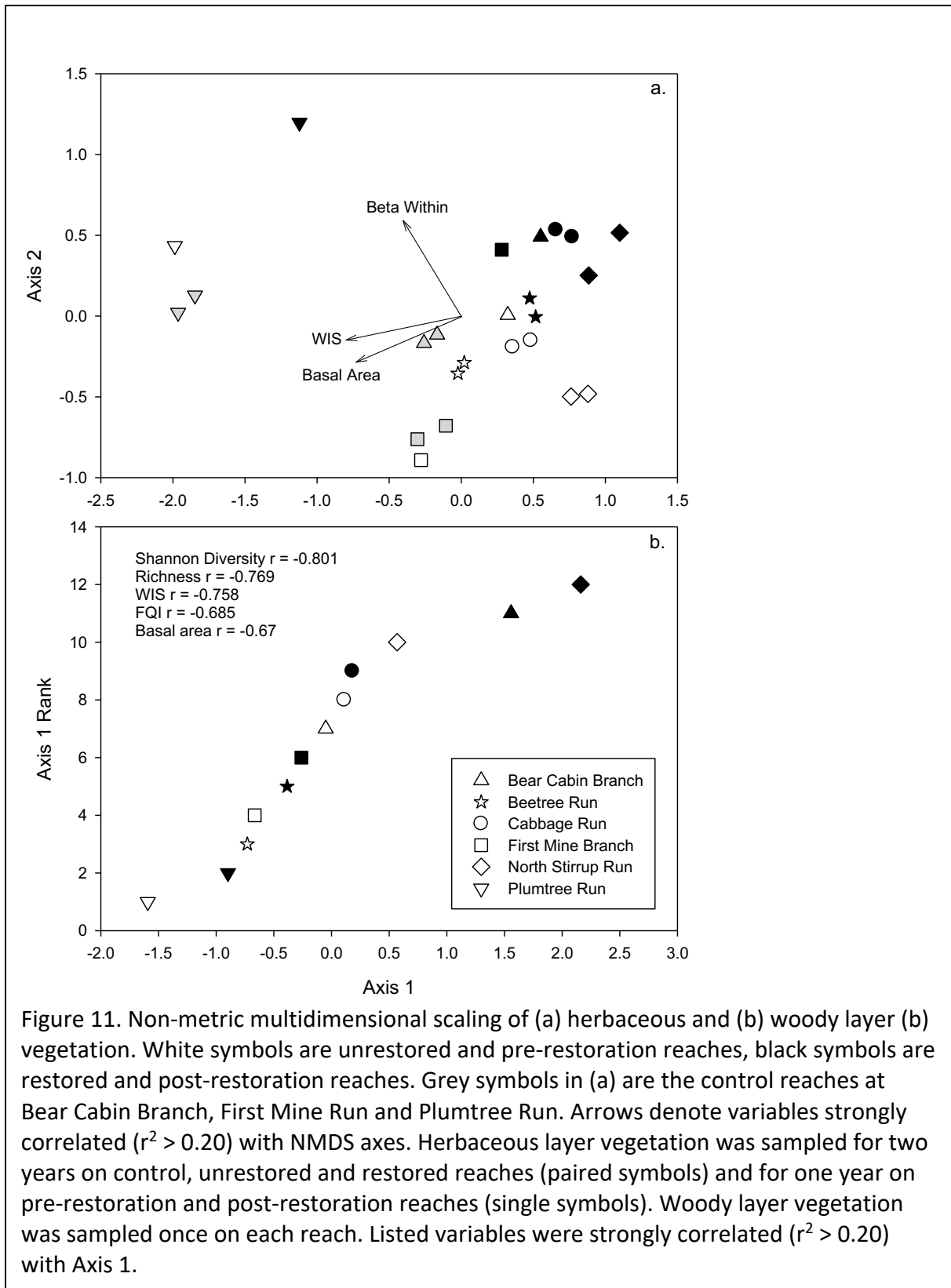
Z = z score from Wilcoxon signed-ranks test which does not have associated degrees of freedom. Paired t-test were used for all other comparisons.



species increased from 10 to 50% ($p < 0.001$). The percentage of annual species in the top 30 most highly ranked species increased from 10 to 30%, but this increase was not significant. Within the top 10 woody species, there was a trend toward an increased proportion of wetland species, a reduction of invasive species, and the elimination of vines, but these changes were not significant (data not shown).



In unrestored reaches, *Rosa multiflora*, *Viola sororia* and *Juglans nigra* were significant herbaceous layer indicators, with *Celastrus orbiculatus* the only significant woody layer indicator of unrestored reaches (Table 3). The importance of these species decreased significantly with restoration. Expanding to include the 10 herbaceous and 5 woody layer species with greatest change in importance value adds six invasive species, three native shrub or vine species, and *Symplocarpus foetidus*, a native obligate wetland species that experienced the largest drop in importance with restoration. All other indicators of unrestored reaches were facultative or upland species. For restored reaches, 21 herbaceous indicator species were identified, only 4 of which were intentionally planted during restoration (Table 4). Of the 18 non-planted indicators, 12 are native species and 10 are wetland species. There were no significant woody layer indicators of restored reaches. The five woody layer species with the largest increase in indicator value with restoration were planted as part of the restoration process. Four of the five were native, wetland species with only *Salix purpurea* being non-native.



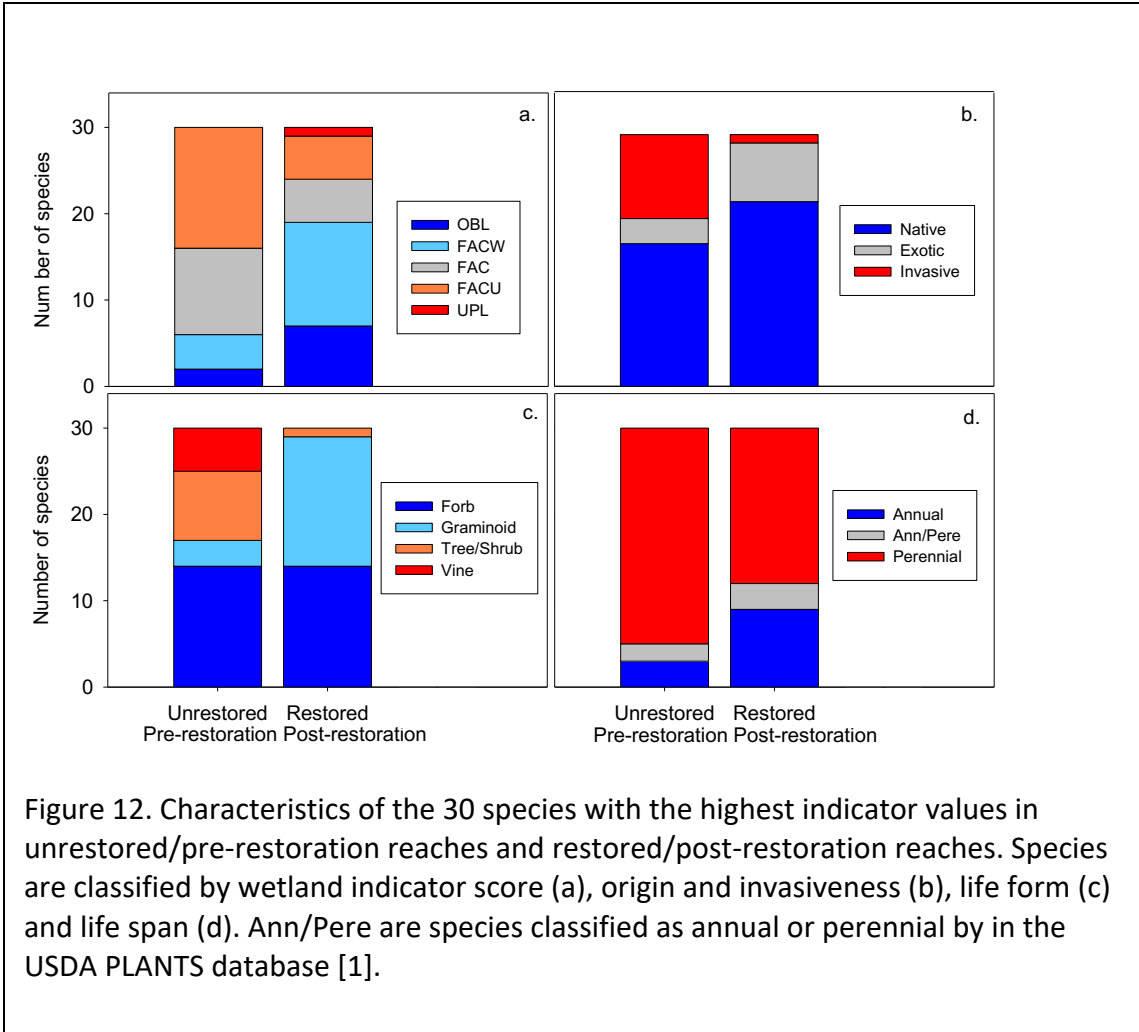


Figure 12. Characteristics of the 30 species with the highest indicator values in unrestored/pre-restoration reaches and restored/post-restoration reaches. Species are classified by wetland indicator score (a), origin and invasiveness (b), life form (c) and life span (d). Ann/Pere are species classified as annual or perennial by in the USDA PLANTS database [1].

Table 3. Indicator species of unrestored reaches.

Layer	Species	U-IV	R-IV	Change	Mean	SD	p	Form	Status	WIS	Life
Herbaceous	<i>Symplocarpus foetidus</i>	91	7	-84	65.7	15.76	0.097	Forb	Native	OBL	Perennial
	<i>Lindera benzoin</i>	84	13	-71	68.8	10.61	0.067	Shrub	Native	FAC	Perennial
	<i>Rosa multiflora</i>	84	16	-68	62.7	10.46	0.007*	Shrub	Invasive	FACU	Perennial
	<i>Alliaria petiolata</i>	84	16	-68	66.0	10.26	0.070	Forb	Invasive	FACU	Annual
	<i>Viola sororia</i>	83	14	-69	60.9	11.56	0.036*	Forb	Native	FAC	A/P
	<i>Amphicarpaea bracteata</i>	77	5	-72	55.2	14.12	0.077	Vine	Native	FAC	A/P
	<i>Juglans nigra</i>	76	3	-73	51.7	14.63	0.037*	Tree	Native	FACU	Perennial
	<i>Allium vineale</i>	74	6	-68	55.7	14.65	0.145	Forb	Invasive	FACU	Perennial
	<i>Glechoma hederacea</i>	72	7	-65	53.1	15.88	0.174	Forb	Invasive	FACU	Perennial
	<i>Berberis thunbergii</i>	65	0	-65	39.9	14.70	0.103	Shrub	Invasive	FACU	Perennial
Woody	<i>Celastrus orbiculatus</i>	80	10	-70	51.5	10.80	0.021*	Vine	Invasive	FACU	Perennial
	<i>Rubus phoenicolasius</i>	60	2	-58	36.7	13.18	0.104	Shrub	Invasive	FACU	Perennial
	<i>Rubus occidentalis</i>	59	2	-57	37.2	13.51	0.155	Shrub	Native	UPL	Perennial
	<i>Lonicera maackii</i>	50	0	-50	44.1	12.51	0.260	Shrub	Invasive	UPL	Perennial
	<i>Vitis riparia</i>	48	2	-46	37.8	14.54	0.223	Vine	Native	FACU	Perennial

Data are sorted by layer and then by indicator value (IV) in the unrestored reach. U = unrestored reach, R = restored reach.

Mean and standard deviation (SD) are IV values derived from 4999 random permutations. P values are the proportion of randomized trials with indicator value equal to or exceeding the observed indicator value. * denotes a species that is a significant ($p < 0.05$) indicator of unrestored reaches.

OBL = obligate (almost always occurs in wetlands); FACW = facultative wetland (usually occurs in wetlands, but may occur in non-wetlands); FAC = facultative (occurs in wetlands and non-wetlands); FACU = facultative upland (usually occurs in non-wetlands, but may occur in wetlands); UPL = upland (almost never occurs in wetlands).

A/P are species classified as annual or perennial by in the USDA PLANTS database [1].

Table 4 – Indicator species of restored reaches.

