



Restoration Research Summary of Results: Final Report
Arundel Rivers Federation (South River Federation prior to January 2019)
and Smithsonian Environmental Research Center
February 9, 2020

Prepared by Jesse Iliff, Wayne Martin, and Sarah Giordano, ARF
and Thomas E. Jordan, SERC

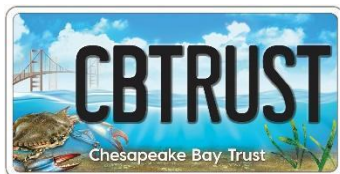
For Restoration Research grant funders:

Chesapeake Bay Trust

Maryland Department of Natural Resources

Montgomery County Department of Environmental Protection

State Highway Association



Project Title:

Assessing Watershed-scale Restoration Effectiveness: Treatment Impacts and Monitoring Requirements

FY 2017 Restoration Research Award Program Questions to be addressed:**A1: What are the cumulative effects of watershed restoration activities within a watershed?**

Specifically: Do best management practices (BMPs) designed to reduce loads of total suspended solids (TSS), total nitrogen (TN) and total phosphorus (TP) from urban watersheds perform as rated for the impervious area being treated? Are the effects of multiple BMPs along a channel simply additive, or could their effects be interactive (could a BMP enhance the effectiveness of a downstream BMP)?

A.3: What are the most efficient methods for measuring effects of stormwater management methods at watershed scales?

Specifically, we tested whether high frequency measurements of flow and turbidity can be used to estimate the fluxes of total suspended solids (TSS), and particulate forms of nitrogen and phosphorus for the purpose of limiting monitoring costs. We also explored the correlations concentration among different nutrient forms and total TSS and correlations among of high frequency measurements of dissolved oxygen, conductivity, temperature, and stream flow. Strong correlations would suggest opportunities to use proxy measurements to reduce redundancy of effort and eliminate some monitoring costs. In addition, we used automated flow-paced sampling as a highly efficient and accurate method for monitoring discharges of nutrients and TSS (Thompson et al. 2020).

Hypotheses to be tested:

1. Remediation efforts are effective in removing Total Suspended Solids (TSS), Nitrogen (N), and Phosphorous (P) from surface water.
2. BMP type and impervious surface percentage in a sub-watershed will affect reductions in the pollution outfall.
3. Correlations between measurements of turbidity and TSS can support the use of turbidity as a proxy measurement for some for TSS and particulate forms of N and P.

Project Summary:

We used automated flow-paced sampling to monitor weekly concentrations and loads of total suspended solids (TSS) and forms of nitrogen (N) and phosphorus (P) in discharges during two years from five watersheds, which included a variety of best management practices (BMPs) in an urban drainage basin. We deployed sondes at the sampling points to record stream temperature, conductivity, dissolved oxygen, and turbidity every 15 minutes. Four of the watersheds overlap in area and include 64-66% impervious surface. One watershed is separate and includes 31% impervious surface. That watershed discharged less water per area than the others. Some of the streams in the watersheds have rock weirs installed at intervals to create series of pools and riffles called regenerative stormwater conveyances (RSCs), which may enhance retention of nutrients and TSS. Comparing points entering and leaving an RSC located between the discharge points of two of the watersheds, we found retention of about 70% of the nitrate entering, while ammonium, phosphate, total phosphorus, and TSS were not retained. Phosphate and total P concentrations did not show clear seasonal patterns, but nitrate concentrations and loads in three of the five study watersheds were lowest in summer highest in winter. Ammonium showed a similar seasonal pattern at one location. The seasonal high concentrations of nitrate and ammonium were similar to concentrations recorded in rain 7 miles from the study watersheds at the Smithsonian Environmental Research Center. The spatial and temporal

patterns of nitrate suggest that biological uptake in the RSCs, may play a role in reducing nitrate concentrations and fluxes during the summer. There were statistically significant differences in the loads of nutrients and TSS among the five watersheds. However we found no correlation between our measured loads and the total credits for removal of total N, total P, and TSS by the BMPs in the watersheds, although the range of removal credits for total N and TSS spanned about 50% of the range of measured loads and the range of removal total P credits spanned about 25% of the range of loads. Nutrient and TSS loads were lowest from the watershed with the least extensive impervious surface mainly because the water discharge was lowest from that watershed. There was some correlation of temporal variability among concentrations of different substances measured in weekly sampling (e.g. TSS and total P) and among parameters measured by the sondes, but none of the correlations were strong enough to suggest redundancy of measurements. We were unable to obtain consistent enough correlations between TSS and turbidity to use turbidity as a surrogate measure of TSS.

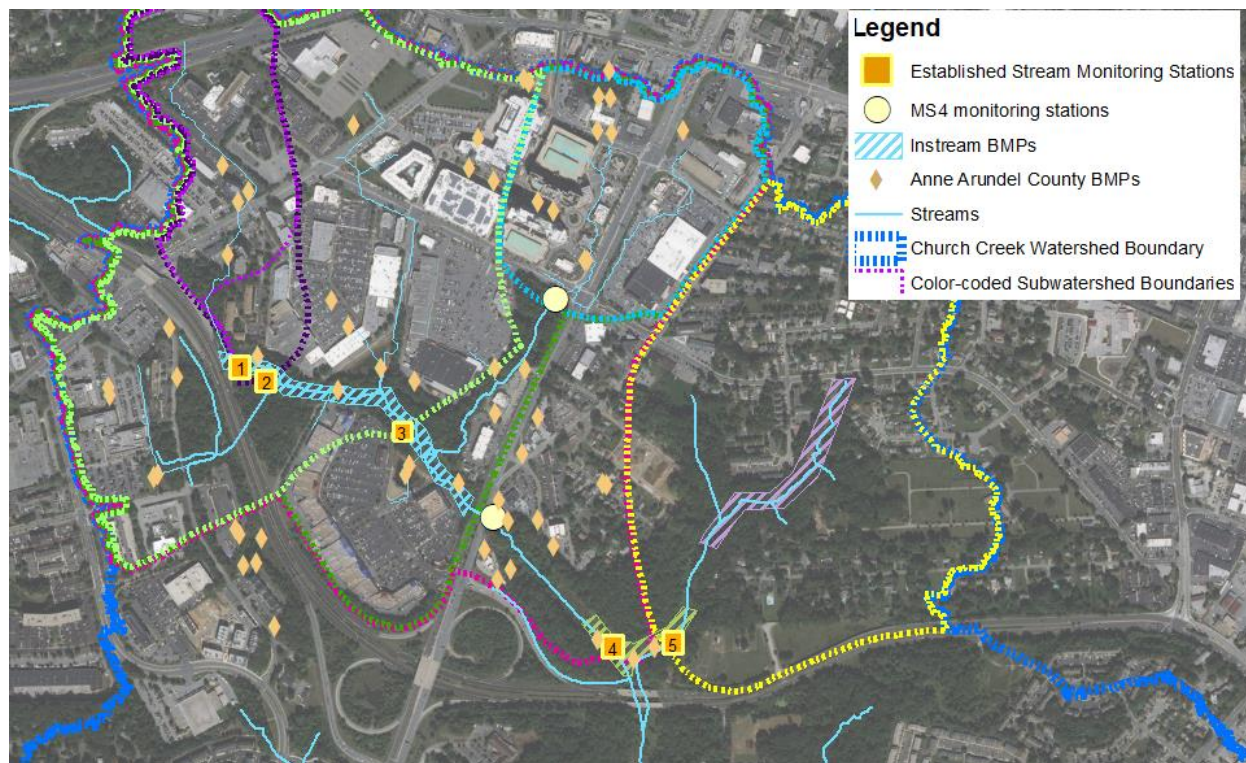
Methods:

Our approach was to compare nutrient and suspended solids loads from five sub-watersheds within the urbanized Church Creek watershed. By partitioning the Church Creek watershed into sub-watersheds, we intended to identify variations in pollution outflow associated with sub-watershed conditions, including the percentage of impervious surface, land usage, as well as type, size, and number of Best Management Practices (BMPs) installed within each sub-watershed. We expected that BMP's would be effective in removing nutrients and suspended solids from surface water. Our goal was to quantify BMP effectiveness and investigate the impact of BMP interactions. Does layering BMPs within the same watershed, as is done in Church Creek's watershed, have additive, synergistic, or antagonistic effects on the BMPs cumulative capacity to reduce nutrient and sediment loading? Knowing this would help us determine the amount of effort and investment needed to restore a watershed.

