

FINAL REPORT

University of Wisconsin – Green Bay *Sub-Award Report*

Grant Number: GL00E01451

Project Title: Targeting Outcome-Based Sediment Reduction in the Lower Fox Watershed

Main Award to: Fox-Wolf Watershed Alliance, Inc

Project Period: 2/23/2015 to 2/28/2021

Evaluation of the impact of a vegetated strip on soil and phosphorus export from a concentrated flow area using a paired edge of field study design

May 2021

University of Wisconsin – Green Bay

Disclaimer: Use of company or product names is for descriptive purposes only and does not imply endorsement by the University of Wisconsin – Green Bay.

BACKGROUND

The objective of this study was to evaluate the effectiveness of installing a vegetated strip in a shallow concentrated flow channel to reduce soil erosion and phosphorus export from an agricultural field catchment. We wished to answer the following question: How effective is a vegetated strip in reducing concentrated flow erosion from a relatively shallow sloped agricultural field?

This conservation practice is formerly referred to as *concentrated flow area treatment* (i.e., critical area planting of vegetation in a shallow concentrated flow channel). This practice would generally not be utilized in deeper flow channels where a *constructed grass waterway* is more suitable because filling and extensive reshaping are required due to the depth of the gully, and/or frequent nature of gully formation. The vegetated (i.e., grassed) strip treatment practice was chosen because it is likely to be more acceptable than an engineered constructed grass waterway which requires more resources to construct and does not permit planting a crop through it. In addition, areas where a grassed strip might suffice to reduce erosion are often more extensive than those that would require a constructed grass waterway. The lowest portion of a concentrated flow path might require a constructed grass waterway due to the depth of the ephemeral or permanent gully that has formed; whereas, a grassed strip would frequently suffice for the upper portion of the concentrated flow path because there is less water volume flowing through it. There is greater need for a grassed strip, grass waterway and eventually a protected ditch or stream as drainage area and water volume and peak flow increase. So, it follows that there are more areas where a grassed strip would suffice compared to where an engineered grass waterway would be required. Installing a grassed strip typically only requires some minor tillage and/or leveling within and along the shallow flow channel slopes with tillage equipment prior to planting the vegetation on the slopes and middle of the concentrated flow path, as was done in this study.

The null hypothesis is that there is no difference in the constituent event-mean concentrations or event yields between the concentrated flow channel with a grassed strip and the one without this targeted practice. If there is a statistically significant difference ($p < 0.05$), the null hypothesis will be rejected in favor of the alternative hypothesis that there is a decline in these constituents that is likely due the grassed strip practice.

The UWGB was solely responsible for conducting the concentrated flow area monitoring study, including the construction, operation and maintenance of all monitoring equipment.

Acknowledgements: The Outagamie County Land Conservation Department (LCD) oversaw the construction of the berms that directed runoff into the catchment outlets, and they assisted in the placement of the wingwalls. The Outagamie and Brown County LCD's were also responsible for overseeing the grassed strip installation, and maintenance. Their contribution to this study is greatly appreciated. The farm owner and the farm operator were very cooperative with maintaining consistent crop rotation and tillage management, and the latter assisted with ensuring that the grassed strip integrity was maintained once it became established. Without the assistance of these people and organizations, this study would not have been possible.

A brief overview of the monitoring methods and site description for this project are provided below, followed by the full methods section.

Overview: The study site is in an agricultural field located in the Plum Creek watershed, which is within the Lower Fox subbasin, in NE Wisconsin (Figure 1). This field is in northern Calumet county, Wisconsin (Figures 1 and 2). Our study utilized a paired edge-of-field (EOF) water quality monitoring design to estimate the ability of a grassed strip to control erosion within the concentrated flow path of the treatment catchment, and reduce total suspended solids (TSS), total phosphorus (TP), and dissolved phosphorus (DP) export. The study consisted of three phases: pre-treatment, transitional and post treatment. The pre-treatment phase ended when statistical analysis of the events determined that there were enough events to detect a change. The transition phase consisted of planting and establishing the vegetated cover treatment in the concentrated flow path of the West catchment. The post-treatment phase began after this vegetated cover treatment was judged to be sufficiently established that it would reduce excessive soil erosion and prevent the formation of an ephemeral gully. The paired study design greatly reduces the influence of climate differences between the pre- and post-treatment phases. The adjacent paired plots were managed the same by the farm operator throughout the study, except for the addition of the grassed strip, thereby greatly reducing the possibility of an unintended effect due to a change in management.

Our EOF monitoring design followed similar protocols to those used for Wisconsin USGS EOF monitoring (Stuntebeck et al. 2008, 2011). Monitoring stations installed at the outlet of each catchment were configured to collect continuous discharge data and automated event samples. Flow from each of the paired field catchments was directed to H-flumes at each of the respective outlets. Flume stage, runoff volume and sampling information were monitored continuously, and recorded by a data logger. Discrete samples were collected during each runoff event by an automated sampler. Flow-weighted composite samples were created by taking sub-samples from each of the collected event samples in proportion to the flow runoff volume that occurred within each sample interval. A flow-weighted composited sample was used to represent the event-mean concentration (EMC) for each storm event. The cumulative flow and EMC were multiplied to calculate the total constituent load for each runoff event. Paired relationships between the East and West catchments were established for flow, TSS, TP and DP parameters during the pre-treatment phase. These relationships were compared to relationships during the post-treatment phase to determine if there were any changes. Detected changes that were determined to be statistically significant ($p < 0.05$) could likely be attributed to the treatment practice. A detailed description of study methods is provided in the methods section.

Site Description: The study site is within a farm field under a corn-grain soybean rotation, with low to moderate tillage. The GIS-estimated drainage areas of the East and West catchment monitoring sites are 6.36 and 8.52 acres, respectively (2.57 and 3.45 ha). These estimates were based on applying the watershed delineation tool of the Soil and Water Assessment Tool model (SWAT, Arnold 1997) to the 2018 elevation contours provided by the Calumet County Land and Information Office. Additional analysis was conducted with ESRI ArcMap 10.5 to obtain other watershed characteristics and to create related images displayed in this report. The mean, maximum and standard deviation of overland slope in the East catchment is 1.76%, 10.32% and 0.72%, respectively. The mean, maximum and standard deviation of overland slope in the West catchment is 1.68%, 9.47% and 0.71%, respectively. The average concentrated flow channel slope is 1.54% in both catchments (highest elevation minus lowest elevation along channel path, divided by channel length). This finding is somewhat surprising because

air photos over many years show that ephemeral gully formation in the west catchment was more frequent, deeper, and covered a greater area compared to the east catchment. The drainage area of the West catchment is estimated to be 34% greater than the East catchment; however, it is not certain that this factor alone explains the difference in concentrated flow erosion.