Layer	Species	U-IV	R-IV	Change	Mean	SD	p	Form	Status	WIS	Life	Planted Sites
Herbaceous	<i>Echinochloa colonum</i>	0	100	100	42.7	14.28	0.003*	Graminoid	Exotic	FACW	Annual	0
	<i>Cyperus esculentus</i>	0	99	99	49.9	14.99	0.003*	Graminoid	Native	FACW	Perennial	0
	<i>Persicaria maculosa</i>	0	99	99	50.7	15.07	0.007*	Forb	Exotic	FACW	A/P	0
	<i>Trifolium repens</i>	1	98	97	65.6	13.33	0.009*	Forb	Exotic	FACU	Perennial	0
	<i>Juncus effusus</i>	3	97	94	63.8	13.30	0.005*	Graminoid	Native	FACW	Perennial	3
	<i>Leersia oryzoides</i>	3	95	92	59.7	14.38	0.003*	Graminoid	Native	OBL	Perennial	0
	<i>Lobelia inflata</i>	2	94	92	54.0	14.59	0.007*	Forb	Native	FACU	Annual	0
	<i>Carex lurida</i>	4	94	90	55.4	10.88	0.003*	Graminoid	Native	OBL	Perennial	0
	<i>Echinochloa crusgalli</i>	3	93	90	53.0	12.59	0.007*	Graminoid	Exotic	FAC	Annual	1
	<i>Bidens vulgata</i>	4	92	88	56.5	13.22	0.007*	Forb	Native	FAC	Annual	0
	<i>Carex vulpinoidea</i>	3	90	87	50.5	13.17	0.008*	Graminoid	Native	OBL	Perennial	1
	<i>Arthraxon hispidus</i>	6	92	86	59.9	14.20	0.042*	Graminoid	Invasive	FAC	Annual	0
	<i>Salix nigra</i>	5	89	84	53.6	12.71	0.005*	Tree	Native	OBL	Perennial	5
	<i>Vicia tetrasperma</i>	6	88	82	52.9	12.25	0.020*	Forb	Exotic	UPL	Annual	0
	<i>Setaria pumila</i>	1	81	80	40.3	12.71	0.014*	Graminoid	Exotic	FAC	Annual	0
	<i>Sagittaria latifolia</i>	1	78	77	42.8	14.49	0.042*	Forb	Native	OBL	Perennial	0
	<i>Acalypha rhomboidea</i>	3	79	76	44.5	12.52	0.011*	Forb	Native	FACU	Annual	0
	<i>Juncus antheratus</i>	10	85	75	56.0	11.18	0.016*	Graminoid	Native	FACW	Perennial	0
	<i>Epilobium coloratum</i>	7	79	72	48.7	12.71	0.036*	Forb	Native	FACW	Perennial	0
	<i>Persicaria sagittata</i>	13	84	71	62.8	11.22	0.048*	Forb	Native	OBL	A/P	0
<i>Ambrosia artemisiifolia</i>	9	74	65	48.7	12.25	0.027*	Forb	Native	FACU	Annual	0	
Woody	<i>Salix purpurea</i>	0	67	67	31.8	12.81	0.062	Tree	Exotic	FACW	Perennial	3
	<i>Platanus occidentalis</i>	16	57	41	47.5	11.59	0.195	Tree	Native	FACW	Perennial	5
	<i>Cornus ammomum</i>	4	44	40	40.4	14.24	0.276	Shrub	Native	FACW	Perennial	6
	<i>Aronia arbutifolia</i>	6	33	27	31.1	13.45	0.562	Shrub	Native	FACW	Perennial	1
	<i>Salix nigra</i>	33	60	27	57.5	8.63	0.334	Tree	Native	OBL	Perennial	5

Data are sorted by layer and then by indicator value (IV) in the restored reach.

See explanations in Table 4.

Discussion

Water Quality Implications

Assessing effectiveness of stream restoration projects is challenging: 1) pre-restoration data are commonly unavailable, 2) collecting sufficient data to accurately characterize changes in fluxes and loads is time-intensive and expensive, 3) many years may elapse before the desired outcomes of restoration are observable due to legacy inputs and/or the time needed to establish key biogeochemical processes, and 4) stream restoration responses may be obscured by weather/climate variability [67, 90]. For example, high retention and removal rates at one Mid-Atlantic restoration resulted in decreased flow-weighted mean concentrations of most constituents, but loads did not significantly decrease for NO_3^- and TSS [91]. Additionally, for constituents with reduced loads, the effects of restoration were not detectable ~600 m downstream except for orthophosphate.

The Chesapeake Bay Program grants regulatory credit to stream restoration projects upon completion via three relevant protocols with each being annual and additive [92]. *Protocol 1 Prevented Sediment during Storm Flow* credits sediment and nutrient reduction based on prevented bank or channel erosion. *Protocol 2 Instream and Riparian Nutrient Processing During Baseflow* credits nitrogen reductions based on design features promoting denitrification during baseflow. *Protocol 3 Floodplain Reconnection Volume* credits sediment and nutrient reduction for reconnecting stream channels to floodplains.

Removal of erodible bank and floodplain deposits that were 1–2.5 m thick at the study sites via LSR/FR restoration likely resulted in decreased TSS loads and associated particulate nitrogen and phosphorus, suggesting that *Protocol 1* was addressed. In mid-Atlantic Piedmont watersheds, bank sediments supply much of the suspended load [19], e.g., 6-90% (average 57%) in Baltimore County, MD and 30–65% in Big Spring Run [23, 93]. Substantial decreases in TSS concentrations and a 69% decrease in TSS loads with accompanying total phosphorus decreases were observed at Big Spring Run, ~50 km northeast of the study region [93]. Among the first LSR restorations, Big Spring Run has a high-frequency data record from multiple USGS gages for three years pre-restoration and four years post-restoration. At another Piedmont LSR/FR site, TSS fluxes increased during restoration with the restoration period extending across the 2018–2019 high discharge years [no post-restoration data available, 94].

We found little evidence for decreased post-restoration nitrogen concentrations or fluxes resulting from denitrification or other biogeochemical processes (*Protocol 2*). Results were similar at the Big Spring Run LSR/FR restoration where NO_3^- concentrations increased (though not significantly) and orthophosphate remained unchanged [93]; groundwater NO_3^- concentrations began decreasing 4–5 years post-restoration with carbon accumulation as an important factor [95]. At a Piedmont site with a similar study period (2017–19), NO_3^- concentrations decreased slightly but fluxes increased during restoration; orthophosphate did not change [94]. A Piedmont site with observed post-restoration nitrogen reductions had initially low nitrogen concentrations (similar to the (sub)urban BCB and PTR sites), and reductions may have resulted from a combination of denitrification and lower nitrogen inputs to restored reaches [96]. Baseflow nitrogen in this study was entirely nitrate and remained so following restoration. By contrast, sites with substantial post-restoration denitrification had large amounts of carbon added during restoration and multiple lines of evidence pointed to

denitrification, *e.g.*, nitrate decreases accompanied by ammonia and dissolved organic nitrogen increase and low dissolved oxygen at times [3].

Other post-restoration factors may contribute to increased denitrification with time. Denitrification may increase in the lower post-restoration floodplains due to increased carbon inputs from more frequent inundation or by a water table closer to the surface. Terrestrial vegetation biomass, and thus carbon inputs, on restored floodplains should increase with time.

Lags between restoration completion and improved water quality are common, especially in agricultural watersheds. Time is needed to establish sufficient rates of biogeochemical processing. Legacy nutrients in soils and groundwater may take years to decades to reach streams [5, 90, 97, 98]. Legacy contributions (older than 1 year) represent 50% of nitrogen exported by the Susquehanna River, and 18% is older than 10 years [99].

We have insufficient data to determine the success of changes related to *Protocol 3* for floodplain reconnection. Achieving success related to this protocol may be challenging, particularly in urban watersheds. In five urban Piedmont watersheds, much of the water accessing the floodplain during overbank events occurred in during large events (50% of floodplain flow during 2–3 events per year), resulting in 0.2–1% of nitrogen retention or removal by Protocol 3 [100].

Results of this study suggest some positive outcomes of LSR/FR stream restoration. Future research on LSR/FR or similar projects should likely focus on data collection related to sediment erosion prevention and floodplain reconnection (*Protocols 1 and 3*). Pre- and post-restoration high-frequency data collection for discharge and sediment transport would greatly facilitate answering questions related whether these protocols are successfully addressed. Groundwater sampling would be useful for addressing *Protocols 2 and 3*, but changes may take several years to emerge, especially related to baseflow nutrient processing.

Riparian Vegetation Implications

Legacy sediment removal and floodplain reconnection resulted in marked changes in the composition of both the woody and herbaceous layers. Changes to the overstory community included a substantial decrease in basal area and woody species richness (Table 3), an expected consequence of the excavation that accompanies these projects [101]. As anticipated, obligate and facultative wetland species increased in importance post-restoration, largely due to removal of facultative upland and upland species with restoration and, to a lesser extent, planting of obligate and facultative wetland species. The similarity of the woody community across plots within a reach decreased, likely because of the removal of a few widely distributed invasive species, leaving pockets of vegetation dominated by different species. However, woody layer similarity increased among sites, indicating that restoration activities and active revegetation may be causing a regional homogenization of the overstory layer of restored riparian zone ecosystems. Priority effects from planting are expected to dominate restored woody communities [28, 29], and if propagule dispersal is inhibited, the artificially selected community may persist into the future. Although not considered invasive, planting of non-native *Salix purpurea* at three of six sites is particularly concerning.

Changes in the herbaceous layer composition (Table 3) included a shift toward towards more hydrophytic and graminoid species [26, 102], but this may change as sites age [102]. Although we did not monitor groundwater depth or the frequency of overbank flooding, the

change in species composition signals that restoration altered the hydrology of these floodplains [27, 43, 103, 104]. While tree removal can decrease evapotranspiration and water table elevation, we did not find a relationship between site level decrease in basal area and change in WIS. The increase in graminoid importance (Figure 11c) and the increase in beta diversity among sites (Table 3) may be fueled by seed banks remaining in the legacy sediment left on site. Significant indicators of restored reaches included many native species that were not intentionally planted (Table 5), suggesting that each site has a unique seedbank that can increase regeneration potential and decrease regional homogenization [105-108]. In the NMDS of the herbaceous layer (Figure 9), the axis separating reaches by site explained far more variation than the axis separating reaches by restoration status, indicating that restoration clearly alters species composition, but not to the point that sites lose their unique herbaceous-layer identity post-restoration.

A common consequence of disturbance is higher abundance of disturbance-adapted exotic plant species [40, 109, 110], with riparian systems particularly susceptible to colonization by exotic species [32-38]. What little research has been done on herbaceous plant communities in riparian restoration projects shows higher numbers of non-native species in restored reaches than reference reaches (Malone, 2011). We found that restoration was associated with a significant change in nativity index score but not in FQI, which highlights the different sensitivities of these indices. The FQI focuses on native species, with exotic and invasive species either not scored or assigned a value of zero. Our nativity index awarded all native species the same value but included separate scores for exotic and invasive species. Comparison of the 30 most important species between restored and unrestored reaches shows that while restoration slightly increased the importance of native species overall, it also resulted in a decrease in invasive species importance and an increase in the importance of non-native exotic species (Figure 11b). We also found that restoration was not associated with a significant change in richness or Shannon diversity in the herbaceous layer, indicating that, on average, plant community quality and diversity do not significantly decline, at least in the first few years after project completion [26, 111].

While the results of the restoration projects we surveyed were largely positive with an increase in hydrophytic plant species and no decrease in herbaceous-layer richness, diversity or native community quality, there are several potential drawbacks to these projects with respect to the riparian vegetation community. First, while tree removal is an inescapable result of this process, extensive tree removal may be detrimental to nutrient cycling, which is the primary motivation of many projects. Loss of leaf litter from trees and from skunk cabbage, a quintessential wetland species that, at least initially, is nearly eradicated by sediment excavation, may deprive floodplains of critical carbon sources that are essential for nitrogen amelioration [103, 112]. In a companion project, McMahan, Beauchamp [113] found that restoration had no effect on the flux of total dissolved nitrogen at any of these six sites, likely due to a limited supply of dissolved organic carbon and legacy effects of nitrogen in groundwater.

We also found that low quality sites, in terms of community composition, richness and diversity improved the most with restoration, but high quality sites improved the least or even decreased in quality after restoration (Figure 8). In many areas, legacy sediment terraces are habitat for native forest understory plant species. While these upland or ecotonal communities

are being replaced with what is, at least initially, good quality riparian habitat, it is certainly a trade-off. Preservation of high-quality forest areas, even if they are atop legacy sediment terraces, should be considered, particularly if losses in tree canopy are not being offset by gains in nutrient cycling. We did find that we found that woody layer vegetation quality increased the most at sites with the highest impervious surface cover, but we saw no other relationships between project length or watershed ISC and site recovery in terms of FQI, nativity score, richness and diversity. More disturbed watersheds may benefit most from removal of exotic trees and revegetation with native species, but for all other variables selecting larger projects or areas with more or less ISC did not affect the outcome of restoration in terms of creating quality riparian and wetland habitat.

Acknowledgements

This project was supported with a Restoration Research grant (award #13974) from the Chesapeake Bay Trust and funding from Towson University. We thank Ginny Jeppi, David Cote and Lisa Wheeler for field assistance and Charlie Davis for help with plant identification. Scott and Colin McGill at Ecotone Inc. were invaluable partners and we thank Howard County, the City of Bel Air and the Edwards, Rigdon and Pitts families for access to their properties. This work was supported by a Restoration Research grant from the Chesapeake Bay Trust (VBB, RAC, CJS, JM) and partially by the National Science Foundation ICER-1540631 (JM). We gratefully acknowledge Caroline Stanley, Colin McGill, Scott McGill, and others at Ecotone, Inc. for their assistance in site selection without which this project would not have been possible, for connecting us with property owners to gain site access, and for their willingness to share site design plans and other relevant information. The authors thank David Cote, Lisa Wheeler Ginny Jeppi, Chris Manning, and Michael Dawson for assistance in the field and laboratory along with Mark Monk for invaluable support with laboratory procedures and instrumentation. We also thank Charlie Davis for help with plant identification and Howard County, the City of Bel Air and the Edwards, Rigdon and Pitts families for access to their properties.