The five study watersheds have differing management practices in place that were expected to retain differing amounts of TN, TP, and TSS. Two of the watersheds had been monitored in a previous CBT-funded study that examined the changes in concentrations and loads at the inlet and outlet of a stream reach restored with a regenerative stormwater conveyance (RSC). The inlet of the RSC is the discharge point of one study watershed and the outlet is the discharge point of another study watershed (watersheds 1 and 2, respectively, Figure 1). During the subsequent study we continued monitoring watersheds 1 and 2 and we added three new monitoring stations (3, 4, and 5, Figure 1). Watersheds 1 through 4 overlap and have 64% to 66% impervious surface. Watershed 5 is separate from the others and has 31% impervious surface. All the study watersheds include extensive RSCs and a variety of other stormwater management practices designed to enhance retention of nutrients and TSS.

Measuring loads required measuring concentrations of TN, TP, and TSS as well as the water discharges. We also measured concentrations and loads of phosphate P (PO₄), ammonium N (NH₄), and the sum of nitrate and nitrite N (NO₃). From these measurements we calculated concentrations of organic P (by subtracting PO₄ from TP) and organic N (by subtracting NH₄ and NO₃ from TN). Methods for analyzing forms of N, P, and TSS were the same as used by Thompson et al. (2018).

Figure 1. The boundaries and stream sampling locations (numbered 1-5) of the study watersheds. Watersheds 1-4 overlap and are located within Anne Arundel County, Maryland. Watershed 5 is separate and located mainly in the City of Annapolis, Maryland. Blue, yellow, or purple hash marks across the streams indicate reaches that were restored by adding rock weirs across the channels to reduce erosion and create pools and riffles yellow diamonds indicate various other BMPs. Yellow discs mark sites monitored by the Anne Arundel County MS4 program.



Although the EPA has set Total Maximum Daily Loads (TMDLs) for TN, TP, and TSS for Chesapeake Bay, it is important to know the loads of different forms of N and P because they have different potentials for promoting algal blooms and other negative consequences of eutrophication. For example, dissolved inorganic forms are generally more bio-available than particulate forms and organic forms (Glibert et al. 2016). Also, the mechanisms of N and P removal from watershed discharges differ for different forms of N and P. For example, NO₃ may be removed from stream water by denitrification in hypoxic environments. The Chesapeake Bay Program has recently concluded that consideration of the different impacts of different forms of N and P will be important for managing eutrophication (Shenk et al. 2020).

We were not able to distinguish particulate from dissolved forms of N and P because our automated sampling requires that acid be added as a preservative to the sample that accumulates during the weekly sampling periods. The acid releases particulate forms PO₄ and NH₄ into solution. However, it is known that PO₄ and TP in streams are mostly particulate (Hartzell et al. 2017), as can be seen from correlations of with TSS (see results below). In contrast TN and NH₄ are mostly dissolved, and NO₃ is entirely dissolved (Jordan et al. 1997). Also, the predominance of particulate forms of N and P increase as the TSS concentration increases.

We used automated flow-paced sampling to monitor concentrations and loads of TN, NH₄, NO₃, organic N, TP, PO₄, organic P, and TSS. Each sampler included a Campbell Scientific data logger that recorded stream water depths every five minutes as measured by a pressure transducer in the stream. The logger calculated the stream flow based on an equation relating flow to depth. At each of our five sampling locations, we determined the relationship between stream depth (stage) and water flow rate by using a portable velocity meter to measure flow rates at points across the stream channel during different stages of stream flow. After a set amount of flow had occurred, the logger triggered an ISCO sampler to pump a water sample of a set volume. The sample was split into two separate jugs, one of which included an acid preservative for the sample to be analyzed for forms of N and P. Because the frequency of pumping depended on stream flow, the volume of sample collected during the sampling period was proportional to the stream flow and the concentrations of substances in the accumulated samples represented the flow-weighted average concentrations in the stream water passing the sampler during the usually week-long sampling period. Samples were collected at each station most weeks for two years. This method of flow-paced sampling has been demonstrated to be the most accurate and cost-efficient among several commonly used methods for measuring discharges of substances in stream water (Thompson et al. 2020).

We calculated weekly loads by multiplying the weekly flow-weighted concentrations by the total stream flow each week. We assumed that water flow per area would be the same for all the nested watersheds 1-4 because the watersheds overlapped and had very similar percentages of impervious surface. We used measurements of water flow based on stage measurements recorded in a culvert at the MS4 monitoring site between the sampling stations for watersheds 3 and 4. The stage versus flow relationship was more stable in the culvert than in the open stream channels where low-flow stage varied with formation and breakdown of debris dams. Flow measurements at watershed 5 were based on stage measured at the sampling station there.

We added the weekly loads to calculate annual loads, estimating loads for the few weeks lacking samples by multiplying the flow-weighted mean concentrations from all weeks measured during the year by the volume of the water flow during the unmeasured weeks. When comparing different watersheds, we expressed loads per area of the watersheds to account for the obvious effect of area on load.

Near each of the automated samplers at watersheds 3-5, we installed a water quality sonde that recorded in-situ measurements of dissolved oxygen, turbidity, temperature, and specific conductivity every 15 minutes. We were particularly interested in testing the use of turbidity measurements for calculating the discharge of TSS. This method could improve accounting for rapid variability of TSS concentration during high flow events while also saving costs of direct TSS analysis (Thompson et al. 2020).

Results:

Comparing Water discharges: The loads of TSS and forms of N and P depended on their concentrations and the amounts of water discharged. All the study watersheds showed the highly variable water discharge that is characteristic of watersheds with impervious urban surfaces (mainly roofs and pavement). Much of the discharge took place during and immediately after rain events. Some steady low-flow discharge continued between rains presumably due to gradual releases from pervious areas or

from flow-control structures such as retention ponds and step pools, as well as groundwater discharge into the streams.

The annual percentage of precipitation discharged from the watersheds depended on the percentage of impervious surface. The watersheds with 64% to 66% impervious surface discharged the equivalent of 77 inches of rain or 72% of the rainfall during the 2-year study period. The watershed with 31% impervious surface discharged the equivalent of 46 inches of rain or 43% of the rainfall during the study. The lower percentage of discharge in the watershed with the lower percentage of impervious surface probably reflects the greater amount of infiltration through more extensive pervious surfaces and the greater amount of evapotranspiration associated with more extensive vegetation. Lower water discharge contributed to lower the loads of nutrients and TSS from that watershed, as described later.

Comparing Concentrations and Loads:

The concentrations of TSS and forms of N and P differed among the watersheds (Tables 1 and 2) and forms of N showed seasonal variations. For most of the year nitrate and ammonium concentrations in stream water were similar to those in rain (Jordan et al. 1995 and recent unpublished data from SERC), but in summer stream-water nitrate concentrations were often below those in rain (Figure 2). NH₄ concentrations also decreased in summer at one site (Figure 2). This suggests that biological activity in summer may have removed NO₃ and NH₄ from stream water or converted them to organic N carried in the stream. Organic N concentrations were generally higher in summer, but the organic N peaks did not exactly mirror the NO₃ and NH₄ dips (Figure 2). The peaks in organic N in rain could reflect periods of pollen deposition as noted in a previous study of atmospheric N deposition (Jordan et al. 1995).