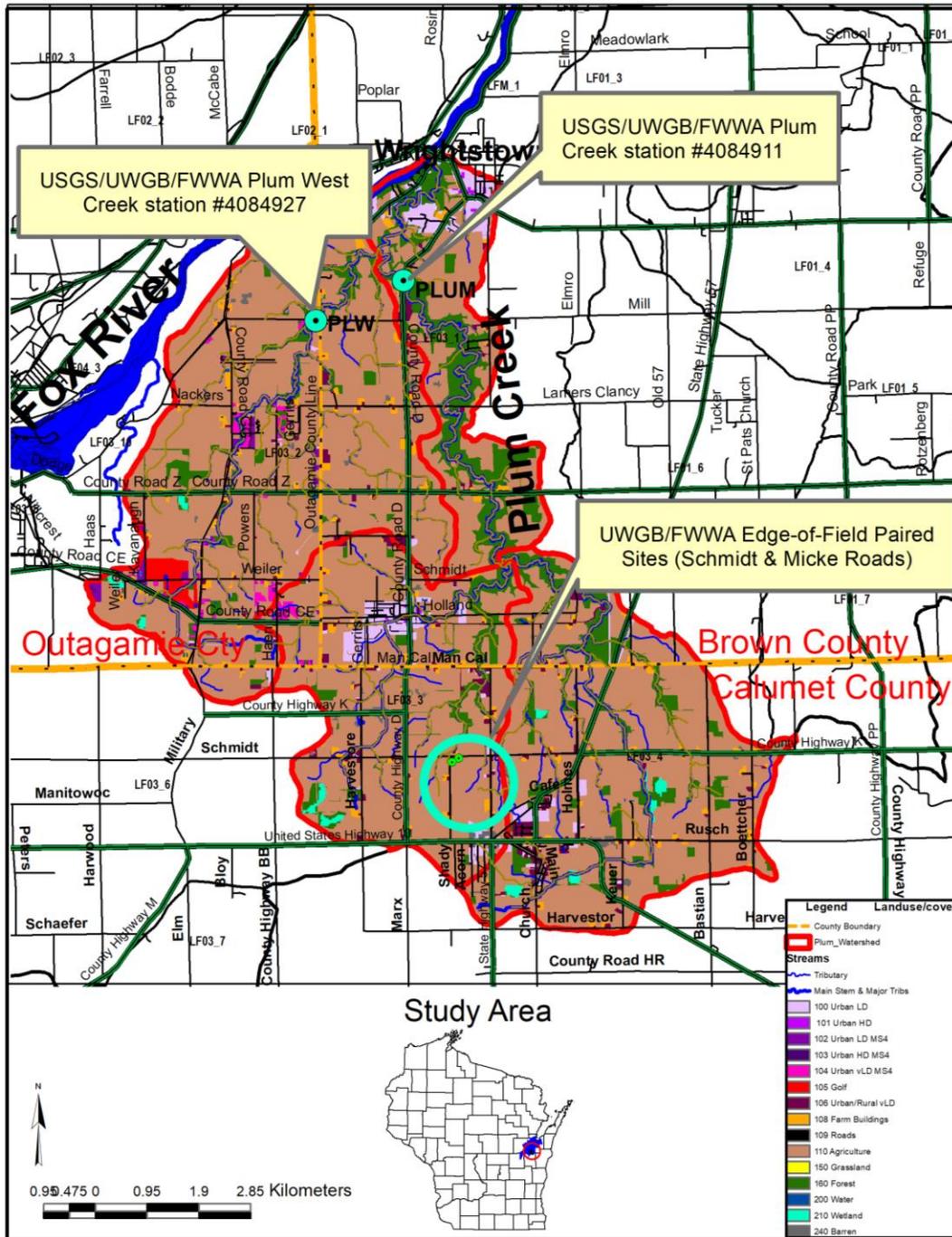


Figure 1. Plum Creek watershed and location of paired edge-of-field monitoring stations. Two USGS water quality monitoring stations that were also funded through the GLRI Fox-Wolf Watershed Alliance project are shown for reference.

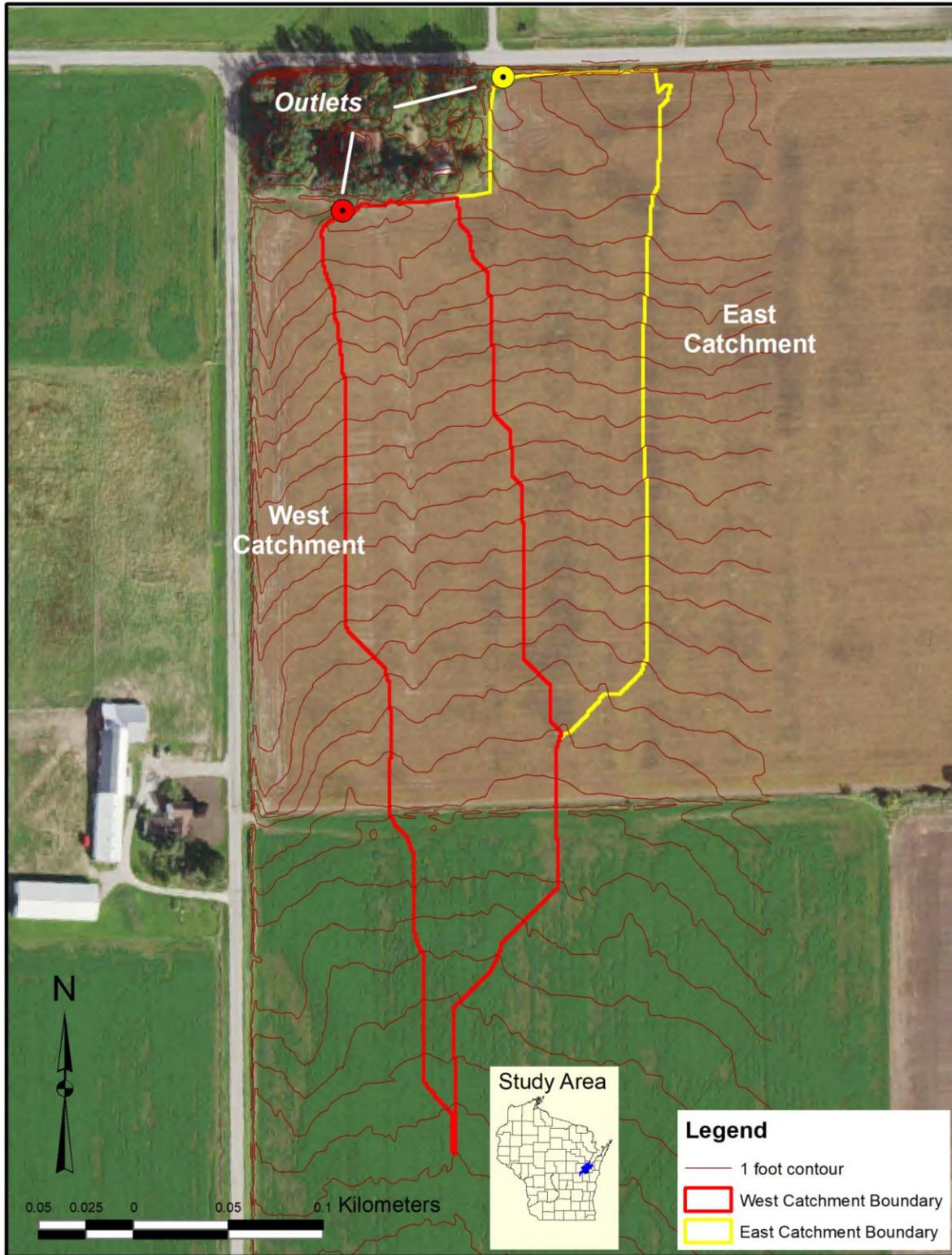


Figure 2. Study site with 1 foot elevation contours and east and west catchment boundaries and outlets to edge-of-field monitoring stations. Contour shapefile based on 2018 LIDAR was obtained from the Calumet County Land Information Office. Background is USDA FSA aerial orthophoto from 2017.