Data availability

Will be publicly available through Dryad.

Literature Cited

1. USDA, NRCS. *The PLANTS Database*. 2020; Available from: <http://plants.usda.gov>.
2. Kemp, W.M., et al., *Eutrophication of Chesapeake Bay: Historical trends and ecological interactions*. Marine Ecology Progress Series, 2005. **303**: p. 1-29.
3. Williams, M.R., et al., *Stream Restoration Performance and Its Contribution to the Chesapeake Bay TMDL: Challenges Posed by Climate Change in Urban Areas*. Estuaries and Coasts, 2017. **40**(5): p. 1227-1246.
4. Bernhardt, E.S., et al., *Synthesizing U.S. River Restoration Efforts*. Science, 2005. **308**(5722): p. 636-637.
5. Lammers, R.W. and B.P. Bledsoe, *What role does stream restoration play in nutrient management?* Critical Reviews in Environmental Science and Technology, 2017. **47**(6): p. 335-371.
6. Beechie, T.J., et al., *Process-based principles for restoring river ecosystems*. BioScience, 2010. **60**(3): p. 209-222.
7. Palmer, M.A., K.L. Hondula, and B.J. Koch, *Ecological restoration of streams and rivers: Shifting strategies and shifting goals*. Annual Review of Ecology, Evolution, and Systematics, 2014. **45**: p. 247-269.
8. Walter, R.C. and D.J. Merritts, *Natural streams and the legacy of water-powered mills*. Science, 2008. **319**(5861): p. 299-304.
9. Merritts, D., et al., *Anthropocene streams and base-level controls from historic dams in the unglaciated mid-Atlantic region, USA*. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2011. **369**(1938): p. 976-1009.
10. Bain, D.J. and G.S. Brush, *Early chromite mining and agricultural clearance: Opportunities for the investigation of agricultural sediment dynamics in the Eastern Piedmont (USA)*. American Journal of Science, 2005. **305**(9): p. 957-981.
11. Groffman, P.M., et al., *Soil nitrogen cycling processes in urban riparian zones*. Environ. Sci. Technol., 2002. **36**.
12. Arnold, C.L. and C.J. Gibbons, *Impervious Surface Coverage: The Emergence of a Key Environmental Indicator*. Journal of the American Planning Association, 1996. **62**(2): p. 243-258.
13. Groffman, P.M., et al., *Down by the riverside: urban riparian ecology*. Frontiers in Ecology and the Environment, 2003. **1**(6): p. 315-321.
14. Voli, M., et al., *Preliminary reconstruction of a pre-European settlement valley bottom wetland, southeastern Pennsylvania*. Water Resources Impact, 2009. **11**(5): p. 11-13.
15. Gold, A.J., et al., *LANDSCAPE ATTRIBUTES AS CONTROLS ON GROUPTH WATER NITRATE REMOVAL CAPACITY OF RIPARIAN ZONES¹*. JAWRA Journal of the American Water Resources Association, 2001. **37**(6): p. 1457-1464.
16. Groffman, P.M. and M.K. Crawford, *Denitrification Potential in Urban Riparian Zones*. Journal of Environmental Quality, 2003. **32**(3): p. 1144-1149.
17. Weitzman, J.N., et al., *Potential nitrogen and carbon processing in a landscape rich in milldam legacy sediments*. Biogeochemistry, 2014. **120**(1): p. 337-357.
18. Weitzman, J.N. and J.P. Kaye, *Nitrate retention capacity of milldam-impacted legacy sediments and relict A horizon soils*. SOIL, 2017. **3**(2): p. 95-112.

19. Hupp, C.R., et al., *Recent and historic sediment dynamics along Difficult Run, a suburban Virginia Piedmont stream*. *Geomorphology*, 2013. **180-181**: p. 156-169.
20. Mukundan, R., et al., *Sediment Fingerprinting to Determine the Source of Suspended Sediment in a Southern Piedmont Stream*. *Journal of Environment Quality*, 2010. **39**(4): p. 1328.
21. Schenk, E.R. and C.R. Hupp, *Legacy effects of colonial millponds on floodplain sedimentation, bank erosion, and channel morphology, MID-Atlantic, USA*. *Journal of the American Water Resources Association*, 2009. **45**(3): p. 597-606.
22. Lyons, N.J., et al., *Bank erosion of legacy sediment at the transition from vertical to lateral stream incision*. *Earth Surface Processes and Landforms*, 2015. **40**(13): p. 1764-1778.
23. Donovan, M., et al., *Sediment contributions from floodplains and legacy sediments to Piedmont streams of Baltimore County, Maryland*. *Geomorphology*, 2015. **235**: p. 88-105.
24. Walter, R., D. Merritts, and M. Rahnis, *Estimating volume, nutrient content, and rates of stream bank erosion of legacy sediment in the Piedmont and Valley and Ridge physiographic provinces, southeastern and central PA*. 2007: A Report to the Pennsylvania Department of Environmental Protection. p. 40.
25. Odhiambo, B.K., et al., *Effects of forested floodplain soil properties on phosphorous concentrations in two Chesapeake Bay sub-watersheds, Virginia, USA*. *Environmental Science and Pollution Research*, 2016. **23**(16): p. 16056-16066.
26. Göthe, E., et al., *Structural and functional responses of floodplain vegetation to stream ecosystem restoration*. *Hydrobiologia*, 2015. **769**(1): p. 79-92.
27. Fraaije, R.G.A., et al., *Functional responses of aquatic and riparian vegetation to hydrogeomorphic restoration of channelized lowland streams and their valleys*. *Journal of Applied Ecology*, 2019. **56**(4): p. 1007-1018.
28. Kaase, C.T. and G.L. Katz, *Effects of Stream Restoration on Woody Riparian Vegetation of Southern Appalachian Mountain Streams, North Carolina, U.S.A*. *Restoration Ecology*, 2012. **20**(5): p. 647-655.
29. Beauchamp, V.B., et al., *Riparian community structure and soil properties of restored urban streams*. *Ecohydrology*, 2015. **8**(5): p. 880-895.
30. Lesage, J.C., E.A. Howard, and K.D. Holl, *Homogenizing biodiversity in restoration: the "perennialization" of California prairies*. *Restoration Ecology*, 2018. **26**(6): p. 1061-1065.
31. Nilsson, C., et al., *The role of hydrochory in structuring riparian and wetland vegetation*. *Biological Reviews*, 2010. **85**(4): p. 837-858.
32. Malone, M.E., *A Comparison of the Vegetation and Soils of Restored Streams and their References in the NC Piedmont*, in *Natural Resources*. 2011, North Carolina State University: Raleigh, North Carolina. p. 67.
33. Cooper, D.J., et al., *Mountain wetland restoration: The role of hydrologic regime and plant introductions after 15 years in the Colorado Rocky Mountains, U.S.A*. *Ecological Engineering*, 2017. **101**: p. 46-59.
34. Catford, J.A. and R. Jansson, *Drowned, buried and carried away: Effects of plant traits on the distribution of native and alien species in riparian ecosystems*. *New Phytologist*, 2014. **204**(1): p. 19-36.

35. Planty-Tabacchi, A.M., et al., *Invasibility of species rich communities in riparian zones*. Conservation Biology, 1996. **10**(2): p. 598-607.
36. Richardson, D.M., et al., *Riparian vegetation: degradation, alien plant invasions, and restoration prospects*. Diversity and Distributions, 2007. **13**(1): p. 126-139.
37. Tabacchi, E., et al., *Seed inputs in riparian zones: Implications for plant invasion*. River Research and Applications, 2005. **21**(2-3): p. 299-313.
38. Vidra, R.L., T.H. Shear, and T.R. Wentworth, *Testing the paradigms of exotic species invasion in urban riparian forests*. Natural Areas Journal, 2006. **26**(4): p. 339-350.
39. Larson, D.M., et al., *Sediment excavation as a wetland restoration technique had early effects on the developing vegetation community*. Wetlands Ecology and Management, 2020. **28**(1): p. 1-18.
40. Wise, M., *A field investigation into the effects of anthropogenic disturbances on biodiversity and alien invasions of plant communities*. Bioscience, 2017. **43**: p. 3-14.
41. McLane, C.R., et al., *Succession of Exotic and Native Species Assemblages within Restored Floodplain Forests: A Test of the Parallel Dynamics Hypothesis*. Restoration Ecology, 2012. **20**(2): p. 202-210.
42. Deane, D.C., et al., *Dispersal-driven homogenization of wetland vegetation revealed from local contributions to β -diversity*. Journal of Vegetation Science, 2017. **28**(5): p. 893-902.
43. Hammersmark, C.T., et al., *Simulated Effects of Stream Restoration on the Distribution of Wet-Meadow Vegetation*. Restoration Ecology, 2010. **18**(6): p. 882-893.
44. Qian, H. and Q. Guo, *Linking biotic homogenization to habitat type, invasiveness and growth form of naturalized alien plants in North America*. Diversity and Distributions, 2010. **16**(1): p. 119-125.
45. Homer, C.G., et al., *Completion of the 2011 National Land Cover Database for the conterminous United States-Representing a decade of land cover change information*. Photogrammetric Engineering & Remote Sensing, 2015. **81**: p. 345-354.
46. American Public Health Association, *Standard methods for the examination of water and wastewater*. 1998: Washington, DC.
47. R Core Team, *R: A language and environment for statistical computing*. 2019, R Foundation for Statistical Computing.
48. R Studio Team, *RStudio: Integrated Development Environment for R*. 2018, RStudio, Inc.: Boston, MA.
49. Golemund, G. and H. Wickham, *Dates and times made easy with lubridate*. Journal of Statistical Software, 2011. **40**(3): p. 1-25.
50. De Cicco, L.A., et al., *dataRetrieval: R packages for discovering and retrieving water data available from U.S. federal hydrologic web services*. 2018, U.S. Geological Survey: Reston, VA.
51. Vanderkam, D., et al., *dygraphs: Interface to 'Dygraphs' Interactive Time Series Charting Library*. 2018.
52. Wickham, H., et al., *dplyr: A Grammar of Data Manipulation*. 2019.
53. Weakley, A.S., et al., *Flora of Virginia*. 1st ed. ed. 2012, Fort Worth, Tex.: Botanical Research Institute of Texas Press.

54. Wentworth, T.R., G.P. Johnson, and R.L. Kologiski, *Designation of wetlands by weighted averages of vegetation data: A preliminary evaluation*. JAWRA Journal of the American Water Resources Association, 1988. **24**(2): p. 389-396.
55. Chamberlain, S.J. and H.M. Ingram, *Developing coefficients of conservatism to advance floristic quality assessment in the Mid-Atlantic region*. The Journal of the Torrey Botanical Society, 2012. **139**(4): p. 416-427, 12.
56. Freyman, W.A., L.A. Masters, and S. Packard, *The Universal Floristic Quality Assessment (FQA) Calculator: an online tool for ecological assessment and monitoring*. Methods in Ecology and Evolution, 2016. **7**(3): p. 380-383.
57. UME, *University of Maryland Extension Home and Garden Information Center Compiled Invasive Plant List*. 2017. <https://extension.umd.edu/hgic/topics/compiled-invasive-plant-list>.
58. Chao, A., K.H. Ma, and T.C. Hsieh, *iNEXT (iNterpolation and EXTrapolation) Online: Software for Interpolation and Extrapolation of Species Diversity. Program and User's Guide published at http://chao.stat.nthu.edu.tw/wordpress/software_download/*. 2016
59. Chao, A. and L. Jost, *Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size*. Ecology, 2012. **93**(12): p. 2533-2547.
60. Chao, A., et al., *Online Program SpadeR (Species richness Prediction And Diversity Estimation in R)*. Program and Users Guide published at http://chao.stat.nthu.edu.tw/wordpress/software_download/, 2015.
61. Chao, A., et al., *A two stage probabilistic approach to multiple community similarity indices*. Biometrics, 2008. **64**: p. 1178-1186.
62. IBM Corp, *IBM SPSS Statistics for Windows, Version 25.0*. Armonk, NY: IBM Corp. 2017.
63. McCune, B. and J.B. Grace, *Analysis of Ecological Communities*. 2002, Glenden Beach, OR: MJM Software Design. 304.
64. McCune, B. and M.J. Mefford, *PC-ORD. Multivariate Analysis of Ecological Data*. 2006, MjM Software: Glenden Beach, Oregon, U.S.A.
65. Mayer, P.M., et al., *Nitrogen Dynamics at the Groundwater–Surface Water Interface of a Degraded Urban Stream*. Journal of Environmental Quality, 2010. **39**(3): p. 810-823.
66. Duncan, J.M., et al., *Dynamics of nitrate concentration-discharge patterns in an urban watershed*. Water Resources Research, 2017. **53**(8): p. 7349-7365.
67. Seybold, E., et al., *Influence of land use and hydrologic variability on seasonal dissolved organic carbon and nitrate export: insights from a multi-year regional analysis for the northeastern USA*. Biogeochemistry, 2019. **146**(1): p. 31-49.
68. Zhi, W. and L. Li, *The shallow and deep hypothesis: Subsurface vertical chemical contrasts shape nitrate export patterns from different land uses*. Environmental Science & Technology, 2020. **54**: p. 11915–11928.
69. Wood, K.L., et al., *H43Q-2313: Observations of riparian groundwater quality following restoration projects to model chemical impacts of tree removal*, in *Fall American Geophysical Union Meeting*. 2019: San Francisco, CA.
70. Filoso, S. and M.A. Palmer, *Assessing stream restoration effectiveness at reducing nitrogen export to downstream waters*. Ecological Applications, 2011. **21**(6): p. 1989-2006.