Table 1. Mean, standard error (SE) and number of measurements (N) of concentrations of forms of P and TSS measured in flow-paced samples collected at the 5 watershed sampling stations (Sta) over the 2-year study. Samples were usually composited over 1-week periods but were sometimes composited over 2-weeks during low flow. Watersheds 1 and 2 were sampled less extensively than watersheds 3-5.

Sta	PO ₄ (µg P/L)			ORGP (µg P/L)			TP (µg P/L)			TSS (mg/L)		
	Mean	SE	N	Mean	SE	N	MEAN	SE	N	Mean	SE	N
1	112.20	13.72	38	124.27	30.55	17	212.29	34.39	18	82.78	50.69	35
2	115.55	10.85	66	66.98	9.56	47	166.77	12.73	47	46.23	8.85	59
3	47.84	2.27	81	46.93	3.50	80	94.69	4.27	80	13.37	1.18	79
4	57.68	3.40	80	57.65	5.03	80	115.33	5.87	80	11.23	0.86	79
5	79.61	5.24	72	50.60	5.50	72	130.21	8.09	72	16.17	2.73	71

Table 2. Mean, standard error (SE) and number of measurements (N) of concentrations of forms of N measured in flow-paced samples collected at the 5 watershed sampling stations (Sta) over the 2-year study. See Table 1 caption for details about sampling.

	NH4 ($\mu\text{g N/L}$)			NO32 ($\mu\text{g N/L}$)			ORGN ($\mu\text{g N/L}$)			TN ($\mu\text{g N/L}$)		
Sta	Mean	SE	N	Mean	SE	N	Mean	SE	N	Mean	SE	N
1	390.5	42.1	38	260.3	32.4	37	1062.4	173.4	25	1726.6	225.4	25
2	442.2	32.3	67	38.3	7.2	67	970.0	180.0	35	1441.3	182.3	35
3	182.9	10.3	81	411.2	16.6	81	563.5	24.5	81	1157.5	26.0	81
4	220.6	20.0	80	208.9	17.0	80	625.5	51.8	79	1057.6	54.9	79
5	202.0	14.7	72	229.2	30.5	72	620.9	39.1	70	1056.4	51.1	70

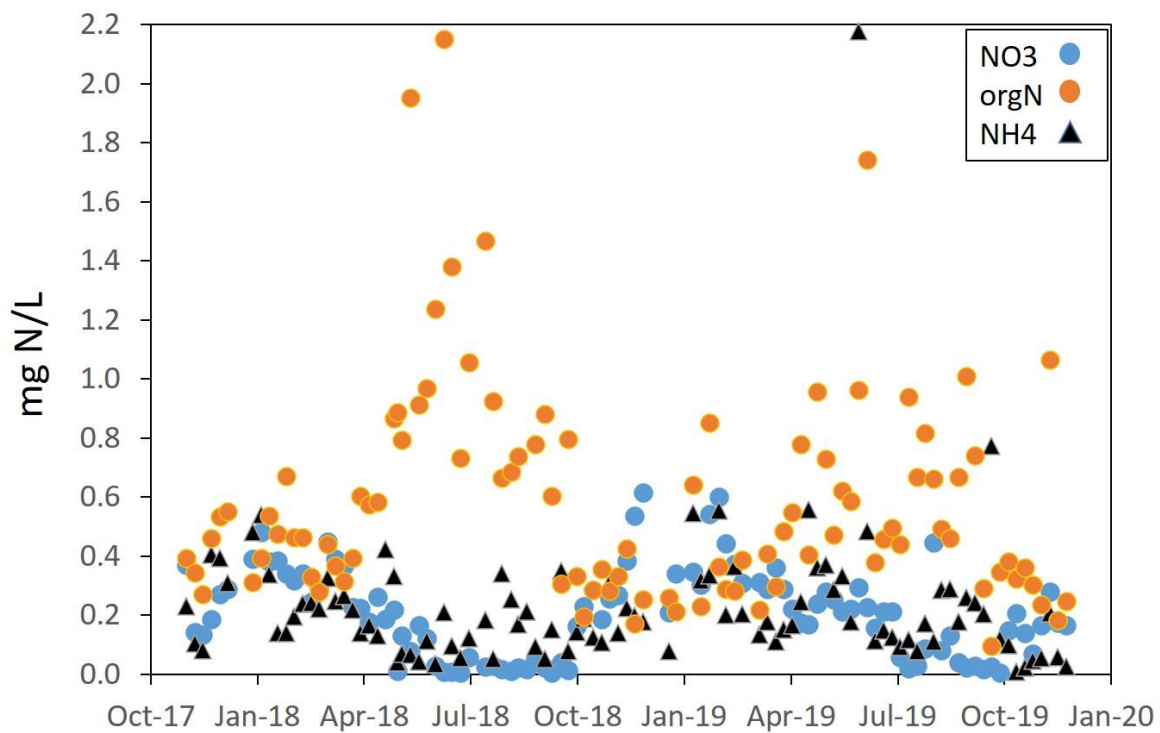


Figure 2. Concentrations of NO3, NH4, and organic N in weekly flow-paced samples collected from watershed 4 during our 2-year study.

NO3 concentrations covaried among the watersheds due to the seasonal variations of NO3 concentration (i.e. NO3 concentrations rise and fall synchronously in all the study watersheds). Covariance of other concentrations would also be expected among the overlapping study watersheds (1, 2, 3, and 4, Figure 1) where water discharged from one watershed enters the next watershed

downstream. Therefore, it is not surprising that the temporal changes in concentrations of all the substances we measured almost always significantly covaried ($p < 0.05$) in overlapping watersheds (one exception was a lack of correlation in PO₄ concentrations between watersheds 3 and 4). Even for the non-overlapping watersheds 4 and 5, all the weekly concentrations except for TSS covaried.

The concentrations of some substances we measured correlated with the amount of water flow during the week-long sampling period. Positive correlations with flow ($p < 0.01$) at watershed 3 for ORGN, PO₄, ORGP, TP, and TSS were the strongest correlations with flow observed. Positive correlations would be expected for TSS and forms of P because high flows tend to suspend solids in stream water and forms of P tend to be associated with particulate matter. This reflects the suspension of TSS by turbulence associated with high water velocity, and the association of TP and phosphate with particulate matter (Correll et al. 1999a). In contrast, NO₃, which is essentially all dissolved, was negatively correlated with weekly flow, as found in other studies, suggesting a tendency for nitrate in streams to be diluted by rain and runoff (Correll et al. 1999b, c). Even NO₃ in rain tends to be diluted during high volume rain events (Jordan et al. 1995) and, as mentioned, NO₃ concentrations in the watershed discharges were like those in rain. Loads of nutrients and TSS covaried among the watersheds even more strongly than concentrations because the water discharge rates of all the watersheds covary with rainfall.

Given the covariation of concentrations and water flow among the watersheds we used paired T-tests to test for statistical significance of differences of concentrations or loads between watersheds. This involved calculating the differences of the weekly concentrations or loads between the watersheds compared and using Student's *t*-test to determine whether the mean of the differences for all the weeks was statistically different from zero. This approach greatly increases the sensitivity for detecting differences between watersheds by controlling for the weekly changes in concentration or water flow that were common between the watersheds.

For the four watersheds that overlapped (1-4, Figure 1), we tested differences in concentrations and loads between upstream and downstream pairs of sampling stations. These comparisons highlighted the effect of the non-overlapping component of the downstream watershed of each pair. The differences in loads per area for those watersheds (see below) were driven by the differences in concentration because the water discharges per area were assumed to be equal for those watersheds, which had very similar percentages of impervious surfaces. We also tested the differences between the non-overlapping watersheds (4 and 5, Figure 1), which discharge near the bottom of the Church Creek watershed. Watersheds 4 and 5 have very different percentages of impervious surface and therefore very different water discharges per area. Thus, the differences in their loads reflect the combined effects of differences in concentration and differences in water discharge.