Soils within both paired catchments are primarily Kewaunee loam and hydrologic group C, which indicates a low infiltration rate. Five soil samples collected by UWGB in 2017 from each catchment were analyzed at the University of Wisconsin-Extension Soil and Forage Analysis lab for Soil Test P (STP). Mean STP was 21 and 27 mg/g Bray-P1 from the East and West sites, respectively. Additional soil samples were collected from the adjacent field that composed the uppermost portion of the catchments. STP from these soils was 32 and 26 mg/g Bray-P1 from the East and West sites, respectively. It is likely that STP concentrations in these catchments do not pose an excessive risk of contributing to high concentrations of runoff DP because they are within the optimum range for growing crops that are typically grown in this region, including corn, soybean, winter wheat and alfalfa (Laboski and Peters 2012).

METHODS

Water quality monitoring stations were installed at the outlets of two concentrated flow paths that drain East (6.35 acres) and West (8.62 acre) catchments, located in a farm field within the Plum Creek watershed. This field is in northern Calumet county, Wisconsin (Figures 1 and 2). The monitored catchments are directly adjacent to each other.

As previously stated, the concentrated flow path monitoring design followed similar protocols to those used by the Wisconsin USGS for edge of field monitoring (Stuntebeck et al. 2008, 2011). However, our study was not intended to cover all seasons. Therefore, winter runoff events when substantial ice was present were generally not captured, although occasional ice-chipping was done to clear the flumes for some cold-weather events that were monitored. The two monitoring stations are configured to collect continuous discharge data and automated event samples from their respective catchments. The EOF monitoring stations became operational in May 2016. A lockable 5' wide x 4' high, and 3' deep aluminum station house with a clam-shell opening was used to house the sampler, logger and related equipment at the outlet of each field plot (Figure 3). All electronic equipment was powered by two large capacity 12 volt DC deep recharge batteries that were charged by a 400 watt solar panel array. No power failures occurred during our study. Campbell Scientific CR-1000 data loggers provided systems control over all monitoring equipment. The USGS had previously provided a Campbell Scientific CR-Basic program for a stream monitoring project on the UWGB campus. We modified this program slightly to suit site-specific conditions for this study. Remote access via cell phone modems enabled real-time monitoring including setting sampling stage thresholds, sampling intervals, and downloading data as needed.



Figure 3. Station housing and monitoring equipment at outlet of west catchment. Turbidity probe is mounted below flume in half-pipe.

Surface water run-off was directed toward an H-flume by installing small berms and treated plywood wing-walls to direct runoff from the catchment flow channel to a single inlet where a 2.5 foot H-Flume served as the control structure. A nitrogen tank and Conoflow sight feed regulator assembly provided a steady supply of gas to a bubbler orifice line that was inserted near the H-flume stage height location and affixed at the bottom of the flume. A Sutron Accubar pressure transducer system measured how much pressure was in the line due to water depth, which was converted to feet and recorded by the datalogger. The higher the stage in the flume, the greater the pressure in the line. A 15-minute data recording interval was used during non-event conditions and a 5 minute interval during storm events.

The data logger program triggered the logger and sampler equipment to enter event monitoring mode at a user defined water level of 0.09 feet. Runoff-event samples were collected with refrigerated Teledyne ISCO 3700 automated samplers that were triggered to collect a discrete sample (~ 900 mL) in response to the initialization of event sampling mode, and stage changes during runoff events, via user-defined stage rise and fall thresholds. Samples were drawn through 3/8" i.d. polyethylene tube that was anchored 0.25 inches above the flume bottom, and 5 inches above the flume outlet. Heat tape was placed along the sampler line, with appropriate foam insulation, and into the flume to extend the monitoring period as much as possible. Samples were collected in 1000 mL semi-clear polyethylene sample bottles. A total of 6 to 18 discrete samples were collected from each EOF site during each event. Flume stage measurements recorded by the Accubar pressure transducer system and datalogger were converted to a record of runoff rate based on a standard 2.5 foot H-flume stage-flow table within an Excel spreadsheet. This spreadsheet was used to create a sample collection field log and sample processing log. Flow-weighted composite samples were created by taking sub-samples from each of the discrete samples collected during an event, in proportion to the runoff volume that occurred within each sample interval, as calculated in the spreadsheet. The flow-weighted composite sample was used to represent the event-mean concentration (EMC) for each runoff event. Cumulative flow volume and EMC were multiplied to calculate the total constituent load for each runoff event.

Sample Retrieval and Processing: Samplers were typically serviced by UWGB field personnel within 24-hours of the end of a storm event. The caps of the 1000 mL ISCO-3700 sample bottles were marked with a unique sample ID that was sequentially tracked by the data logger; for example, PF-E-6001 and PF-W-5001 represented the first ISCO samples from the east and west sampling stations, respectively. Sample bottles were then placed in coolers and processed in the UWGB water lab where they were composited based on flow volume. Flow composited samples therefore represented the mean concentration for each event and site. Flow composited samples were split with a Decca 10-port splitter to divide the sample prior to subsample preservation and analyses. Subsamples included four separate smaller polyethylene bottles that were labeled for TSS (~400 mL), TP (~100 mL), and DP (~100 mL) analysis; plus an extra bottle (~400 mL) was kept for later retesting or confirmation if issues came up with a missing, or unexpected result from the lab. The DP subsample was filtered with a 0.45-micrometer pore size filter prior to preservation. All water sample bottles except the TSS and reserved samples were preserved with H₂SO₄. A unique sequentially numbered sample ID label was placed on each of the composited bottle samples, based on event order (e.g, PF-E-101 and PF-W-101 represented

the first event composites from the east and west stations). A Chain-of-Custody (COC) form was filled out with the sample date, time, sample ID, and parameters to be analyzed. Samples were then transported 4.5 miles to the New Water Laboratory for analysis. The NEW Water lab is USGS accredited and approved by the USGS Branch of Water Quality Systems, and it is certified by the State of Wisconsin under Ch. NR 149, Wis. Adm. Code by the Wisconsin DNR Bureau of Science Services Laboratory Certification and Registration Program. All samples were analyzed by NEW Water, except those collected after mid-March 2020, when samples were directly delivered to Pace Analytical lab in Bellevue, Wisconsin for analysis due to COVID-19 restrictions at NEW Water lab. These 6 sets of samples were refrigerated and preserved per standard protocol, but delivery was delayed for several months until we were certain that analysis by NEW Water would be delayed indefinitely.