71. Sivirichi, G.M., et al., *Longitudinal variability in streamwater chemistry and carbon and nitrogen fluxes in restored and degraded urban stream networks*. Journal of Environmental Monitoring, 2011. **13**: p. 288-303.
72. Roley, S.S., et al., *Floodplain restoration enhances denitrification and reach-scale nitrogen removal in an agricultural stream*. Ecological Applications, 2012. **22**(1): p. 281-297.
73. Smith, R.M., et al., *Influence of infrastructure on water quality and greenhouse gas dynamics in urban streams*. Biogeosciences, 2017. **14**(11): p. 2831-2849.
74. Kaushal, S.S., et al., *Freshwater salinization syndrome on a continental scale*. Proceedings of the National Academy of Sciences of the United States of America, 2018. **115**(4): p. E574-E583.
75. Kaushal, S.S., et al., *Human-accelerated weathering increases salinization, major ions, and alkalinization in fresh water across land use*. Applied Geochemistry, 2017. **83**: p. 121-135.
76. Moore, J., et al., *Nonpoint source contributions drive elevated major ion and dissolved inorganic carbon concentrations in urban watersheds*. Environmental Science & Technology Letters, 2017. **4**: p. 198-204.
77. Bird, D.L., et al., *Steady-state land cover but non-steady-state major ion chemistry in urban streams*. Environmental Science & Technology, 2018. **52**(22): p. 13015-13026.
78. Olson, J.R. and S.M. Cormier, *Modeling spatial and temporal variation in natural background specific conductivity*. Environmental Science & Technology, 2019. **53**(8): p. 4316-4325.
79. Mount, D.R., et al., *Statistical models to predict the toxicity of major ions to Ceriodaphnia dubia, Daphnia magna and Pimephales promelas (fathead minnows)*. Environmental Toxicology and Chemistry, 1997. **16**(10): p. 2009-2019.
80. Cormier, S.M., et al., *Assessing causation of the extirpation of stream macroinvertebrates by a mixture of ions*. Environmental Toxicology and Chemistry, 2013. **32**(2): p. 277-287.
81. Clements, W.H. and C. Kotalik, *Effects of major ions on natural benthic communities: an experimental assessment of the US Environmental Protection Agency aquatic life benchmark for conductivity*. Freshwater Science, 2016. **35**(1): p. 126-138.
82. Mount, D.R., et al., *The acute toxicity of major ion salts to Ceriodaphnia dubia: I. influence of background water chemistry*. Environmental Toxicology and Chemistry, 2016. **35**(12): p. 3039-3057.
83. Cormier, S.M., L. Zheng, and C.M. Flaherty, *A field-based model of the relationship between extirpation of salt-intolerant benthic invertebrates and background conductivity*. Science of the Total Environment, 2018. **633**: p. 1629-1636.
84. Blaszczak, J.R., et al., *Watershed urban development controls on urban streamwater chemistry variability*. Biogeochemistry, 2019. **144**: p. 61-84.
85. Fanelli, R.M., K.L. Prestegard, and M.A. Palmer, *Urban legacies: Aquatic stressors and low aquatic biodiversity persist despite implementation of regenerative stormwater conveyance systems*. Freshwater Science, 2019. **38**.
86. Kaushal, S.S. and K.T. Belt, *The urban watershed continuum: Evolving spatial and temporal dimensions*. Urban Ecosystems, 2012. **15**(2): p. 409-435.

87. Filoso, S., et al., *The Efficacy of Constructed Stream-Wetland Complexes at Reducing the Flux of Suspended Solids to Chesapeake Bay*. Environmental Science and Technology, 2015. **49**(15): p. 8986-8994.
88. Koskelo, A.I., et al., *Biogeochemical storm response in agricultural watersheds of the Choptank River Basin, Delmarva Peninsula, USA*. Biogeochemistry, 2018. **139**(3): p. 215-239.
89. Jost, L., *Entropy and diversity*. Oikos, 2006. **113**(2): p. 363-375.
90. Hamilton, S.K., *Biogeochemical time lags may delay responses of streams to ecological restoration*. Freshwater Biology, 2012. **57**(SUPPL. 1): p. 43-57.
91. Thompson, J., et al., *The multiscale effects of stream restoration on water quality*. Ecological Engineering, 2018. **124**: p. 7-18.
92. Schueler, T. and B. Stack, *Recommendations of the expert panel to define removal rates for individual stream restoration projects*. 2013, <https://owl.cwp.org/mdocs-posts/final-recommendations-of-the-expert-panel-to-define-removal-rates-for-individual-stream-restoration-projects/> (accessed September 2020).
93. Langland, M.J., et al., *Effects of legacy sediment removal on nutrients and sediment in Big Spring Run, Lancaster County, Pennsylvania, 2009–15*. 2020, U.S. Geological Survey. p. 40.
94. Mattern, K., et al., *Stream restoration for legacy sediments at Gramies Run, Maryland: Early lessons from implementation, water quality monitoring, and soil health*. Water (Switzerland), 2020. **12**(8).
95. Forshay, K.J., et al., *H51D-07: Nitrate decrease in surface and groundwater after legacy sediment removal restoration in a floodplain stream, Big Spring Run, PA USA*, in *Fall American Geophysical Union Meeting*. 2018: Washington, DC.
96. Kaushal, S.S., et al., *Effects of stream restoration on denitrification in an urbanizing watershed*. Ecological Applications, 2008. **18**(3): p. 789-804.
97. Meals, D.W., S.A. Dressing, and T.E. Davenport, *Lag time in water quality response to best management practices: A review*. Journal of Environmental Quality, 2010. **39**(1): p. 85-96.
98. Van Meter, K.J. and N.B. Basu, *Catchment legacies and time lags: A parsimonious watershed model to predict the effects of legacy storage on nitrogen export*. PLoS ONE, 2015. **10**(5).
99. Van Meter, K.J., N.B. Basu, and P. Van Cappellen, *Two centuries of nitrogen dynamics: Legacy sources and sinks in the Mississippi and Susquehanna River Basins*. Global Biogeochemical Cycles, 2017. **31**(1): p. 2-23.
100. Doll, B.A., et al., *Flood flow frequency analysis to estimate potential floodplain nitrogen treatment during overbank flow events in urban stream restoration projects*. Water (Switzerland), 2020. **12**(6).
101. Violin, C.R., et al., *Effects of urbanization and urban stream restoration on the physical and biological structure of stream ecosystems*. Ecological Applications, 2011. **21**(6): p. 1932-1949.
102. Lennox, M.S., et al., *Development of vegetation and aquatic habitat in restored riparian sites of California's North Coast rangelands*. Restoration Ecology, 2011. **19**(2): p. 225-233.

103. McMillan, S.K., et al., *Influence of restoration age and riparian vegetation on reach-scale nutrient retention in restored urban streams*. Journal of the American Water Resources Association, 2014. **50**(3): p. 626-638.
104. Dietrich, A.L., C. Nilsson, and R. Jansson, *A phytometer study evaluating the effects of stream restoration on riparian vegetation*. Ecohydrology, 2016. **9**(4): p. 646-658.
105. Gurnell, A.M., et al., *Seed bank, seed dispersal and vegetation cover: Colonization along a newly-created river channel*. Journal of Vegetation Science, 2006. **17**(5): p. 665-674.
106. Goodson, J.M., et al., *Riparian seed banks: structure, process and implications for riparian management*. Progress in Physical Geography, 2001. **25**(3): p. 301-325.
107. Vosse, S., et al., *Can riparian seed banks initiate restoration after alien plant invasion? Evidence from the Western Cape, South Africa*. South African Journal of Botany, 2008. **74**(3): p. 432-444.
108. O'Donnell, J., K.A. Fryirs, and M.R. Leishman, *Seed banks as a source of vegetation regeneration to support the recovery of degraded rivers: A comparison of river reaches of varying condition*. Science of The Total Environment, 2016. **542**: p. 591-602.
109. MacDougall, A.S. and R. Turkington, *Are Invasive Species the Drivers or Passengers of Change in Degraded Ecosystems?* Ecology, 2005. **86**(1): p. 42-55.
110. Alpert, P., E. Bone, and C. Holzapel, *Invasiveness, invasibility and the role of environmental stress in the spread of non-native plants*. Perspectives in Plant Ecology, Evolution and Systematics, 2000. **1**: p. 52-66.
111. Kuglerová, L., K. Botková, and R. Jansson, *Responses of riparian plants to habitat changes following restoration of channelized streams*. Ecohydrology, 2017. **10**(1).
112. McCarty, G.W., S. Mookherji, and J.T. Angier, *Characterization of denitrification activity in zones of groundwater exfiltration within a riparian wetland ecosystem*. Biology and Fertility of Soils, 2007. **43**(6): p. 691-698.
113. McMahan, P., et al., *Water quality and restoration by legacy sediment removal / floodplain reconnection*. In Prep., 2020.

Supplementary Information

1. Methods Expanded

1.1. Site description

1.1.1. Ecotone, Inc. restoration approach: Ecotone Inc. restored the six restoration study sites with a legacy sediment removal and/or floodplain reconnection (LSR/FR) approach. Post-restoration, all sites were single-channel, meandering streams at low discharge. All restored sites had low $< \sim 0.5$ m banks on both sides of the channels, often with off-channel depressions to capture water. Where possible Ecotone uses materials from the site and tries to use as little off-site material as possible. Most engineered features such as cross veins, sills, and toewood were constructed with wood (again, mostly from on-site at each restoration) with very little rock visible at any site except the highly urban PTR site. Trees removed during excavation of legacy sediments were commonly used as toewood on the outside of meander bends with root balls projecting into the stream. A substantial amount of woody debris from removed trees was placed in the floodplains next to the restored portions of the streams to provide roughness during flooding events and habitat. Another creative way that Ecotone made use of on-site materials was at First Mine Branch where they began growing sod a year previous to the completion of the restoration. When the restoration was complete, the sod was cut and then placed adjacent to the channel to reduce erosion while planted vegetation became established.

1.1.2. First Mine Branch (FMB): This site had 732 m of channel restored in May and June 2017. It is primarily agricultural with the remainder being mostly forest (Table 1). Beyond the floodplain on one side of the stream is active row crop agriculture parallel to much of the study reach along with facilities for an outdoor recreation club. Beyond the floodplain on the other side of the stream is largely forested. Most of the watershed upstream of the restored site is row crops or pasture/hay (Figure S1). Before restoration, FMB had a substantial riparian forest buffer on both sides that sat atop an incised floodplain with banks approximately 0.5–2 meters in height (Figure S2). Post-restoration banks were ~ 0.5 m in height with several depressions created parallel to the channel just beyond the levees. Toe wood was installed around meanders, and log vanes were constructed instream. Several low-radius meanders were removed during restoration and, as a result and somewhat unusually, the channel length was slightly shortened during restoration. Four oxbow wetlands were created in the high-radius portions of the pre-restoration channel, which were in the post-restoration floodplain.

1.1.3. Cabbage Run (CBR): CBR was restored in 2014 along 408 m of channel. This watershed contains a mix of row crops, pasture/hay, and low-density development (Figure S1, Table 1). Approximately 0 – 1 m of bank and floodplain sediments were removed, and several meanders were armored with livestakes and toewood, producing a mean bankful depth of ~ 0.5 m. Following sediment removal and restoration, cultivated crops (row crop corn) were planted within a few meters of the stream channel. Sand deposits on the agricultural fields give qualitative evidence of the floodplain being inundated at least occasionally.

1.1.4. Beetree Run (BTR): BTR was restored in 2016 along 1621 linear meters of channel. It contains a mix of agriculture (row crop and pasture/hay), low-density development, and forest (Figure S1, Table 1). Restoration involved removal of legacy sediments in select areas, armoring

of meanders with toewood, and installation of instream log vanes. Evidence of bank collapse was present just upstream and downstream of the restored reach. One side of the restored reach is classified as cultivated crops though little agricultural activity occurred adjacent to the restored reach during the study period.

1.1.5. North Stirrup Run (NSR): NSR was restored along 792 m of channel in 2015 with a similar approach to Beetree Run. The watershed contains a mix of forest, cultivated crops, and pasture/hay. The restored reach is on the grounds of a property that has been farmed by a single family for centuries. Active pasture lies on either side of the restored reach with several animal crossing. The restored reach has a relatively narrow (<10 m) riparian buffer that is primarily herbaceous.