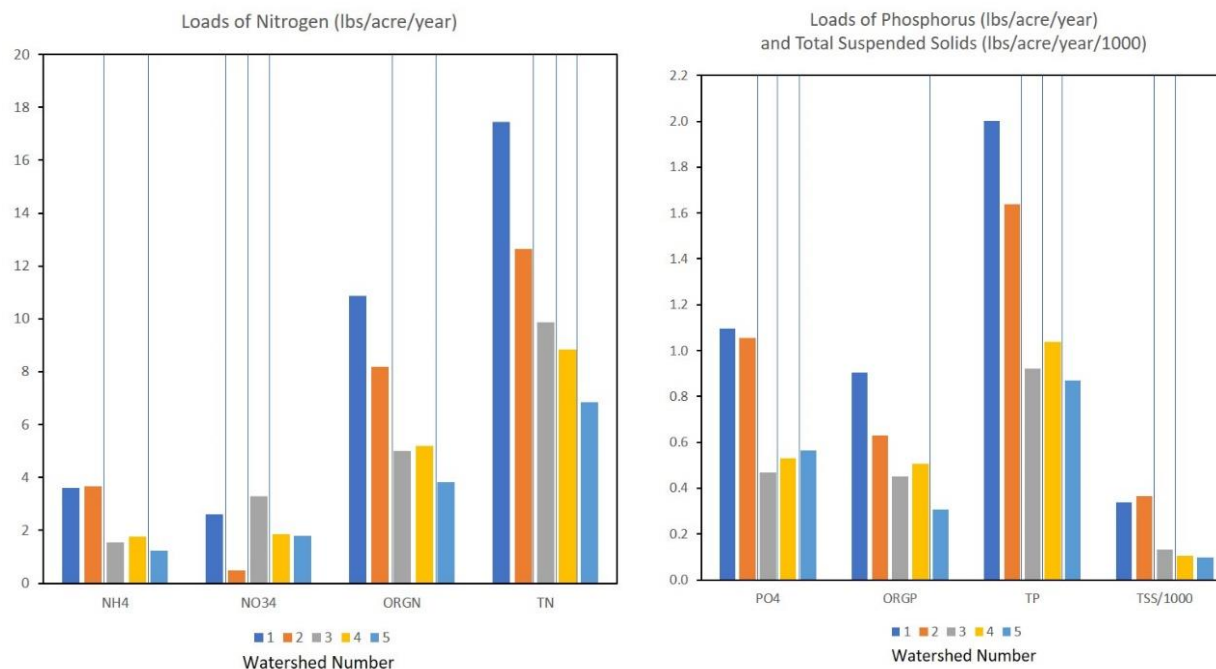
The annual average loads of organic N per acre generally accounted for about half the TN loads in the watersheds, with the sum of the inorganic forms (NH₄ and NO₃) making up the remainder (Figure 3). There were statistically significant differences of loads between many pairs of watersheds (Figure 3). For example, NO₃ loads were significantly lower from watershed 2 than from watershed 1, apparently due to uptake of NO₃ in the pools of the RSC installed between the sampling points for those watersheds (Figure 3). In contrast, NO₃ loads from watershed 3 were markedly higher than from watershed 2, indicating that the part of watershed 3 that did not overlap with watershed 2 has a strong source of NO₃ release. In contrast, difference in loads of NH₄ from watersheds 2 and 3 is opposite of the difference of their NO₃ loads. As with the N loads, the loads of inorganic and organic P forms were about equal. The

load of PO₄ from watershed 3 was significantly lower than from watershed 2 (Figure 3). The significantly lower organic P load from watershed 5 compared to watershed 4, largely reflects the difference in water flow rather than the difference in concentration. In general, the loads of TP, TN, and TSS declined from the upstream to the downstream nested watersheds (Figure 3).

Besides differing among watersheds, loads differed between the two years of the study, with higher water discharges in 2018 than in 2019, leading to higher loads in 2018 than in 2019. However, the difference in loads for NO₃ and NH₄ was less than the difference in water discharge because NO₃ and NH₄ concentrations were less in the wetter year. Similarly, NO₃ and NH₄ concentrations rain have been observed to decline with increases in rain volume (Jordan et al. 1995). This suggests that atmospheric deposition of NO₃ and NH₄ onto our watersheds may be a major source of the NO₃ and NH₄ loads. In contrast, the TSS loads were disproportionately higher in 2018 than discharges of N and P forms. This may be due to increased stream bank erosion and sediment resuspension that accompanies increased stream flow velocities and due to the lower N and P content of the coarser particles transported during higher flows.

An important finding of our load comparisons is that there were statistically significant differences among the study watersheds in loads per area of N and P forms as well as TSS. This suggests that differences of loads might reflect differences among the watersheds in the cumulative effectiveness of management practices designed to reduce loads of N, P, and TSS.

Figure 3. Loads of NH₄, NO₃, organic N, total N, PO₄, organic P, TP, and TSS measured at the 5 watershed discharge points over the 2-year study. Loads are shown as annual rates in pounds of N or P per acre, and lbs./1000 of TSS. The colors of the bars indicate the watershed. Lines between bars indicate that the significant differences ($p < 0.01$) between loads of the stations plotted as adjacent bars, based on paired comparison of weekly loads.



Comparing Management Practices

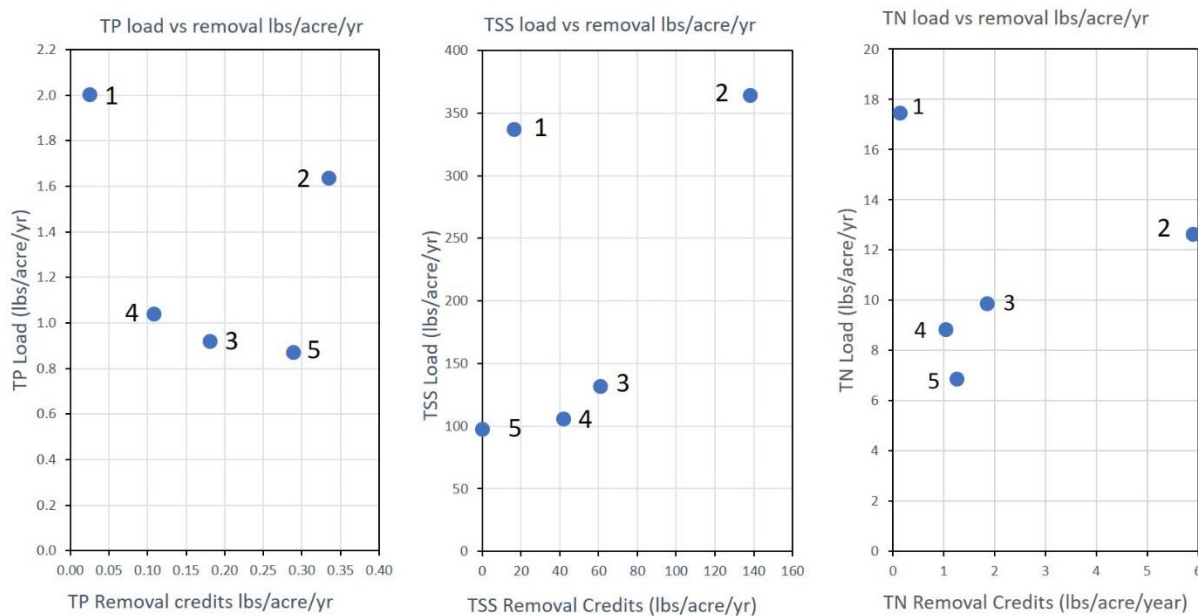
Our study watersheds each have a variety of BMPs that were in place at the start of the study (e.g. retention ponds, rock weirs in streams, infiltration swales, etc.). We could not study each practice separately, instead we measured loads of nutrients and TSS discharged from the whole watersheds with the goal of assessing the combined effects of the practices. We obtained information about the practices in the watersheds such as the type of BMP, its size, the area of drainage to the practice, and the date of construction. Anne Arundel County Office of Watershed Protection and Restoration was very helpful in supplying this information for most of our study watersheds. Less information was obtained from the City of Annapolis, which had jurisdiction over the land in watershed 5, which had a lower percentage of impervious surfaces than the others.