Lab results in Adobe Acrobat format were usually received about 10 days after sample drop-off via email. These results were checked for potential errors based on comparisons between the paired samples, and the three constituents. Retests were requested if there were any issues. An Excel spreadsheet of all NEW Water analytical data that UWGB requested was occasionally received and these data were used to confirm the originally transcribed lab results. Resulting constituent concentrations from the composited samples and the discharge record from each event were used to calculate event-mean concentrations and loads from each of the concentrated flow channels. The paired data were then used to assess the effectiveness of the grassed strip flow channel compared to the standard tilled concentrated flow channel.

Flume Maintenance and Levels: Debris and dirt buildup were removed from flumes during sample collection: both were relatively minimal during all events, and did not affect sample water quality, stage height or flow velocity. Flume levels were measured when event samples were collected at both stations to ensure accurate flow volume measurements which require that the H-flumes be level in both directions (side to side, and inlet to outlet). In addition, flume levels were checked periodically, including late winter and early spring prior to expected spring runoff, because flume height and level are often displaced by frost heave. Corrections were made when the flumes were not level. Nearly all corrections were minor, except when ground frost melted, and these changes typically took place prior to a major spring runoff event. One major correction at the inlet of the East station was required because frost heave caused about a 100 mm increase in elevation above the soil surface at the flume inlet. This issue did not affect the flume level, because the flume outlet level was adjusted accordingly. However, this problem could have reduced runoff volume because of increased ponding, and potentially increased suspended sediment deposition above the flume. Therefore, it was deemed enough of an issue by spring 2020, that it was remedied 4/15/20 by cutting a portion of the plywood that holds the flume and lowering the flume to the original installation state, near the soil surface level.

Direct stage as measured inside the flume, and stage as measured by the pressure transducer system were compared and recorded when there was sufficient water in the flume to perform this check during site visits. An offset was used to correct the recorded stage based on these measurements. In general, this offset amounted to subtracting about 0.014 feet and 0.026 feet (on average) from the recorded stages at the East and West stations, respectively.

Precipitation: Calibrated eight inch diameter Rain Wise and Texas Instrument tipping bucket rain gauges were employed to measure non-frozen precipitation (0.01 inch increments) near each of the station houses, and this data was recorded by the data logger. Precipitation intensity and volume data were utilized to characterize those periods responsible for runoff and erosion, but these parameters were not critical to success of the study.

Turbidity: A second means of potentially computing event mean constituent concentrations and loads was employed by combining continuous discharge data with continuous turbidity data from a Campbell Scientific OBS-501 turbidity probe that measures both backscatter and sidescatter turbidity units (up to 4000 BTU and 1000 STU, respectively). This probe has a retractable head to protect the optics from fouling and it greatly decreases clogging normally associated with standard wiper mechanisms. The probes were mounted vertically such that the sensor end was placed in a half-pipe below the flume outlets (with a “weir” plate to maintain sufficient water depth) until the end of 2018, after which they were placed in the upper part of the flume, opposite of the pressure transducer bubbler line. These two turbidity probes were deployed from mid-June 2016 until they were retrieved in October 2020.

A total of 101 and 129 discrete samples from the East and West stations, respectively, were analyzed for TSS (or SSC), of which 87 samples were analyzed for TP. Chemical analysis was conducted at either the NEW Water (TSS and TP) or UWGB (SSC and TP) labs. These results were used to establish regression-derived relationships between concentration and concurrent in-situ turbidity measurements, which were then used to estimate continuous concentrations throughout the study period. For each event and site, the continuous flow data were combined with turbidity-derived continuous estimates of TSS (or SSC) and TP concentrations to calculate the event-mean concentrations and loads of TSS (or SSC) and TP. Loads were computed in a spreadsheet that sums the 5 minute instantaneous loads computed over the duration of the storm. If the relationships are robust, the resulting event loads and event-mean concentrations can offer an alternative means of assessing the efficacy of the grass waterway (compared to the untreated concentrated flow channel).

In addition to the commercial turbidity probe deployment, UWGB also constructed, bench-tested and deployed about 10 low-cost turbidity probes within the Lower Fox River sub-basin for in-situ testing, including the two stations for this investigation. The low-cost turbidity probes were mounted vertically inside the upper portion of the flumes, alongside the commercial probes, which had previously been placed in a half-pipe below the flume outlets (with a “weir” plate to maintain sufficient water depth). at the East and West catchment monitoring stations for in-situ testing. Except for some caveats (like interference from sunlight), these low-cost probes performed reasonably well. More information can be found in the MS thesis of Schmitz (2020) which investigated the development and deployment of low-cost turbidity probes at many sites in the Lower Fox River Watershed, including the two sites in this report.

Total Kjeldahl Nitrogen (TKN): TKN was originally included as a non-critical constituent to track. However, TKN was no longer analyzed once it was determined that a) it was not a critical parameter that was crucial to the success of the study, per QAPP (2015); b) total nitrite-nitrate would be missed in this analysis, and accounted for about half of the total nitrogen; c) doing both TKN and total nitrite-nitrate would double the cost of nitrogen series analysis; and d) an existing stream at USGS East River CTH ZZ had all water samples analyzed for both constituents, and there was an excellent relationship between TP and TKN; thereby negating the need to duplicate these efforts.

Quality Control: Field blank results and duplicate sample quality checks are included in Appendix B. Copies of field station logs, event sample processing and field collection logs, and lab analysis chain of custody forms are included in Appendix C, D and E, respectively, as attachments to this report.

Crops, Tillage and Nutrients: The field plots were managed the same by the farm operator throughout the study, except for the addition of the grassed strip. Cropping practices consisted of soybean in 2015; corn in 2016, soybean in 2017; corn in 2018; no crop in 2019 except a late “sacrifice” soybean crop followed by a late cover crop; soybean in 2020 and winter wheat in 2021 (planted fall 2020). Corn was harvested as corn grain, not silage. Corn grain and soybean are generally equally erosive row crops, particularly when moderate to aggressive tillage is the dominant practice. Light to moderate spring tillage took place prior to planting. Fall tillage occurred in all years, except when it was too wet in Fall 2018, when the corn crop was not harvested until the soil froze. The farm operator was not going to till the field the following spring because no crop was intended to be harvested due to the continued wet conditions. However, there was a substantial amount of crop residue from the corn crop that was harvested after the ground froze, so the farm operator helped support the project objectives by tilling the field in late spring 2019 to ensure that the tillage practices remained reasonably consistent throughout our study period. Commercial fertilizer was the primary form of nutrients that were applied to the study site, with incorporation at planting or with tillage. However, solid beef manure with a high quantity of bedding material was applied over about one-third of the site field in January or February 2017, which was not incorporated until spring tillage.