1.1.6. Bear Cabin Branch (BCB): BCB was restored along 1120 m of channel from Dec. 2017 – Apr. 2018. Prior to restoration, the land adjacent to the study reach was predominantly grassland with scattered trees, and the stream was deeply incised with banks approaching 2.5 meters in places. A buried legacy stone dam was visible in the study reach at the base of legacy sediments. During restoration, 1.5–2 m of bank and floodplain sediments were removed, meanders were armored with toewood, several log veins were installed, the channel was raised, and >10 acres of wetlands were created or restored. Approximately half of the watershed is developed (up to medium intensity) with the pre-urban development land use being primarily agricultural (Figure S1). The remainder of the watershed is a mix of cultivated crops, forest, and pasture/hay. A housing development was under construction adjacent to the upstream location during the early part of the study period. See pre- and post-restoration site pictures (Figure S2).

1.1.7. Plumtree Run (PTR): PTR was restored during the study period along 377 meters of channel. It is a highly urban watershed with mostly medium- and high-density development (Figure S1; Table 1). The site is located between a secondary road and a school with the upstream end coming out of a culvert with low but continuous discharge and the downstream end draining into three culverts under a road and then resurfacing downstream of the road. Pre-restoration, the site had a relatively mature forest buffer up to 10 m in width and banks up to 2.5 m high in places. A portion of the stream approximately 400 m upstream of the study site was restored several years previously and is currently has little discharge at baseflow conditions. Reconstruction removed sediments leaving ~0.5 m high, gently sloped banks, meanders armored with toewood, several in-stream log veins, and a floodplain wetland parallel to the channel. A sewage leak occurred upstream of the restored reach during the pre-restoration period though a signal of the leak is not detectable in the study data. See pre- and post-restoration site pictures (Figure S2).

1.1.8. Regional reference sites: The forested Pond Branch (POBR) and predominantly forested Baisman Run (BARN) regional reference sites are mostly located in a county park (Figure S1, Table 1). POBR is a small 100% forested stream completely in the park that drains into BARN. BARN is mostly forested with low density development and low pasture in upper reaches. The site 1700 m downstream of the highly urban PTR restoration site is Plumtree Gage (PTRG). PTRG is largely developed with medium to high-density housing/commercial businesses (Figure S1, Table 1).

1.2. Stage-discharge at First Mine Branch

The rating curves for First Mine Branch (FMB) were established by the correlation of stage with discharge as measured with area-velocity or salt dilution methods. The stage was determined with staff plates that were installed at the upstream and downstream sampling locations. Staff plates were installed in stream reaches with relatively stable geomorphology: riffles with gravel to rocky bottoms with minimal deposition or erosion. Any sediments deposited by the staff gauges was removed during site visits. During stormflow events, three stage measurements were taken: before, midway through, and after the discharge measurement by the area-velocity method. Stage changes that exceeded an increase of 0.03 m while discharge was being measured with the area-velocity method were not used in the creation of the rating curve. The upstream and downstream stage-discharge data were fit with a logarithmic regression to non-log-transformed data to establish rating curves ($R^2 = 0.80$ upstream, 0.90 downstream). Predicted and actual discharge matched well (Figure S8). High-frequency stage data were collected every five minutes from April 2019 through June 2020 using Onset Hobo U20 pressure sensors along with a third sensor for barometric pressure. The high-frequency stage data do have some gaps, particularly in late fall 2019, due to a sensor failure. Discharge was determined at the upstream and downstream sites using the rating curves.

1.3. US Geological Survey high-frequency data time series at Plumtree Gage (PTRG) downstream of the highly urban (PTR) restored site

All US Geological Survey (USGS) data were downloaded from the USGS National Water Information System (NWIS) database using the *dataRetrieval* library in R [50] for the period January 2014 through February 2020. Discrete water quality and high-frequency discharge and water quality data came from the Plumtree Run gage (#01581752). Discrete data used were nitrate (NO_3^-), Total Dissolved Phosphorus (TDP), and Total Suspended Sediments (TSS) concentrations. High-frequency data downloaded were 5-minute discharge and turbidity data as well as daily mean discharge and turbidity data. USGS precipitation data (15-minute interval) are from a site approximately 1700 m away (#01581753 Atkisson Reservoir).

Gaps in the high-frequency discharge and turbidity time series data were filled with estimated values. Linear interpolation was used to fill gaps of ≤ 1 hour, and daily mean values were used for gaps >1 hour. From 2014–2020, 104 days lacked daily mean turbidity values. For those days, the daily mean turbidity was estimated using a regression model using daily mean discharge as the predictor variable ($R^2 = 0.72$). The turbidity record had five gaps ≥ 5 days: July 3–8, 2014, Aug. 5–10, 2017, Aug. 16 – Sept. 11, 2017, Feb. 25 – Mar. 8, 2018, and Apr. 15–22, 2019.

High-frequency time series were developed for NO_3^- , TSS, and TDP concentrations based on regression relationships. NO_3^- concentrations were negatively correlated with discharge ($R^2 = 0.71$, Figure S7b) with concentrations decreasing as discharge increased. Turbidity values were highly correlated with TSS and TDP ($R^2 = 0.87$ and 0.90 , respectively, Figure S7c, d). Loads for NO_3^- , TSS, and TDP were calculated for each 5-minute interval and then summed to produce daily loads using the *dplyr* package in R [52]. Total dissolved nitrogen (TDN) and

orthophosphate did not correlate with discharge, specific conductance, or turbidity, so high-frequency time series could not be constructed

Individual precipitation events were identified as periods of recorded rainfall where rainfall was either continuous or, if intermittent, intervals with no recorded rainfall did not exceed four consecutive hours. For each event, the duration, hourly rate of precipitation, and total precipitation were calculated. Six precipitation events prior to the restoration period (July – October 2017) were paired with six events with similar precipitation rates and totals following the restoration period. The hydrological response to each precipitation event was identified as the period when discharge initially increased from baseflow conditions to the time when discharge returned to pre-event or inter-event values. NO_3^- and TSS loads were summed for each event.

2. Results Expanded

2.1. Baseflow non-nutrient chemistry expanded

Specific conductance (SC) values are driven by dissolved ion concentrations and cluster into three groups (Figure 5a): 1) lowest ($28 \mu\text{S}/\text{cm}$) at the forested watershed (POBR); 2) about an order of magnitude higher ($178\text{--}279 \mu\text{S}/\text{cm}$) at the low-density suburban watershed (BARN) and the restored agricultural and suburban sites (FMR, CBR, BTR, BCB); and 3) substantially higher ($\sim 900 \mu\text{S}/\text{cm}$) at the restored highly urban site (PTR). The ions driving elevated SC values differ among sites. Median chloride concentrations in the suburban and agricultural sites are at least an order of magnitude higher than the forested site, and chloride at the highly urban site (PTR) is approximately another order of magnitude higher (Figure 5b);

Thus, downstream values are reported for cross-site comparisons of non-nutrients. The DO and pH are similar among sites with median DO values of 94–110% (Figure S4a) and baseflow pH values of 6.86–7.58 (not shown). Median DO at the reference and agricultural sites was 94–100% and at the suburban and urban restored sites (BCB, PTR) was 103–110% with few observations <90% except for the downstream end of the urban PTR site.

Median specific conductance (SC) values are driven by dissolved ion concentrations and cluster into three groups (Figure 5a): 1) lowest ($28 \mu\text{S}/\text{cm}$) at the forested watershed (POBR); 2) about an order of magnitude higher ($178\text{--}279 \mu\text{S}/\text{cm}$) at the low-density suburban watershed (BARN) and the restored agricultural and suburban sites (FMR, CBR, BTR, BCB); and 3) substantially higher ($\sim 900 \mu\text{S}/\text{cm}$) at the restored highly urban site (PTR). The ions driving elevated SC values differ among sites. Median chloride concentrations in the suburban and agricultural sites are at least an order of magnitude higher than the forested site, and chloride at the highly urban site (PTR) is approximately another order of magnitude higher (Figure 5b); sodium concentrations behave similarly though the differences are somewhat smaller (not shown). The highly urban PTR restored site had the highest concentrations of most other ions, including HCO_3^- and K^+ (Figure S4b, d) and Ca^{2+} (Figure S4c), followed by the highly agricultural FMB site. Elevated SC values and ion concentrations at agricultural and urban sites result from inputs from fertilizer and other agricultural amendments such as lime (K^+ , NO_3^- , Ca^{2+} , HCO_3^-), septic systems (Na^+ , NO_3^-), deicing salts (Na^+ , Cl^-), and breakdown of materials such as concrete (Ca^{2+} , HCO_3^-).

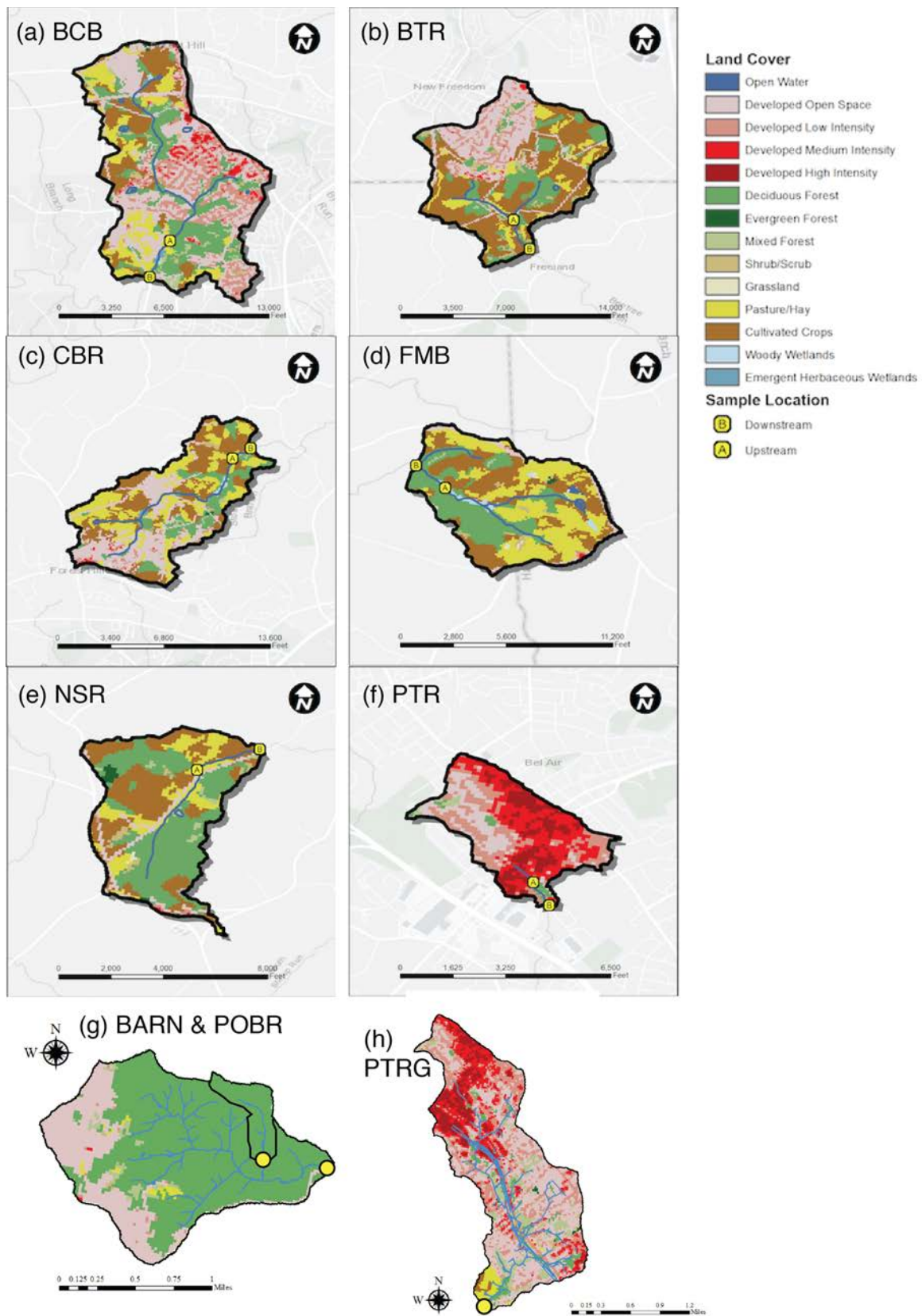


Figure S1. Land use maps for all study watersheds based on 2011 NLCD [45].