The newer BMPs are credited with expected removals of N, P, and TSS toward achieving goals of reducing their loads to Chesapeake Bay. The credits are assigned according to guidelines established by panels of experts that rate features and design criteria of various practices that are thought to reduce loads of N, P, and TSS. The decision to fund the BMPs is strongly influenced by these credits. Thus, it is important that the credits provide a reliable predictor of performance.

After reviewing the information on the BMPs, we decided to use the credits assigned in accordance with the expert panel guidelines rather than develop our own system of calculating expected performance. Here we are considering the credits to be hypothetical estimates of the performances of the practices. Comparing watersheds with differing combined credits from a variety of BMPs would provide a test of the expected performance of the BMPs. Also, testing the predictive value of the credits would be of great value to the decision makers who rely on them for selecting BMPs.

We compared the sums of those credits for the BMPs in each watershed with the corresponding loads from each watershed. The range of removal credits is wide compared to the range of the loads among our watersheds (Figure 4). For example, the maximum and minimum TN loads of the watersheds spanned a range of 10.6 lbs./acre/year while the maximum and minimum TN removal credits of the watersheds spanned a range of 5.8 lbs./acre/year, which is more than half of the span of the loads. Similarly, the TSS removal credits spanned a range of about half the difference between the highest and lowest TSS loads. Given the relatively wide range of credits compared to the range of loads, it seems plausible that there would be some relationship between the loads and the removal credits, but we found no correlation between credits and loads for either TN or TSS among the five watersheds we compared (Figure 4). Similarly, there was no correlation between removal credits and loads for TP (Figure 4). However, the range of the TP credits was only about one quarter of the range of the TP loads.

Figure 4. Annual loads of TP, TSS, and TN (lbs./acre/yr) over the 2-year study plotted against the total removal of the same analytes. The numbers adjacent to points on the graphs indicate the watershed.



We tried a similar analysis subtracting the loads and credits of the nested watersheds from the loads and credits of the watersheds in which they are nested to yield an estimate of the loads and credits for the part of the watershed excluding the nested watershed. This led to estimates of negative loads for watershed 2 in which the nested part occupied most of the watershed. Negative loads might be plausible for that watershed because it includes a regenerative stormwater conveyance (RSC) designed to remove N, P, and TSS from the stream channel that receives water from the nested watershed. Even when excluding the nested portions of the watersheds, there was no clear relationship between the credits and the loads for TN, TP, and TSS.

These findings do not support our hypothesis that remediation efforts are effective in removing TSS, TN, and TP from surface water, but do not conclusively disprove our hypothesis because of the possibility that there are differences in the sources of TSS, TN and TP to surface water in the watersheds. However, the similarity of percent impervious surface in four of our watersheds suggests that they might be well matched for the sources of nutrient and TSS discharge. Watershed 5 with about half the extent of impervious surface than the others had lower loads, mainly due to lower water discharge rather than lower concentrations. That difference did not account for the lack of correlation between loads and removal credits. Lack of correlations could be due to the small number of watersheds compared and the variability of measuring nutrient loads and estimating removal credits. Also, many of the stormwater management practices in the watersheds had no associated estimates of removal credits between the credits and the loads for TN, TP, and TSS.

Improving the Efficiency of Monitoring Methods

We tested correlations among several of the parameters we measured to see if we could streamline our monitoring by eliminating redundant measurements, by reducing the numbers of locations monitored, or by reducing the frequency of measurements. We compared the concentrations of different substances in weekly flow-paced samples from the same location and we compared concentrations of the same substance sampled at different locations. We also compared in-situ measurements of different water quality parameters recorded at the same location and in-situ measurements of the same parameters recorded at different locations. Finally, we tried a method for predicting weekly TSS concentrations from in-situ measurements of turbidity.

We compared the measurements of TSS and total P concentrations in the weekly flow-paced samples because typically total P in stream water is mostly in particulate form (e.g. Hartzell et al. 2017). We found that total P and TSS concentrations showed similar patterns of fluctuation (Figure 5) with peaks usually occurring during weeks with high water flow. However, the correlation was not tight enough to use for predicting one concentration from the other. In contrast, we found that weekly measurements of TP were not at all correlated with weekly measurements of NH₄ at the same location. Comparing the weekly measurements of TP at the outlets of the nested watersheds 3 and 4, we found some correlation (Figure 6), as would be expected considering that watershed 3 makes up half of watershed 4. However, the relationship was not tight enough to predict TP concentration effectively at one watershed from TP concentration at another.

Comparing in-situ water quality measurements of the same parameters at different locations, we found that conductivity and temperature were closely correlated among different sites but the correlations could change over time, for example, with differing shifts in conductivity happening at different sites (Figure 7) or with temperature gradually becoming higher at one site than another (Figure 8). Conductivity and temperature measurements at one location could be used to predict measurements at another for short time periods, but it would require long time series of measurements to construct a model for predictions over long time periods. In contrast, turbidity measurements at one site could not be used to predict turbidity and another site. Turbidity at all sites showed upward spikes during high flow events but the spikes were of very short duration and variable heights (Figure 9).

Figure 5. An example time series of weekly concentrations of TP ($\mu\text{g P/L}$ and TSS mg/L). Collected by the automated flow-paced sampler at the discharge point of watershed 3.

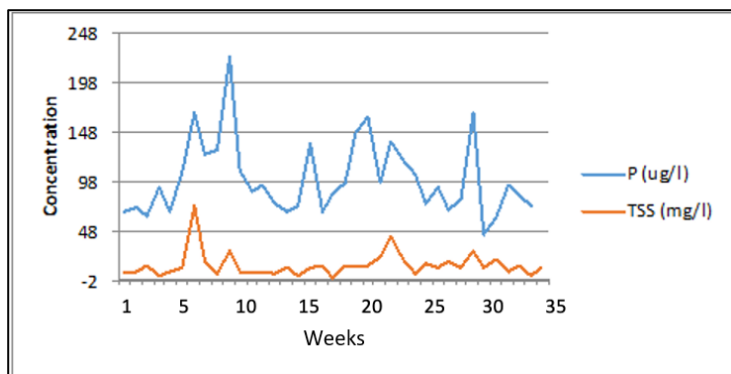


Figure 6. TP concentration ($\mu\text{g P/L}$) at watershed 4 vs. at watershed 3. The line of equal concentration (1:1) is plotted for comparison.

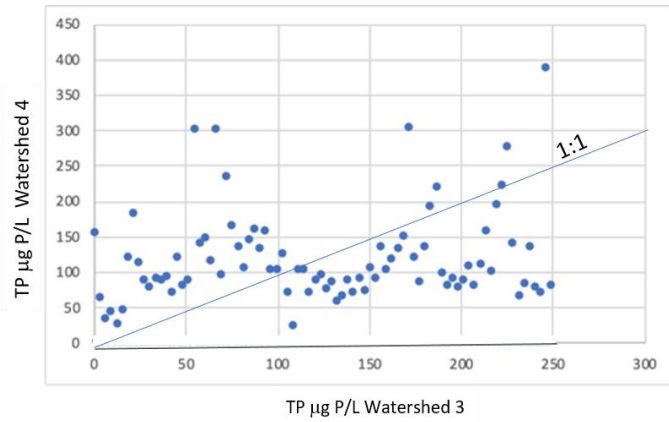


Figure 7. Conductivity (expressed as salinity equivalent ppt) measured every 15 minutes at watersheds 3 and 4 versus time.