Treatment Phase Initiation: Statistical analysis was conducted with the SAS 9.4 statistical analysis program in early 2018, which indicated that data quality and quantity were sufficient to proceed with transitioning to the treatment phase. Therefore, on March 15 2018, the concentrated flow area in the West treatment catchment was frost seeded with winter rye, oats and fescue ground cover plant varieties, with additional seeding later to ensure good cover conditions. Unfortunately, the ground cover in the concentrated flow path area was greatly disturbed by ephemeral gully erosion in early spring (Figure 4), and portions of the grassed strip were killed by inadvertent spraying of herbicide and tillage by the farm operator during spring planting. Therefore, the concentrated flow path was reseeded in late spring, with occasional reseeding to ensure better coverage of trouble areas throughout 2018. Tillage equipment was used to fill in the ephemeral gully. Brown and Outagamie County LCD installed and maintained the grassed strip, with assistance from the farm operator. Although the grassed strip cover was not 100% effective yet, it was determined to be sufficient to start treatment phase monitoring with the 8/27/18 event, by which time, 38 events had been collected for the pre-treatment

phase of the study (6/15/16 to 6/18/18 events). Additional reseeding and limited leveling was re-done in spring and summer 2019 to ensure good cover and to remove a small gully that had reformed. Because of the frequent runoff events, it was difficult to get good cover throughout the study, so protection was never 100% (Figure 5). It was especially difficult to establish more cover during the extremely wet conditions in 2019; however, by late fall 2019 the grassed strip was well developed (Figure 6).



Figure 4. Ephemeral gully erosion in West catchment concentrated flow path, May 10, 2018. Frost-seeded grassed strip planted in early spring was not well enough established to prevent substantial gully erosion, so it was replanted in late spring and the treatment phase was delayed until fall.

A narrow ephemeral gully was observed to be present in the lower section of the gully in fall 2019 and very early spring 2020. This gully deepened greatly in some spots sometime in mid-spring 2020, so we did not expect consistent treatment effect in our paired study after this period. Therefore, we initially decided, and later concluded that all 8 events sampled in 2020 would be excluded from our statistical analysis (these 8 paired sets of analyzed samples are excluded from Table 4, in the results section). The treatment phase started with the 8/27/18 event and ended with the 12/30/19 event.

Related UWGB Project: Kalk (2018) conducted an evaluation of the APEX model to simulate runoff, sediment, and phosphorus loss from agricultural fields in northeast Wisconsin, which included the two sites in our present study.



Figure 5. Photo of grassed strip looking upslope of West catchment outlet (8/6/19). Although this photo was taken during the treatment phase, the grassed strip is not fully protective, as some ephemeral gullies have formed at this spot, and further upslope.



Figure 6. Photo of grassed strip looking further upslope of West catchment outlet during the treatment phase (11/20/19), when the grassed strip was more protective than in August 2019. However, it was still not fully protective, as some ephemeral gullies were still present at this location such that grass was not fully established in the channel center along the entire length of the strip.

RESULTS and DISCUSSION

A weight-of-evidence approach was utilized to judge treatment effectiveness to ensure that a finding of statistical significance alone would not be mis-interpreted. Therefore, other lines of evidence and evaluation were included besides the Analysis of Covariance (ANCOVA) test recommended by Clausen and Spooner (EPA 1993) for a paired watershed study design: a) sequential event plots showing before and after treatment periods; b) plots of ranked differences before/after treatment; c) before/after treatment boxplots; d) scatter plots of east versus west catchments before/after treatment; e) simple regression plots to visualize before/after treatment differences; and f) non-parametric Wilcoxon Ranked sum difference test. Statistical analysis was primarily conducted with the SAS program. SAS was also used to confirm results of regression analysis performed in MS Excel spreadsheets.

Event runoff volume, precipitation, and TSS, TP and DP EMC's are listed by event in Tables 1 for the pre-treatment period and Table 2 for the post-treatment period. Precipitation listed in Tables 1 and 2 are based on either the East or West station gauge, depending on which one was determined to provide the most accurate result, because there were times when the wind and nearby trees affected the precipitation amount, or one of the gauges was plugged, etc. Runoff volume in mm is based on the total event runoff (flow) volume divided by the estimated drainage area of each catchment. An expanded table that includes these parameters, plus flow in liters, and loads and yields is included in the Appendix A. Three sets of paired DP concentrations were not included in the analysis or in Tables 1 and 2 because they were deemed outliers: a) 0.065 mg/L at West plot because it was very low relative to TP of 0.336 mg/L, especially given the relatively low TSS concentration of 59 mg/L; b) 1.2 mg/L at West plot on 7/15/17 due to likely excessive commercial fertilizer that had been recently applied on the surface of the hay field just upslope and south of the study farm field per communication with Outagamie LCD staff; and c) 0.863 mg/L at West plot on 2/25/18 due to possible excessive commercial fertilizer or manure that was recently applied on the surface of the hay field just upslope and south of the study farm field per communication with Outagamie LCD staff.

Prior to the treatment phase, the mean and median West catchment TSS EMC were about twice as high as the East Catchment, and the mean and median West catchment TP EMC were about 21% and 43% higher than the East catchment EMC, respectively (Table 1). After the grassed strip was added to the West catchment (treatment phase), the mean and median West catchment TSS EMC were about 26%, and 37% higher than the East Catchment, respectively (Table 2). After the treatment phase, the mean and median West catchment TP EMC was about 5% and 16% higher than the East Catchment EMC, respectively (Table 2). Mean and median DP EMC from the catchments were similar, before and after treatment.

TSS and TP event-mean concentrations during the pre-treatment (6/6/16 to 5/10/18) and post-treatment (8/27/18 to 12/29/19) periods are plotted by event in Figures 7 and 8 for both East and West plots. A total of 38 events were captured during the pre-treatment phase, and 34 events during the post-treatment phase. Event-mean concentrations of TSS and TP at the East and West plots were similar until an ephemeral gully re-formed in the West plot concentrated flow channel, after which there was a marked increase in concentrations from the West plot during the 5/1/17 event (Figures 7 and 8). This pattern changed after the vegetation in the grassed strip became established, such that the concentrations at both plots were similar from the 8/27/18 event onward, thereby indicating that there was a relative decline in TSS and TP EMC's from the West treatment due to the grassed strip.

Table 1. Event mean constituent concentration (mg/L) and runoff volume (mm) at East and West Paired Field Catchments: pre-treatment period.