Figure S2. Site pictures

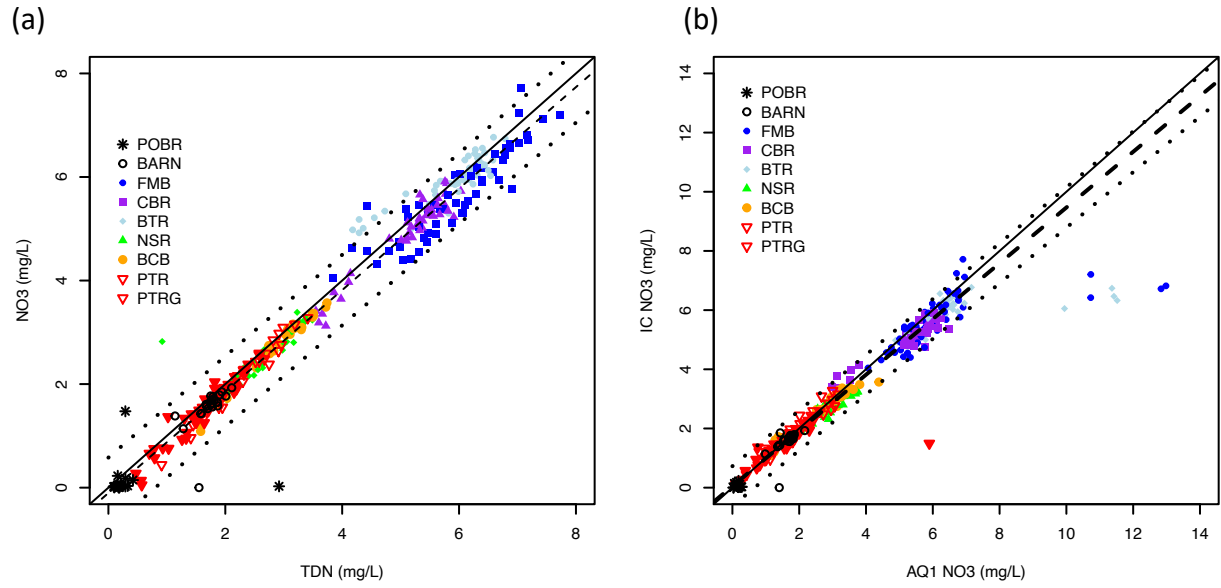


Figure S3. (a) NO₃⁻ as measured by ion chromatograph (IC) versus TDN concentrations for all sites. A 1:1 line is plotted for reference. Note that NO₃⁻ is below the IC detection limit (<0.4–0.5 mg/L) at the forested reference site POBR with NO₃⁻ plotted as zero/near-zero for that site. (b) NO₃⁻ as measured by ion chromatograph versus as measured colorimetrically with Seal AQ1 Discrete analyzer. Nitrate at the POBR forested site also is typically below the detection limit (0.05 mg/L) for the AQ1 analyzer.

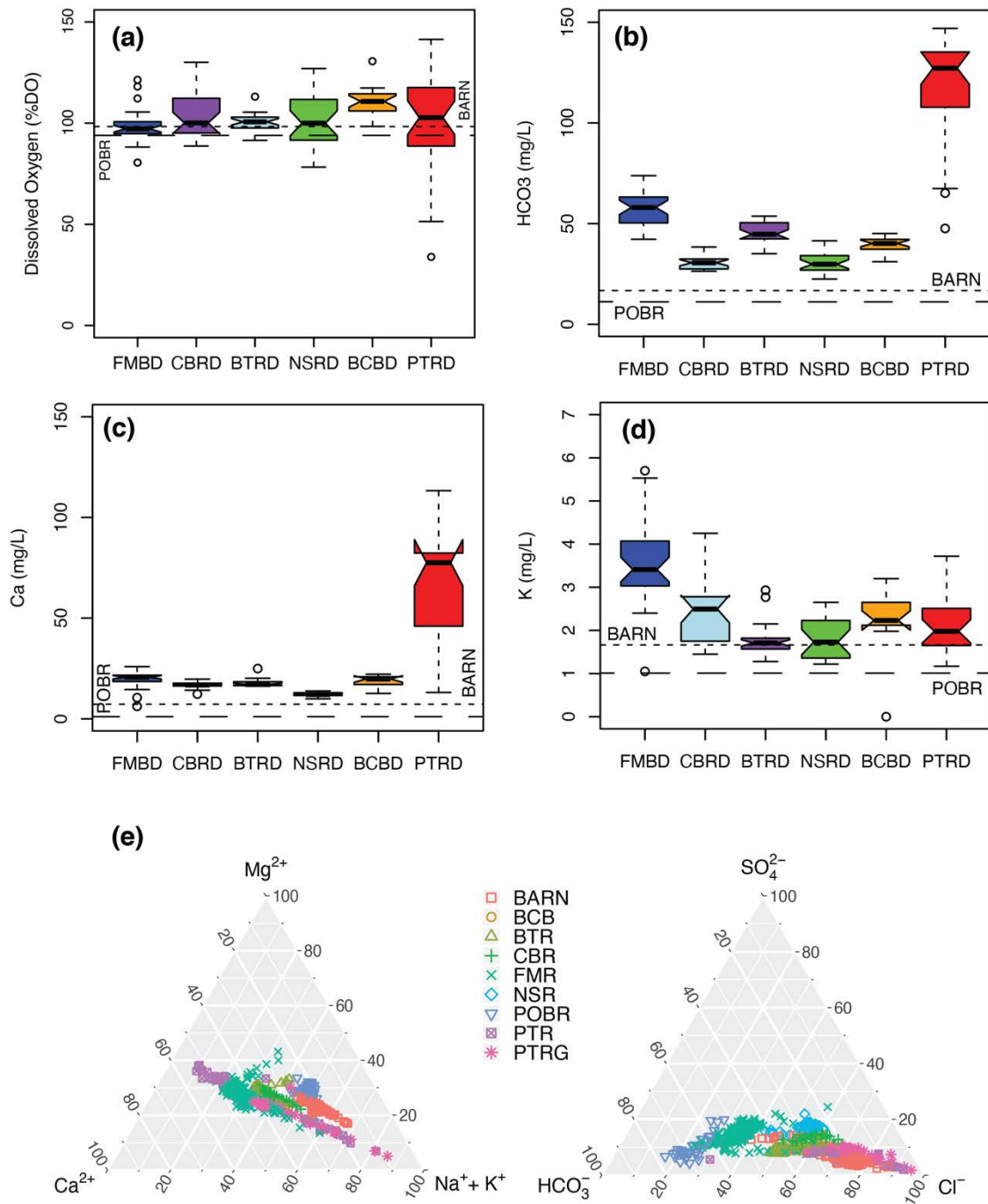
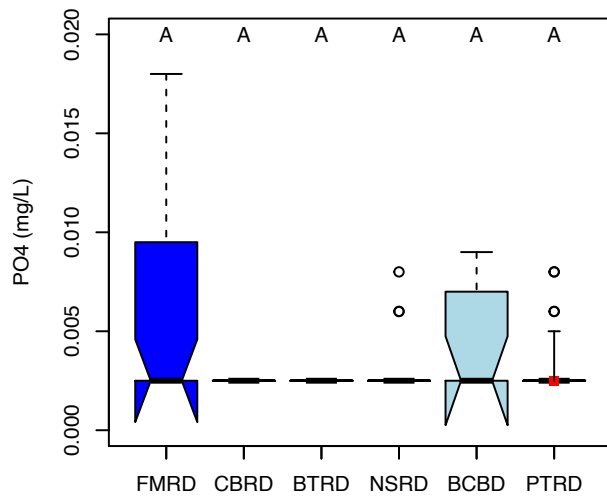


Figure S4: (a) dissolved oxygen (DO), (b) bicarbonate/alkalinity, (c) calcium, (d) potassium, and (e) ternary diagrams based on the total contribution of the positive or negative charge ($\mu\text{eq/L}$) contributed by cations or anions, respectively for each of restored sites. Lines show median values for the regional reference sites.

(a)



(b)

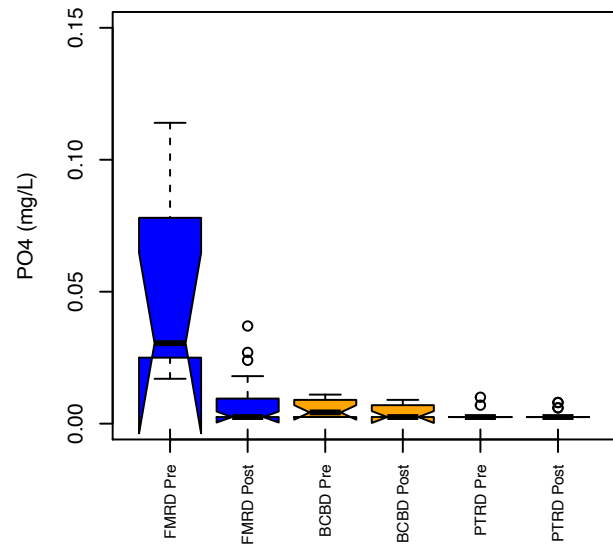


Figure S5. (a) Post-restoration orthophosphate concentrations at the downstream ends of the restored sites. The median oP values for the CBR, BTR, NSR, PTR as well as at the forested and mostly forested regional reference sites are 0.0025 mg/L, which is 50% of the 0.005 mg/L detection limit established for the method. (b) Pre- and post-restoration orthophosphate concentrations at the downstream ends of the three sites with pre-restoration data. Note the different y-axis scale. Only FMB had a significant change (decrease) in concentrations after restoration (Table S6). Median values for all sites are reported in Table S6 along with some total dissolved phosphorus data.

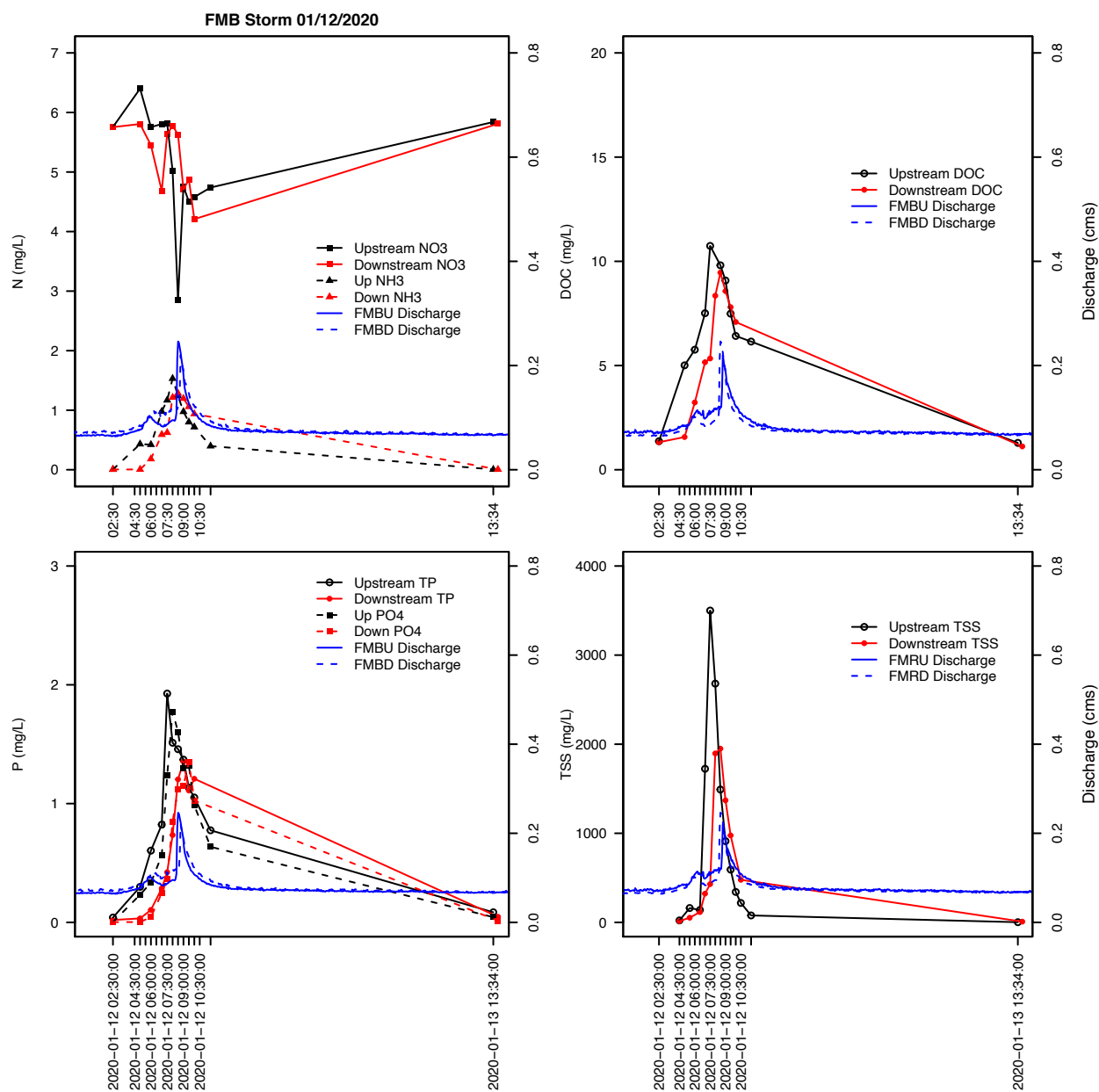


Figure S6, Part 1. January 12, 2020 storm data for discharge and concentrations of (a) nitrogen, (b) dissolved organic carbon, (c) dissolved phosphorus [TP = total dissolved phosphorus], and (d) total suspended sediments.