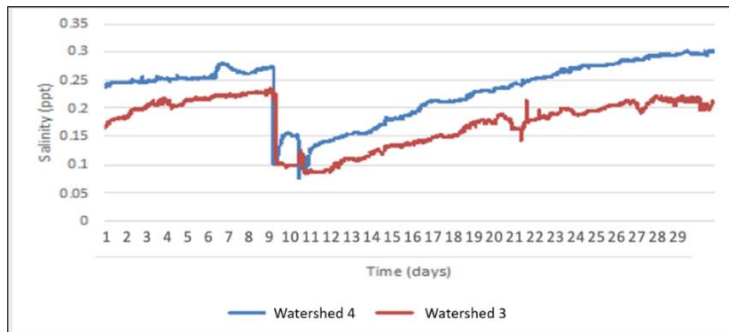


Figure 8. Temperature ($^{\circ}\text{C}$) measured every 15 minutes at watersheds 3 and 4 versus time.

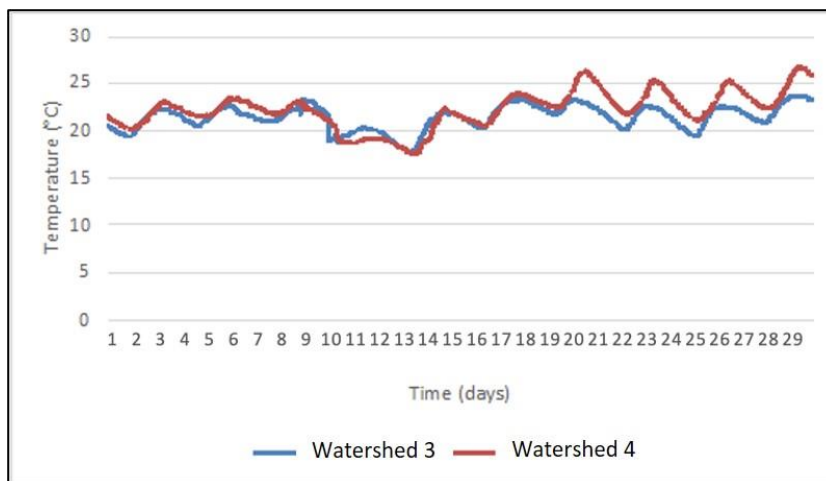
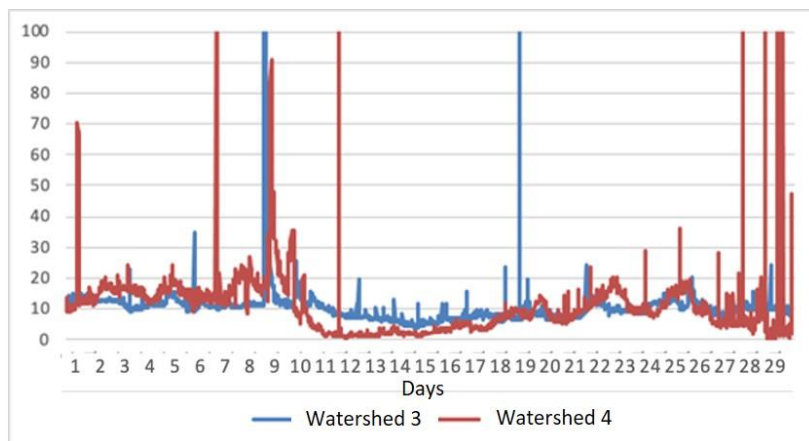


Figure 9. Turbidity measured every 15 minutes at watersheds 3 and 4 versus time.



We were especially interested in developing a way to predict weekly flow-weighted TSS concentrations for continuous in-situ turbidity measurements as others have done (e.g. Thompson et al. 2020). However, we were unable to establish a consistent correlation between our measurements of turbidity and our measurements of TSS. This may have been due to a lack of synchrony between our in-situ measurements of turbidity and the automated acquisition of water samples to analyze for concurrent TSS concentration. Our attempts to modify the sample acquisition to improve the synchrony did not result in consistently useable correlations. Acquiring concurrent measurements is especially difficult in urban watersheds where flow and turbidity can vary greatly over time scales of minutes. It is also possible that suspended particles with differing optical properties pass by the sampling station at different times in a flow event. For example, we have observed brief periods at the beginning of flow events when black particles flow by the sampling station followed by lighter colored particles later in the event. In such cases, the relationship between turbidity and TSS concentration may differ at different times during the flow event.

Although the in-situ sonde data did not provide proxies for direct measurement of nutrient and suspended solids, the data are potentially valuable in other ways. For example, conductivity was very successful in recording release of road deicing salts. In addition, monitoring dissolved oxygen documented episodes of oxygen depletion during warm low-flow conditions. Oxygen depletion can enhance removal of nitrate by denitrification and may be related to the decreases in nitrate concentration observed during warm weather.

Reviewing Research Questions and Hypotheses:

A1: What are the cumulative effects of watershed restoration activities within a watershed?

We did not find clear evidence of BMP effects. The nutrient and TSS loads from our five study watersheds were not correlated with the total credited rates of nutrient and TSS removals by the BMPs. This does not prove that there were no effects but does not provide evidence for effects either.

A.3: What are the most efficient methods for measuring effects of stormwater management methods at watershed scales?

We did not find any new ways to increase the efficiencies of measuring the effects of stormwater management at watershed scales. However, other recent research has shown that our method of automated flow-paced sampling is more efficient and accurate than other methods for measuring watershed discharges (Thompson et al. 2020).

Hypotheses to be tested:

Hypothesis 1. Remediation efforts are effective in removing Total Suspended Solids (TSS), Nitrogen (N), and Phosphorous (P) from surface water.

It is unclear whether the remediation efforts are effective. The lack of correlation between the total expected removal of TP, TN, and TSS and the loads of these substances from the 5 watersheds suggests a lack of effectiveness, but it is possible that the different watersheds differ in the amount of N and P released into the watershed from various sources and also differ in the factors affecting mobilization of suspended solids into the stream. Moreover, the data on the expected effects of the remediation efforts are incomplete, relying only on the credits assigned to newer BMPs.

Although the test of the hypothesis was inconclusive, the following two findings suggest possible effects of the BMPs:

- 1) Reduction of the concentration of nitrate flowing through the RSC between the sampling stations of watersheds 1 and 2 could be a direct effect of the RSC installed between the sampling stations. The reduction is too high to be attributed to dilution by water discharged by watershed 2.
- 2) The seasonal reductions in nitrate concentration below that in rainwater in 4 of the 5 study watersheds suggests that there was biotic uptake of nitrate in summer, which could reflect the expected effects of constructed pools in the stream channels present in all the streams we sampled.

Apparent removal of nitrate by BMPs in our watersheds is consistent with the observation of a 40% decline in TN load from the entire Church Creek watershed (including the areas of our study watersheds) after extensive stream restorations (Filoso 2020). Comparing nine other streams, the same study found that declines in TN loads occurred in all but one after restorations, while responses of TP and TSS loads were less consistent. In Church Creek there was no clear change in TP load and TSS seemed to increase after the restorations (Filoso 2020).

In general, the nutrient concentrations in the streams we studied were relatively low. The nitrate and ammonium concentrations were comparable to those in rainfall, except in summer when nitrate concentrations in the streams was lower than typical for rainfall. The low nutrient concentrations might reflect low nutrient inputs to the watersheds and/or effective removal of nutrients by management practices already deployed in the watersheds. If we had data on nutrient and TSS loads from before as well as after the BMPs were constructed, we would have a better basis for judging their effects.

Hypothesis 2. BMP type and impervious surface percentage in a sub-watershed will affect reductions in the pollution outfall.

We could not compare the effects of different BMP types because a wide variety already were present in all the watersheds. Only the sampling at watersheds 1 and 2 focused on a single type of BMP, an RSC with a series of pools and rock weirs.

Our comparison of the effects of percentage of impervious surface were also limited because 4 of the 5 watersheds we studied had very similar percentages (64-66%). However, we found that the watershed with only 31% impervious surface discharged about 60% less water during the 2-year study than did the other study watersheds and that the lower water discharge resulted in lower loads of N, P, and TSS.