		Runoff		TSS		Total Phosphorus		Dissolved Phosphorus		
Mean		6.4	10.2	448	888	1.11	1.59	0.27	0.25	
Median		3.5	7.0	164	330	0.86	1.04	0.23	0.21	
maximum		30.4	52.7	2,540	3,090	3.60	4.28	0.61	0.59	
Event	#	Rain	East	West	East	West	East	West	East	West
6/15/16	1	47.8	8.7	8.1	2,540	2,960	3.27	3.73	0.23	0.26
9/22/16	2	47.5	2.6	6.7	122	344	0.92	1.22	0.45	0.41
10/12/16	3	11.9	0.8	2.1	183	104	1.15	1.05	0.61	0.49
10/26/16	4	33.3	9.2	13.2	632	736	1.89	1.99	0.53	0.50
11/28/16	5	15.7	2.8	4.7	144	128	0.88	1.09	0.42	0.32
11/29/16	6	18.8	11.2	11.7	96	32	0.60	0.75	0.17	0.20
12/6/16	7	5.8	1.2	1.5	5	17	0.30	0.29	0.24	0.11
2/18/17	8	melt	1.7	5.6	22	51	0.43	0.35	0.32	0.21
2/21/17	9	5.6	2.8	4.7	31	59	0.46	0.35		
2/28/17	10	11.7	12.3	17.9	17	33	0.30	0.25	0.20	0.11
3/6/17	11	melt	16.3	19.3	8	15	0.22	0.19	0.18	0.10
3/7/17	12	4.1	2.6	4.6	49	95	0.53	0.42	0.31	0.13
3/24/17	13	13.7	1.4	7.5	27	130	0.55	0.44	0.35	0.12
3/26/17	14	13.7	8.6	17.9	33	49	0.28	0.36	0.19	0.15
3/30/17	15	7.1	3.1	7.0	17	25	0.35	0.20	0.24	0.13
4/4/17	16	15.0	6.5	13.8	47	76	0.32	0.29	0.16	0.14
4/15/17	17	18.3	0.3	3.3	86	188	0.83	0.93	0.42	0.23
4/16/17	18	10.2	3.6	7.0	190	246	0.81	0.88	0.23	0.14
4/20/17	19	22.6	9.9	15.1	326	250	1.72	1.49	0.25	0.28
4/27/17	20	23.9	5.9	11.7	40	62	0.69	0.73	0.27	0.24
5/1/17	21	20.6	10.6	14.8	296	316	1.12	1.01	0.18	0.17
6/3/17	22	27.9	0.7	1.6	1,887	2,660	3.34	4.28	0.53	0.40
6/4/17	23	12.4	4.9	5.9	1,873	2,770	3.60	4.20	0.44	0.38
6/14/17	24	20.6	1.4	2.5	828	1,580	1.74	2.37	0.30	0.25
6/22/17	25	15.0	0.3	0.6	884	2,140	1.73	2.87	0.20	0.18
6/22/17	26	11.2	6.5	6.1	1,920	2,850	2.42	3.60	0.19	0.19
6/23/17	27	31.0	25.0	26.0	256	540	0.71	1.02	0.17	0.28
6/29/17	28	16.5	1.4	1.6	792	1,610	1.95	2.89	0.20	0.21
7/2/17	29	15.5	2.0	1.6	832	2,800	1.59	2.88	0.15	0.23
7/15/17	30	36.6	11.6	10.5	448	916	0.98	2.39		
7/18/17	31	8.9	0.5	0.8	484	1,867	1.14	2.74	0.22	0.29
2/25/18	32	10.2	3.4	12.6	490	1,300	1.51	3.13		
4/21/18	33	melt	11.7	15.6	11	108	0.12	0.19	0.11	0.09
4/22/18	34	melt	7.8	21.8	96	610	0.20	0.96	0.10	0.23
5/2/18	35	31.5	0.7	2.5	80	785	0.65	1.26	0.23	0.18
5/3/18	36	51.6	30.4	52.7	555	1,850	1.16	2.55	0.12	0.50
5/9/18	37	25.4	2.9	6.1	141	344	0.50	0.74	0.14	0.20
6/18/18	38	76.7	11.2	21.0	530	3,090	1.23	4.21	0.41	0.59

Table 2. Event mean constituent concentration (mg/L) and runoff volume (mm) at East and West Paired Field Catchments: post-treatment period.

			Runoff		TSS		Total Phosphorus		Dissolved Phosphorus	
mean			7.1	13.6	358	452	0.90	0.95	0.21	0.24
median			6.4	12.2	262	359	0.80	0.94	0.19	0.21
maximum			19.3	31.4	1,806	1,450	3.34	2.06	0.52	0.56
Event	#	rain	East	West	East	West	East	West	East	West
8/27/18	39	140.2	13.4	21.5	206	304	0.64	0.82	0.33	0.44
9/3/18	40	75.2	19.3	30.8	212	368	0.79	1.06	0.33	0.42
9/5/18	41	10.7	2.8	7.3	66	86	0.67	0.78	0.47	0.56
10/3/18	42	10.7	0.9	2.4	340	482	1.29	1.30	0.52	0.43
10/8/18	43	24.4	8.5	20.9	500	520	1.32	1.31	0.37	0.50
10/10/18	44	23.4	5.7	13.6	204	350	0.71	1.07	0.32	0.49
11/4/18	45	15.7	4.3	6.2	110	115	0.63	0.51	0.29	0.28
11/5/18	46	15.7	7.1	14.7	332	296	0.94	0.85	0.21	0.17
4/12/19	47	9.7	4.7	8.5	4	6	0.14	0.09	0.11	0.08
4/17/19	48	12.4	3.1	6.6	5	15	0.15	0.12	0.12	0.08
4/22/19	49	9.7	1.4	3.7	8	17	0.18	0.13	0.12	0.07
5/1/19	50	8.9	2.7	6.4	7	13	0.15	0.14	0.10	0.08
5/8/19	51	26.7	4.6	9.9	15	23	0.23	0.20	0.16	0.11
7/19/19	52	48.8	2.9	5.4	1,806	1,333	3.34	2.06	0.24	0.26
7/20/19	53	7.6	2.2	7.5	674	880	1.64	1.63	0.26	0.29
8/3/19	54	61.7	8.2	11.9	979	1,450	1.54	1.77	0.20	0.25
8/5/19	55	45.0	9.5	16.9	965	1,060	1.77	1.78	0.22	0.25
8/7/19	56	28.7	7.2	12.5	1,120	1,090	2.09	1.98	0.21	0.23
9/10/19	57	34.8	8.1	16.2	119	394	0.54	0.79	0.19	0.20
9/11/19	58	29.5	11.6	20.7	444	810	0.94	1.21	0.19	0.18
9/12/19	59	20.6	7.6	13.8	420	647	0.94	1.06	0.21	0.19
9/19/19	60	22.1	4.2	8.9	340	532	0.88	1.13	0.21	0.24
9/22/19	61	48.0	13.1	23.5	332	853	0.81	1.39	0.15	0.17
10/1/19	62	9.4	2.8	7.8	368	368	0.96	1.02	0.19	0.25
10/2/19	63	29.0	16.2	31.4	588	692	1.22	1.26	0.16	0.16
10/3/19	64	7.4	3.5	7.7	140	126	0.64	0.51	0.17	0.20
10/5/19	65	7.9	1.9	5.1	171	210	0.82	0.76	0.16	0.18
10/11/19	66	30.5	9.0	18.5	416	576	0.96	1.04	0.14	0.14
10/21/19	67	10.7	2.4	2.9	260	280	0.74	0.73	0.14	0.12
10/27/19	68	10.4	2.6	4.3	264	218	0.79	0.72	0.20	0.18
11/20/19	69	31.0	16.4	27.4	165	252	0.48	0.63	0.19	0.26
11/27/19	70	22.1	10.2	18.3	407	720	1.03	1.42	0.13	0.21
12/29/19	71	18.5	14.0	30.2	123	196	0.38	0.50	0.10	0.18
12/30/19	72	14.2	9.2	20.0	51	77	0.25	0.47	0.10	0.27