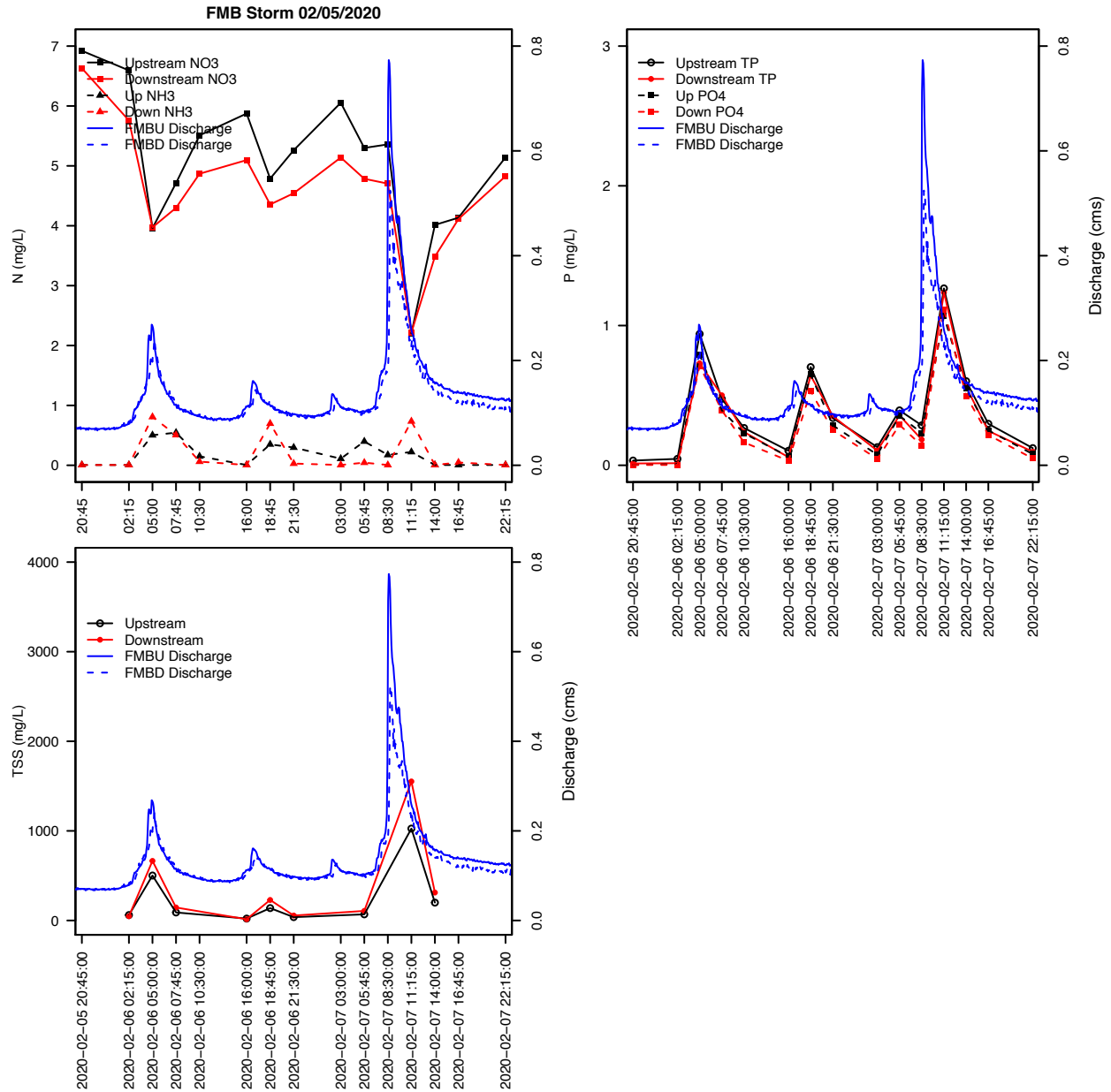


Figure S6, Part 2. February 5, 2020 storm data for discharge and concentrations of (a) nitrogen, (g) dissolved phosphorus [TP = total dissolved phosphorus], and (c) total suspended sediments.

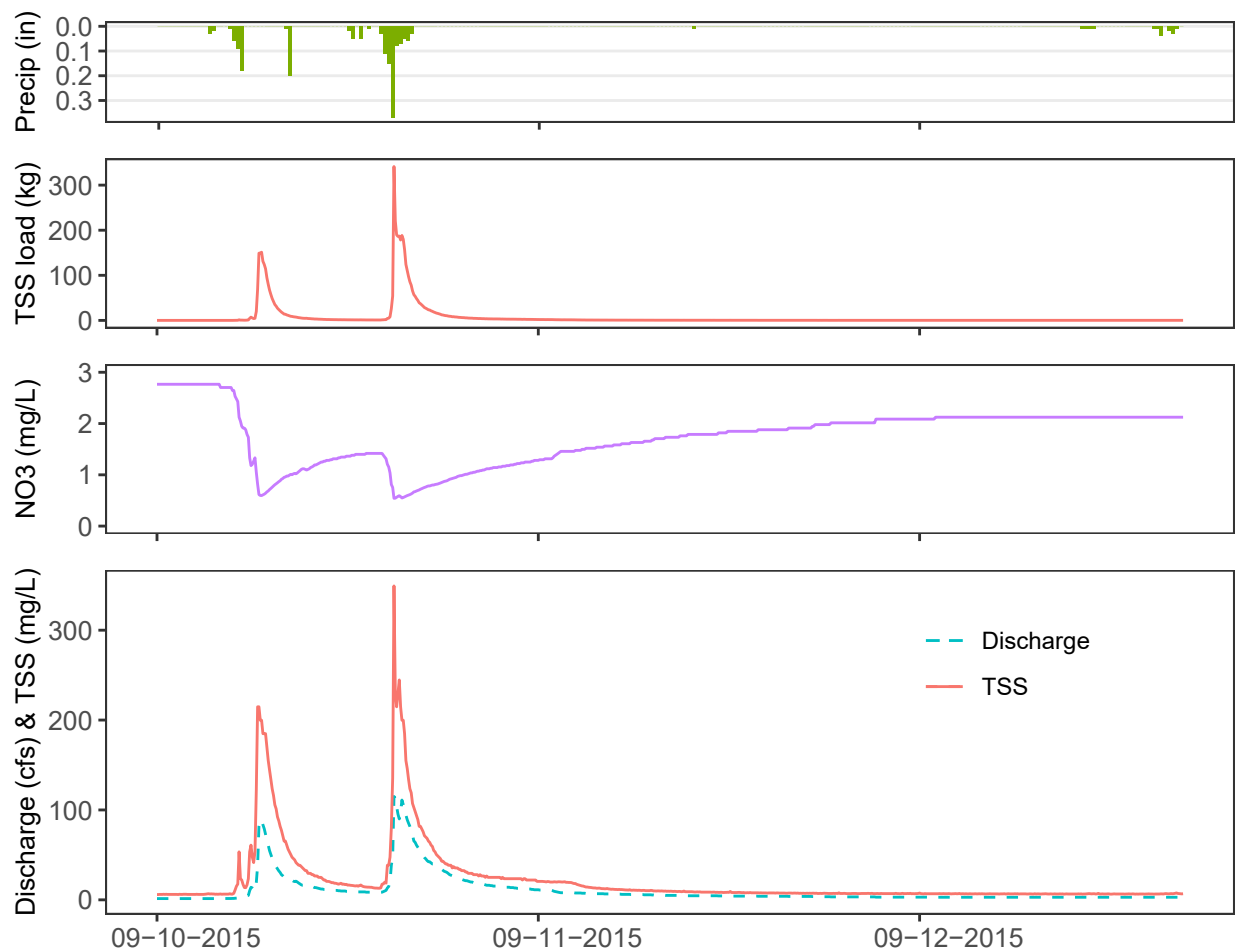


Figure S7: Example storm event at Plumtree Run gage with USGS data for precipitation, TSS load (calculated from regression between discrete sample data for TSS concentrations and high-frequency turbidity data), NO₃ concentrations, and discharge and TSS concentrations.

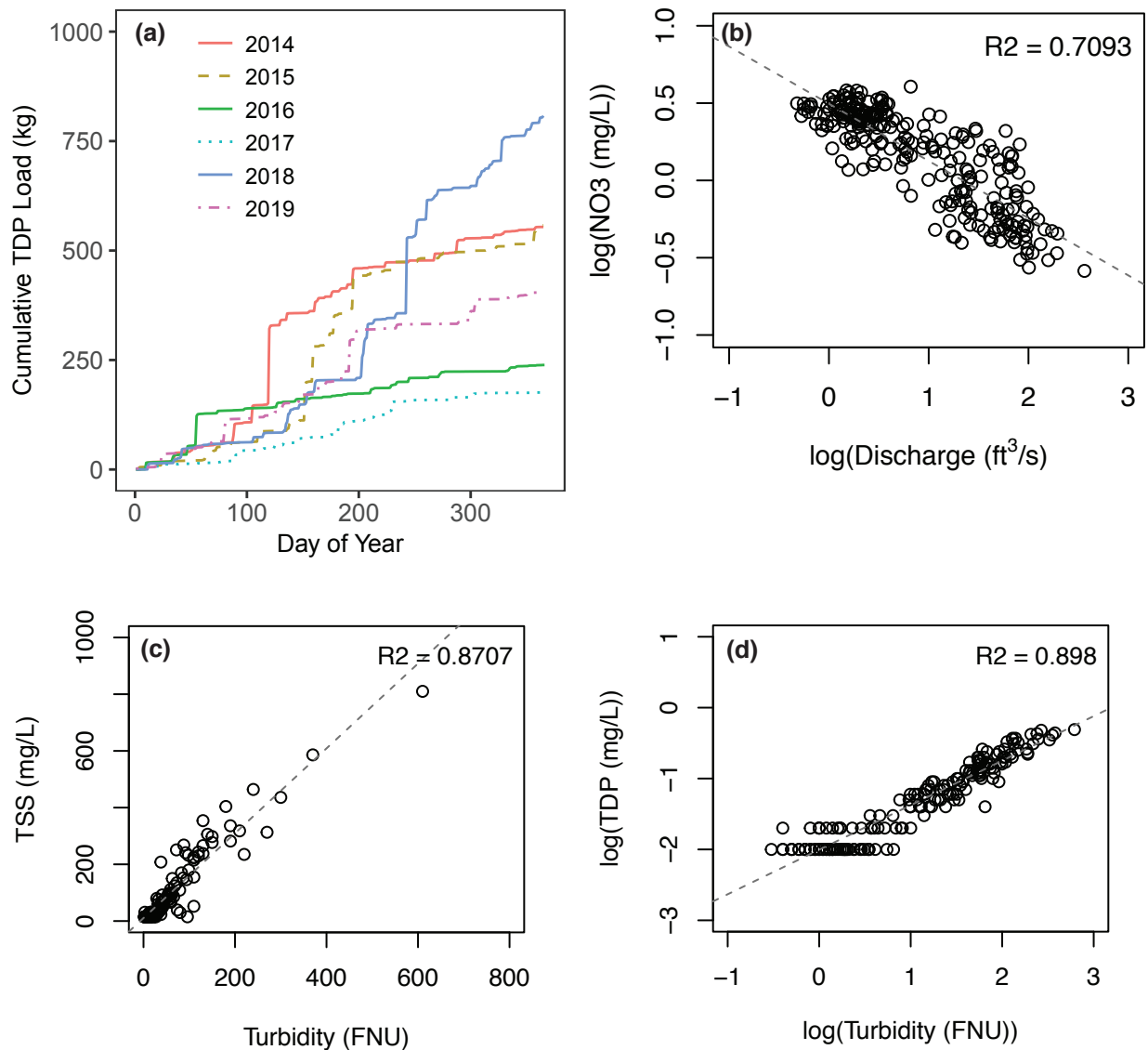


Figure S8: Data from the Plumtree Run gage (PTRG) for (a) cumulative total dissolved phosphorus loads, (b) log discrete nitrate concentrations versus discharge, (c) total discrete suspended sediments concentrations versus high-frequency turbidity values, (d) log total discrete dissolved phosphorus concentrations versus log high-frequency turbidity values.