Hypothesis 3. Correlations between measurements of turbidity and TSS can support the use of turbidity as a proxy measurement for some for TSS and particulate forms of N and P.

We did not identify any clear proxy measurements. We had hoped to correlate turbidity with TSS concentration to enable calculation of TSS load based on low-cost monitoring of turbidity. We and others have had success with that approach (Thompson et al. 2020). However, in the present study we were unable to obtain consistent correlations between turbidity measurements in the stream and TSS concentrations in samples pumped concurrently from the stream. We tried to measure correlations during several flow events. For some events or parts of events there were correlations but the equations for those differed. We suspect that rapidly changing flow rates and TSS concentrations in urban streams may make this method difficult to calibrate.

How can the findings be used for the regulatory community, for practitioners, for researchers, and others?

Regulators and practitioners:

Some urban lands might not be major non-point sources of nutrient discharge. The nitrogen discharge rates from the watersheds we studied are much less than commonly found in agricultural lands (e.g. Jordan et al. 2003). Our study watersheds had NO₃ and NH₄ concentrations and loads similar to those in rainfall. This might be expected because urban lands are designed to rapidly shed rain. However, it does not suggest that urban lands *per se* are necessarily the source of those nutrients, as NO₃ enters the atmosphere from fossil fuel combustion while NH₄ enters the atmosphere from agricultural sources, mainly livestock waste (Li et al. 2016). It is possible that our study watersheds already have enough BMPs to lower the nutrient and TSS discharges. However, before investing in BMPs designed to reduce nutrient discharges from urban watersheds, we advise measuring nutrient concentrations in their discharges to evaluate the need for nutrient removal.

The credits for nutrient and TSS removal assigned to BMPs need to be verified. The best way to check the credits of the BMPs is to measure discharges before and after installation. To verify the performance of BMPs information about the BMPs should be publicly available. Anne Arundel County provided detailed information about BMPs in four of our study watersheds in their jurisdiction, but we were not able to obtain complete information about BMPs in the watershed within the City of

Annapolis. This made it difficult to compare the expected performance of the BMPs across all our watersheds.

Researchers:

It may be problematical to use turbidity measurements as predictors of TSS in urban streams. While it is widely known that TSS and Turbidity are positively correlated, we were unable to find a consistent enough relationship to predict TSS from turbidity, possibly due to the variability of water flow, TSS concentrations, and optical properties of TSS in the urban streams.

We also found it difficult to measure water flow based on stream stage in restored urban streams. Water discharges from watersheds with extensive impervious surfaces change rapidly and it is difficult to obtain flow measurements at a range of stages when stage is changing rapidly. Also, the stream restorations tend to broaden the channels, which makes a small increase in depth represent a larger increase in water flow than would be the case in a narrow channel. The broad channels are also likely to accumulate ephemeral debris dams that cause the baseflow stage to vary over time. This led us to using stage measurements in a culvert to calculate the flow per area of our nested watersheds upstream of the culvert.

Future Research

More research is needed about the sources of N and P in urban watersheds. These nutrients might not increase directly with the proportion of impervious surface. A suburban watershed with septic systems and fertilized turf might release more nutrients than a more highly impervious watershed lacking turf and septic systems. In some cases, nutrient releases from urban watersheds may come from leakage of sewer pipes.

With knowledge of the sources of nutrients in a watershed, regulators may decide to address the sources directly rather than constructing BMPs to remove the nutrients after they are released into the streams. If necessary, improving sewage and septic systems could be more effective at reducing nutrient discharges than would restoring streams. In the watersheds we studied, it is possible that the dissolved inorganic N comes from atmospheric deposition. Recent reductions in releases of nitrogen oxides to the atmosphere have already reduced this source of NO₃ (Li et al. 2016). Most of the total N concentration in the streams we studied was organic N. Organic N is generally less available to phytoplankton than the dissolved inorganic forms. However, it would be useful to investigate the potential availability of organic N in urban runoff.

It is likely that most of the total P in our streams is particulate P. Our analyses could not distinguish between dissolved and particulate P because we used acid to preserve our water samples while they accumulated in the automated samplers. However, particulate P is usually the main form of P transported in streams (Föllmi, 1996). Dissolved inorganic P is much more available to phytoplankton than particulate P and it is possible for particulate P to be converted to dissolved P that is released into estuarine waters. However, there is evidence that most of the particulate P entering Chesapeake Bay is deposited and buried in the sediments of the low-salinity portions of tributaries (Hartzel et al. 2017). Future research should evaluate whether the particulate forms of N and P released from urban watersheds would be buried in the sediments of estuarine waters or be released to the water column and potentially stimulate phytoplankton blooms. Such knowledge could help indicate what regulatory actions are needed to prevent the negative consequences of eutrophication.

We also recommend further study of one type of BMP present in our study watersheds: retention ponds. Although retention ponds are widely used, their effectiveness has seldom been directly measured and the published studies show variable results. For example, Hancock et al. (2010) found that four of five ponds evaluated “fail to achieve regulatory goals” for water discharge and argued that more studies were needed to assess the performance of these types of BMPs. While some studies have produced results that demonstrate moderate ability to reduce nutrient and sediment pollution, others have found that some ponds have a statistically insignificant or even inverse pollution reducing capacity (Borne, 2014). The effects of retention ponds on deicing salt pollution should also be investigated because the ponds may enhance infiltration of salt-contaminated surface water into aquifers.

References:

- Borne, K. E. 2014. Floating treatment wetland influences on the fate and removal performance of phosphorus in stormwater retention ponds. *Ecological Engineering* 69: 76-82.
- Correll, D. L., T. E. Jordan, D. E. Weller. 1999a. Precipitation effects on sediment and associated nutrient discharges from Rhode River watersheds. *Journal of Environmental Quality* 28:1897-1907.
- Correll, D. L., T. E. Jordan, D. E. Weller. 1999b. Effects of precipitation and air temperature on nitrogen discharges from Rhode River watersheds. *Water, Air, and Soil Pollution* 115: 547-575.
- Correll, D. L., T. E. Jordan, D. E. Weller. 1999c. Transport of nitrogen and phosphorus from Rhode River watersheds during storm events. *Water Resources Research* 35:2513-2521.
- Filoso, S. 2020. Evaluating the Effectiveness and Sustainability of Novel Stream Restoration Designs for Coastal Plain Streams in Maryland: Integrating Existing and New Data from Restoration Monitoring, Final Report to the Chesapeake Bay Trust.
- Föllmi, K.B. 1996. The phosphorus cycle, phosphogenesis and marine phosphate-rich deposits. *Earth-Science Reviews* 40: 55–124.
- Glibert, P. M., Wilkerson, F. P., Dugdale, R. C., Raven, J. A., Dupont, C. L., Leavitt, P. R., Parker, A. E., Burkholder, J. M., and Kana, T. M. (2016), Pluses and minuses of ammonium and nitrate uptake and assimilation by phytoplankton and implications for productivity and community composition, with emphasis on nitrogen-enriched conditions. *Limnology and Oceanography*, 61: 165-197. Doi:10.1002/lno.10203.
- Hancock, G.S., J.W. Holley and R.M. Chambers. 2010. A field-based evaluation of wet retention ponds: how effective are ponds at water quantity control? *Journal of the American Water Resources Association* 46(6):1145-1158.
- Hartzell, J. L., T. E. Jordan, J. C. Cornwell. 2017. Phosphorus Sequestration in Sediments Along Salinity Gradients of Chesapeake Bay Subestuaries. *Estuaries and Coasts*. DOI 10.1007/s12237-017-0233-2.
- Jordan, T. E., D. E. Weller, and D. L. Correll. 2003. Sources of nutrient inputs to the Patuxent River estuary. *Estuaries* 26:226-243.
- Jordan, T. E., D. L. Correll, D. E. Weller, and N. M. Goff. 1995. Temporal variation in precipitation chemistry on the shore of the Chesapeake Bay. *Water, Air and Soil Pollution*. 83:263-284.