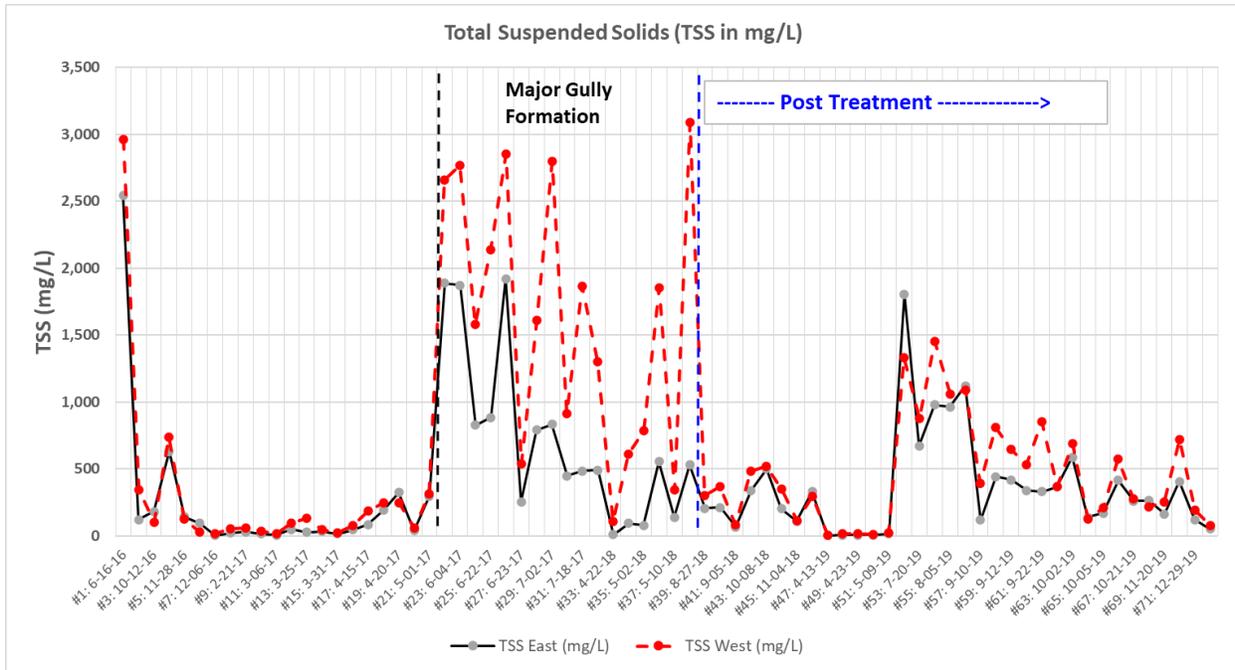


Figure 7. Event-mean concentrations of TSS by event, from the East and West catchments: pre- and post-treatment. West and east catchments are similar after grassed strip installation in West catchment.

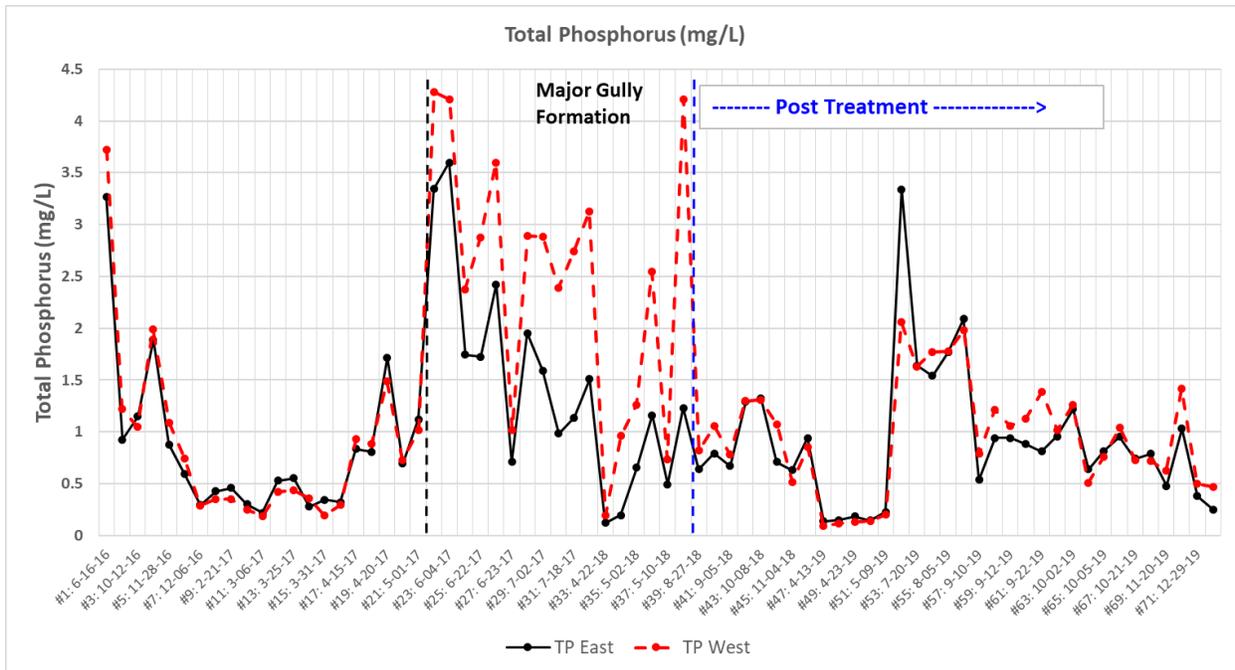


Figure 8. Event-mean concentrations of total phosphorus by event, from the East and West catchments: pre- and post-treatment. West and east catchments are similar after grassed strip installation in West catchment.

Plots of ranked events of event-mean TSS and TP concentration differences between the West and East plots are shown in Figures 9 and 10, respectively. These plots compare the pre-treatment and post-treatment periods to illustrate any change in the difference between the two plots. The displayed values are calculated by subtracting the East plot event-mean concentrations from the West plot, and then ranking these values from the largest to the smallest observed difference. This procedure was used to better understand the pattern and importance of observed differences between the two periods. As can be seen in Figure 9, differences in TSS event-mean concentrations are much greater during the pre-treatment period compared to the treatment period. During the pre-treatment period, a total of 12 events had differences greater than 550 mg/L TSS, with the highest being 2,600 mg/L. In contrast, the difference never gets above 550 mg/L TSS, and mostly hovers a bit above and below zero difference during the treatment period. A similar pattern was observed for TP (Figure 10), where the pre-treatment period had a total of 14 events with differences greater than 0.60 mg/L TP, with the highest being nearly 3.0 mg/L TP. In contrast, the difference never gets above 0.60 mg/L TP during the post-treatment period.