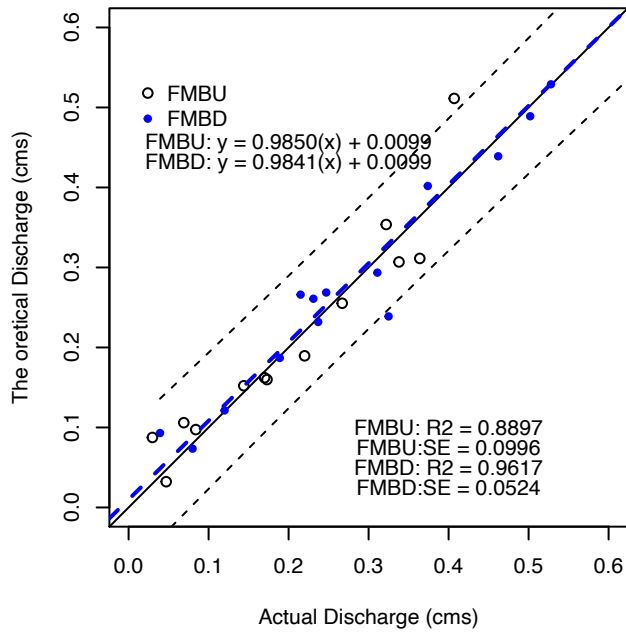


Figure S9: Predicted (or theoretical) versus actual discharge values

Table S1: p values for significant relationships

Plot / Relationship	FMB ^a	CBR	BTR	NSR	BCB	PTR
<i>Figure 2 – Post-restoration</i>						
<i>Diff Down–Up Post TDN (mg/L)</i>	<0.001↓	<0.001↑	<0.001↓	<0.001↑	<0.001↓	<0.001↓
<i>Diff Down–Up Post TDN load (g/km²)</i>	–	–	–	–	0.001↓	<0.001↓
<i>Figure 3 – Pre vs Post-restoration</i>						
Upstream TDN (mg/L) <i>Pre vs post</i>	–				0.010↑	–
Downstream TDN (mg/L) <i>Pre vs post</i>	–				0.006↑	–
<i>Diff Down–Up Pre TDN (mg/L)</i>	0.016↓				<0.001↓	0.008↓
Upstream TDN load (g/km ²) <i>Pre vs post</i>	0.004↑				<0.001↑	–
Downstream TDN load (g/km ²) <i>Pre vs post</i>	–				<0.001↑	–
<i>Diff Down–Up Pre TDN load (g/km²)</i>	0.031↑				0.039↓	0.016↓

^a – is >0.05; blank is not measured/applicable; ↓ and ↑ indicate a decrease or increase, respectively

Table S2: Post-restoration baseflow data

Site	Median TDN	Median TDN daily load	Median DOC	Median DOC daily load	DOC:TDN		TSS ⁻	
	mg/L	g/km ²	mg/L	g/d/km ²	mM/mM	N	mg/L	N
FMBU	6.40 ^a	6936	1.78	1592	0.31 ^b	33	2.82 ^a	31
FMBD	6.01	6343	1.43	2234	0.36	33	6.10	31
CBRU	5.30 ^a	7832	1.07	1820	0.24	26	3.04	24
CBRD	5.37	8935	1.02	1606	0.20	26	1.90	24
BTRU	6.13 ^a	9448	0.92	1464	0.19	29	4.16	26
BTRD	5.90	8771	0.95	1339	0.19	29	3.81	27
NSRU	2.80 ^a	3425	1.16 ^a	1801	0.45	26	2.04 ^a	24
NSRD	2.85	4082	1.32	1680	0.54	26	2.00	24
BCBU	2.96 ^a	4311 ^a	1.55 ^a	2270	0.63 ^a	22	1.06 ^a	19
BCBD	2.72	3062	1.78	1664	0.69	22	3.05	19
PTRU	1.91 ^a	712.8 ^a	2.09 ^a	766.5 ^a	1.40 ^a	29	2.13	28
PTRD	1.42	477.5	2.65	815.4	2.64	29	3.15	28
POBR	0.19	216.6	1.77	1179	11.47	36	4.27	20
BARN	1.78	2134	1.02	1490	0.62	31	2.99	25
PTRG	2.52	1877	2.05	1580	0.93	45	0.84	36

^a Indicates statistically significant differences. See Table S1 for TDN p statistics.

[DOC] post-restoration difference up–down: NSR p = 0.021, BCB p = 0.004, PTR <0.001

DOC load post-restoration load difference up–down: PTR = 0.019

DOC:TDN post-restoration difference up–down: FMB = 0.001, BCB & PTR <0.001

Note orthophosphate (PO₄³⁻) data are found in Table S6 for baseflow and storms.

Table S3: Annual precipitation & mean streamflow

Calendar Year / Period	<i>Precip.</i>		<i>Discharge</i>	
	BWI m/yr	POBR m/yr	BARN m/yr	PTRG m/yr
2014	1.34	0.610	0.622	0.772
2015	1.30	0.405	0.399	0.813
2016	1.03	0.480	0.397	0.490
2017	0.97	0.196	0.201	0.407
2018	1.82	0.526	0.631	0.981
2019	0.97	0.771	0.589	0.674
<i>Long-term mean values</i>	<i>1.06</i>	<i>0.450</i>	<i>0.412</i>	<i>0.640</i>
<i>Period for Mean</i>	<i>1981–2010</i>	<i>1999–2019</i>	<i>2000–2019</i>	<i>2002–2019</i>

Data sources: BWI (<https://www.weather.gov/media/lwx/climate/bwiprecip.pdf>)
 POBR https://waterdata.usgs.gov/nwis/uv/?site_no=01583570&agency_cd=USGS
 BARN <https://waterdata.usgs.gov/usa/nwis/uv?01583580>
 PTRG https://waterdata.usgs.gov/md/nwis/uv/?site_no=01581752&agency_cd=USGS

Table S4: Median baseflow discharge (Q)

Site	2017 Median Q	N	2017 Median Q	N	2018 Median Q	N	Pre- restoratio n	N	Post- restoratio n	N
	m ³ /s		m ³ /s		m ³ /s		mg/L			
FMBU	0.024^a	18	0.040^a	17	0.065 ^a	9	0.028 ^a	12	0.037 ^a	33
FMBD	0.031	18	0.061	15	0.111	9	0.038	12	0.052	33
CBRU	0.043	2	0.077 ^a	15	0.097 ^a	9	–	–	–	–
CBRD	0.047	2	0.088	17	0.125	9	–	–	–	–
BTRU	0.069	7	0.086^a	18	0.167	9	–	–	–	–
BTRD	0.046	6	0.130	18	0.175	9	–	–	–	–
NSRU	0.010	2	0.025^a	16	0.041	9	–	–	–	–
NSRD	0.015	2	0.033	16	0.040	9	–	–	–	–
BCBU	0.056	16	0.163	11	0.114 ^a	11	0.056 ^b	19	0.129	22
BCBD	0.057	17	0.116	11	0.093	11	0.058 ^b	22	0.106	22
PTRU	0.003	15	0.005	16	0.005	9	0.003	13	0.003	30
PTRD	0.002	16	0.004	17	0.004	9	0.003 ^b	14	0.004	29
POBR	–	–	0.004	14	0.005	19	–	–	–	–
BARN	–	–	0.052	16	0.065	24	–	–	–	–
PTRG	0.0375	12	0.066	18	0.061	20	–	–	–	–

Bold indicates a statistically significant difference ($p < 0.05$) between the indicated year and following year.

^a Indicates a statistically significant difference between the upstream and downstream ends of the reach

^b Indicates a statistically significant difference between pre-restoration and post-restoration

Table S5: Pre- versus post-restoration baseflow differences from upstream to downstream

Site	Pre-restoration			Post-restoration		
	Upstream	Downstream	N	Upstream	Downstream	N
FMB						
TDN (mg/l)	6.17 ^a	5.66	7	6.40 ^a	6.01	33
TDN load (g/km ²)	4235^a	4353		6936	6343	
BCB						
TDN (mg/l)	2.45^a	2.33	16	2.99^a	2.72	21
TDN load (g/km ²)	1542^a	1443		4311^a	3062	
PTR						
TDN (mg/l)	1.85 ^a	1.50	8	1.91 ^a	1.42	29
TDN load (g/km ²)	446 ^a	355		713 ^a	477	

^a Indicates statistically significant differences between the upstream and downstream values. See Table S1 for p statistics.

Bold indicates statistical differences between pre-restoration and post-restoration at the upstream or the downstream sites, respectively. See Table S1 for p statistics.

Table S6: Median orthophosphate (oP, PO₄³⁻) and total dissolved phosphate (TDP)

Site	Baseflow			Stormflow ^c				
	Median PO ₄ ³⁻ mg/L ^a	Total N	BDL ^b N	Median PO ₄ ³⁻	N	BDL N	TDP mg/L	N
<i>Pre-restoration</i>								
FMBU ^d	0.0310	6	0					
FMBD	0.0305	6	0					
BCBU	0.0025	15	8					
BCBD	0.0043	14	7					
PTRU	0.0060	9	4					
PTRD	0.0025	9	7					
<i>Post-restoration</i>								
FMBU ^d	0.0070	27	11	0.4610	77	3	0.5495	76
FMBD	0.0025	28	18	0.2540	60	5	0.3320	73
CBRU	0.0025	14	13					
CBRD	0.0025	14	14					
BTRU	0.0025	18	18					
BTRD	0.0025	17	17					
NSRU	0.0025	14	13					
NSRD	0.0025	15	12					
BCBU	0.0025	9	7					
BCBD	0.0025	10	7					
PTRU	0.0025	19	14					
PTRD	0.0025	18	14					

^a Results are reported for primarily unfrozen and some frozen samples. The unfrozen samples are likely to be somewhat lower than actual due to storage time.

^b BDL = Below detection limit; for median calculations, 0.0025 mg/L was used for BDL samples (50% of the 0.005 mg/L DL)

^c Stormflow samples were frozen until analysis.

^d FMB was the only site with significantly different pre-restoration and post-restoration concentrations: FMBU pre vs post p = 0.040, FMBD pre vs post p <0.001. Additionally, for FMB post-restoration, upstream versus downstream concentrations were significantly different p = 0.003. At FMB, 9 frozen post-restoration samples for upstream and 9 for downstream were run for TDP: 0.0520 and 0.0360 mg/L, respectively. The oP for those same 9 samples was 0.0260 and 0.005 mg/L for upstream and downstream, respectively with 3 BDL for upstream and 2 BDL for downstream.

Note: median oP values for the regional forested (POBR) and mostly forested sites (BARN) were 0.0025 mg/L.

Table S7: Storm versus baseflow concentrations for First Mine Branch (FMB) dissolved nitrogen, carbon, and total suspended sediments

Parameter	Location	Baseflow		Stormflow			
		Median ⁻ mg/L ^a	N	25% mg/L	Median mg/L	75% mg/L	
TSS	Upstream	2.84	31	44.1	164	619	70
	Downstream	6.10	31	53.6	174	616	75
TDN	Upstream	6.40	33	4.01	5.02	5.84	40
	Downstream	6.01	33	3.94	4.55	4.99	40
NO ₃ ⁻	Upstream						
	Downstream	5.93	33	2.98	4.29	4.85	58
DOC	Upstream	1.78	33	4.30	6.67	9.86	61
	Downstream	1.84	33	3.69	6.34	8.68	58

Table S8: PTRG Mean discharge and annual loads

Calendar Year	Mean Discharge m ³ /s	NO ₃ ⁻ kg/yr	TDP kg/yr	TSS kg/yr
2014	0.1584	6617	554.0	588,600
2015	0.1674	7007	557.7	600,300
2016	0.1006	5278	238.7	253,600
2017	0.0832	4533	175.8	152,100
2018	0.2013	7809	807.8	988,800
2019	0.1381	6252	406.7	421,800

Data sources: PTRG https://waterdata.usgs.gov/md/nwis/uv/?site_no=01581752&agency_cd=USGS

Table S9: Storm sets for Plumtree Run (PTRG)

Storm Set	Storm start time	Storm end time	Peak discharge ft ³ /s	Pre- or post-restoration	Total precip. mm	Precip. duration hr	Precip. rate mm/hr	TSS load ^a kg	TDP load ^a kg	NO3 load kg
1	7/30/2015 14:35	8/1/2015 16:05	188	Pre	735.48	0.75	38.61	6590	6.44	56.95
1	6/19/2019 22:10	6/22/2019 22:50	203	Post	716.13	1.50	18.8	9688	9.25	72.76
2	9/10/2015 3:55	9/12/2015 16:35	115	Pre	1083.9	12.75	3.35	5356	6.64	75.62
2	7/27/2018 8:50	7/29/2018 12:25	576	Post	1154.8	13.75	3.31	45224	32.48	89.02
3	11/6/2014 0:10	11/7/2014 6:15	34	Pre	451.61	11.25	1.58	550	1.08	35.60
3	11/7/2017 12:20	11/11/2017 7:10	30.5	Post	496.77	10.75	1.82	607	1.16	50.05
4	7/8/2016 17:15	7/11/2016 14:20	42.7	Pre	703.22	2.75	10.07	1424	1.68	39.85
4	7/17/2019 18:10	7/22/2019 13:35	589	Post	619.35	2.25	10.84	24422	20.65	116.1
5	10/4/2014 0:45	10/5/2014 18:40	25.7	Pre	264.52	8.50	1.23	217	0.41	24.98
5	1/12/2020 1:20	1/14/2020 15:20	48.7	Post	270.97	7.50	1.42	1050	1.56	46.05
6	1/10/2014 8:55	1/13/2014 16:55	139	Pre	806.45	38.75	0.82	11006	11.96	106.4
6	2/5/2020 23:55	2/9/2020 18:35	68.1	Post	793.55	37.25	0.84	3097	4.87	92.37

^a TSS = Total Suspended Sediments, TDP = Total Dissolved Phosphorous