Li, Y., B. A. Schichtel, J. T. Walker, D. B. Schwede, X. Chen, C. M. B. Lehmann, M. A. Puchalski, D. A. Gay, J. L. Collett Jr. 2016. Increasing importance of deposition of reduced nitrogen in the United States. Proceedings of the National Academy of Sciences, 113: 5874-5879.

Shenk, G., Wainger, L., Wu, C., Capel, P., Friedrichs, M., Hubbart, J., Iho, A., Kleinman, P., Sellner, K., Stephenson, K. 2020. Assessing the environment in outcome units. STAC Publication Number 20-00x, Edgewater, MD. 34 pp.

Thompson J., C.E. Pelc, W.R. Brogan III, T.E. Jordan. 2018. The multiscale effects of stream restoration on water quality. Ecological Engineering 127:7-18.

Thompson, J., C. E. Pelc, T. E. Jordan. 2020. Water quality sampling methods may bias evaluations of watershed management practices. Science of the Total Environment.

Presentations:

December 7, 2018, Maryland Water Monitoring Council Conference (MWMC), Maritime institute in Linthicum, MD

Title: Stream Team: Citizen Science and Urban Best Management Practices.

Speaker: Sarah Giordano

Number of attendees: ~ 25 people

Abstract:

South River Federation's citizen scientists have become an invaluable resource in our evaluation of stormwater best management practices (BMPs). In 2017, South River Federation and Smithsonian Environmental Research Center, supported by the Chesapeake Bay Trust, joined forces to study Church Creek, an urban tributary previously known as one of the most polluted creeks in the South River Watershed. Since 2010, Church Creek has been the focus of restoration efforts and now has over 50 BMPs installed within its watershed. Continuous monitoring equipment was placed at key locations to assess the cumulative and individual effectiveness of the BMPs on reducing nutrient and sediment pollution. However, the majority of these devices provide a limited view of the system, collecting weekly flow-based averages that represent the whole of upstream sub-watersheds. In order to fine-tune our understanding of the nutrient pollution input to Church Creek, the Federation's volunteers collect synoptic data from all small tributaries within the research area during high flow events. Through the effort of our dedicated citizen scientists, we have been able to not only identify hot spots, but also to pinpoint areas of high reductive efficiency. Ultimately, these efforts will allow for more informed decision making when it concerns impaired stormwater systems, such as Church Creek.

November 2019, Chesapeake Watershed Forum, Shepherdstown, MD

Title: Advancing Stormwater Research with Citizen Science

Speakers: Sarah Giordano and Nancy Merrill

Number of attendees: ~ 35 people

Abstract:

Looking to track stormwater pollution? Not sure how to effectively involve your citizen scientists? In this session, we share techniques to engage volunteers in water pollution detection. As a case study, we discuss the invaluable role citizens have played in our research project, evaluating the cumulative effectiveness of stormwater Best Management Practices. With their help, we have been able to broaden our research scope and collect data that will ultimately inform the management of impaired stream systems. The volunteer training and coordination tools presented will provide avenues to reach the true capacities of citizen scientists and achieve your research goals.

December 2019, Maryland Water Monitoring Council Conference (MWMC), Maritime institute in Linthicum, MD

Title: Evidence of Nitrate Uptake in Step Pool Stormwater Conveyances in an Urban Watershed

Speaker: Tom Jordan

Number of attendees: ~50

Abstract:

We used automated flow-paced sampling to monitor concentrations and fluxes of nutrients and total suspended solids (TSS) at five points in a stream network draining an urban watershed. The stream network includes extensive step pool stormwater conveyances (SPSCs), which may enhance retention of nutrients and TSS. Comparing points entering and leaving an SPSC, we found that average nitrate concentration declined from 0.24 mg N/L to 0.07 mg N/L, indicating a retention of about 70% of the nitrate entering, while ammonium, phosphate, total phosphorus, and TSS were not retained. At a point downstream of that SPSC and downstream of a major tributary that lacked an SPSC, nitrate increased to 0.42 mg N/L, but, further downstream of additional SPSCs, nitrate dropped to 0.23 mg N/L. Nitrate concentrations also showed strong seasonal changes with the lowest concentrations in summer and the highest in winter. Ammonium showed a similar seasonal pattern at one location but not at others. Phosphate and total phosphorus concentrations did not show clear seasonal patterns. The spatial and temporal patterns of nitrate suggest that biological uptake in the SPSCs, possibly enhanced by summer hypoxia in the step pools, may play a role in reducing nitrate concentrations and fluxes.

February 18, 2020: 2020 Earth Optimism Lecture Series, Smithsonian Environmental Research Center

Title: City Stream, Country Stream: Getting a Clearer Picture of Stream Restorations

Speaker: Tom Jordan

Number of attendees: ~110

Recorded audio and PowerPoint posted at www.serc.si.edu

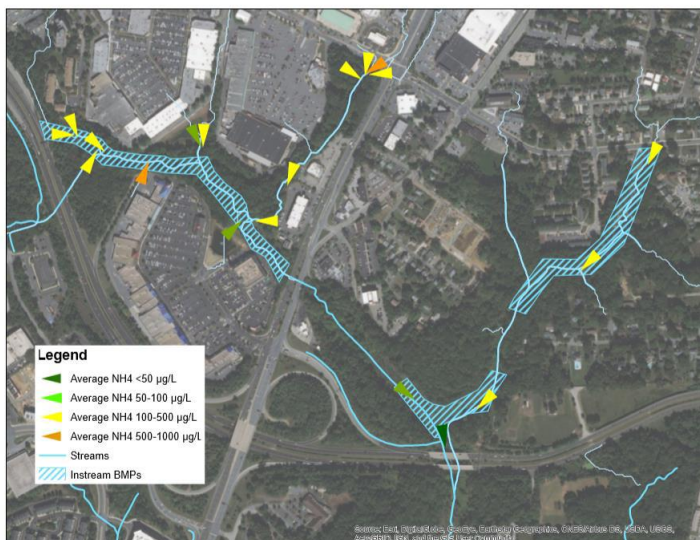
Lecture Announcement: Stream restorations, especially in urban watersheds, are a popular strategy for reducing nutrient loads to Chesapeake Bay. However, the latest research suggests their effectiveness can vary. In this talk, SERC nutrient ecologist Tom Jordan will compare two stream restorations his lab has been tracking: an urban on and a rural one. He'll reveal the different approaches each restoration took, and how each measured up in terms of improving water quality.

Appendix: Synoptic Sampling (Stream Team):

To look for possible source areas of TSS or dissolved NO_3 , NH_4 , or PO_4 , we supplemented our continuous monitoring of discharges from five sub-watersheds of Church Creek with synoptic sampling at 30 locations within the sub-watersheds. We selected sampling locations mainly at junctions in the stream network to highlight contrasts among the tributaries. Typically, tributaries and the larger streams they join were sampled just upstream of the junction. A team of volunteer citizen scientists collected the synoptic samples over one-hour periods during four different runoff events, thereby capturing the first flush of runoff when pollutant concentrations are liable to be elevated.

Comparing four synoptic sampling dates, we found that certain locations had consistently high or low concentrations. For example, Figures 10, 11, and 12 show points with consistently of high NH_4 , PO_4 , and NO_3 , respectively entering a tributary that connects with a stream flowing from the watershed 3 sampling station to the watershed 4 sampling station. Currently, we can only speculate about why high concentrations were found there, but these results suggest that synoptic sampling could point toward specific locations of nutrient sources. Locations with exceptionally high or low concentrations tended to be in the upper reaches of the watersheds, as would be expected due to homogenization as waters from various locations mix while flowing downstream through the drainage basin.

Figure 10. Averaged synoptic NH_4 concentrations at 19 stations throughout the Church Creek watershed research area (only those with datasets from all three storm events were included), show the elevated ammonium concentrations flowing from the upstream-most portions of the research area, which are surrounded by impervious surfaces.



[illegible][illegible]