Boxplots of event-mean concentrations of TSS and TP in Figures 11 and 12, respectively, illustrate this same pattern where the concentrations in the West plot are higher than the East plot during the pre-treatment phase, but they are very similar during the post-treatment phase. This pattern is especially apparent for the range, which should be expected given that ephemeral gully impacts are not expected to be consistent from event to event. Little difference between periods was observed for event-mean concentrations of DP (Figure 13), which was expected given that installing a grassed strip in the concentrated flow channel should have little impact on soluble constituents like DP. Interestingly, water runoff volume seemed to increase in the West plot, relative to the East plot, during the treatment period (Figure 14).

The plotted relationships between the East and West plots for runoff volume (mm), and event-mean concentrations of TSS, TP and DP are shown in Figures 14a to 14d for the pre-treatment and post-treatment periods. These same relationships are plotted in natural log space in Figures 15a to 15d. These plots include pre- and post-treatment periods, for which the associated best-fit trend regression lines and R^2 statistics are shown (and verified with SAS). The relatively high R^2 statistics lend credence to the validity of the relationships between the East and West catchments, for both TSS and TP. Both normal and log-space regressions visually indicate that it is likely that the relationship between the East and West catchment changed (pre- vs post-treatment), such that both TSS and TP likely declined in the grassed strip treated West catchment. As with the boxplots, little difference between periods was observed for event-mean concentrations of DP.

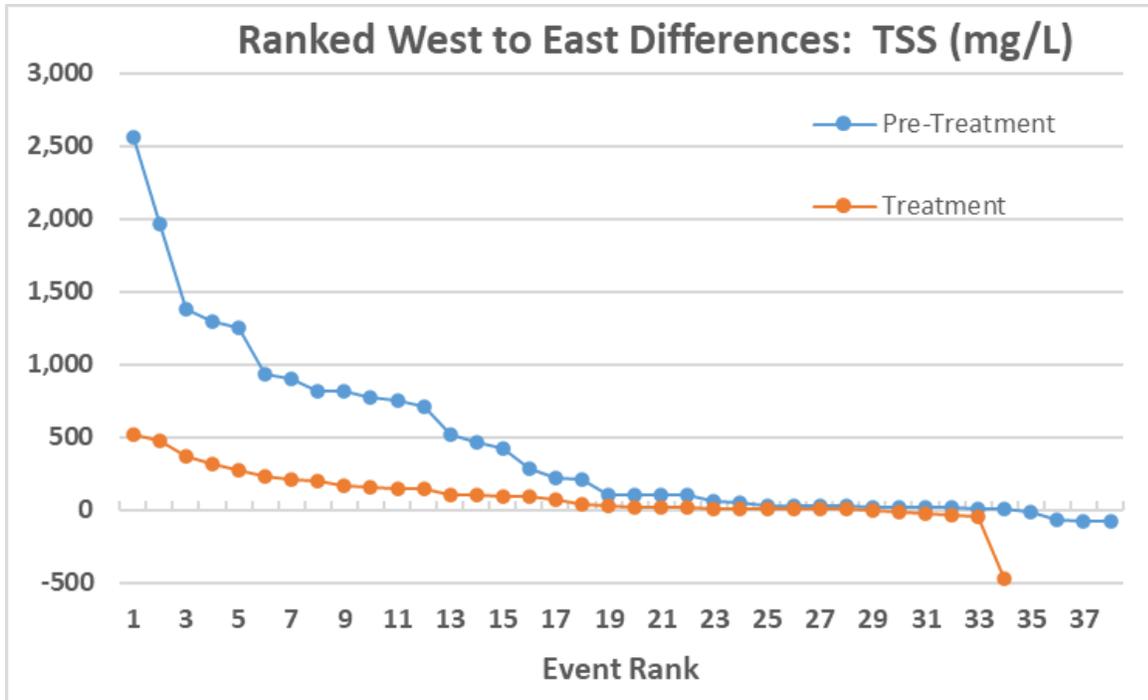


Figure 9. Ranked West to East catchment differences of TSS event-mean concentrations (mg/L): pre and post treatment. West catchment had grassed strip in flow channel during post-treatment period.

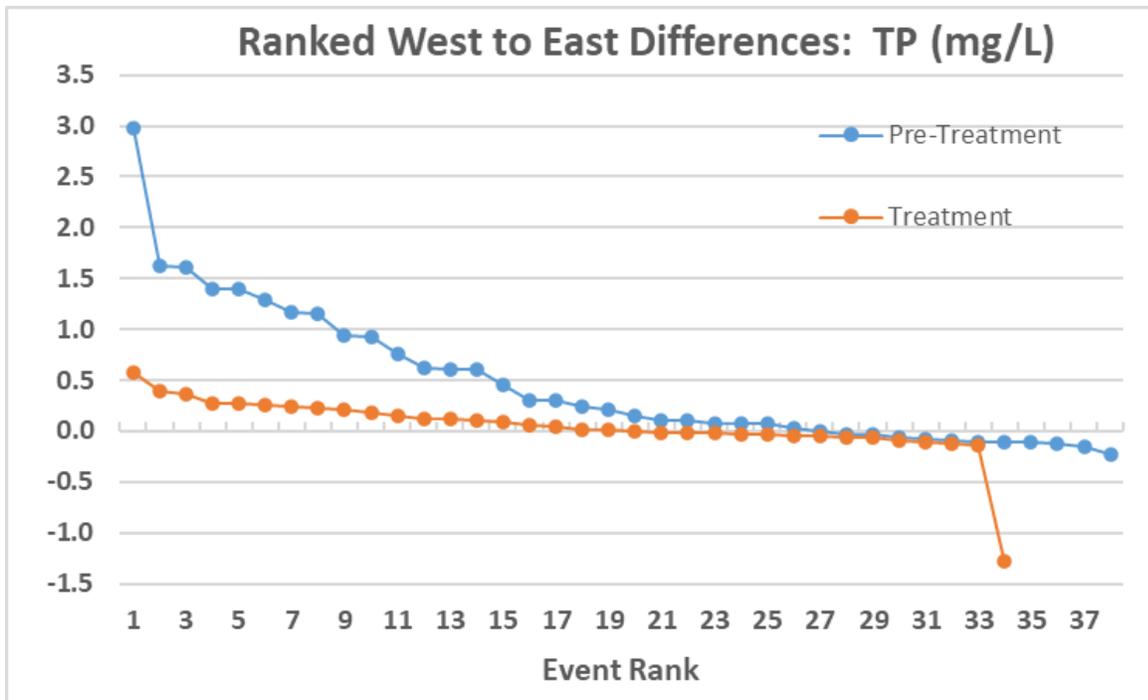


Figure 10. Ranked West to East catchment differences of total phosphorus event-mean concentrations (mg/L): pre and post treatment. West catchment had grassed strip in flow channel during post-treatment period.

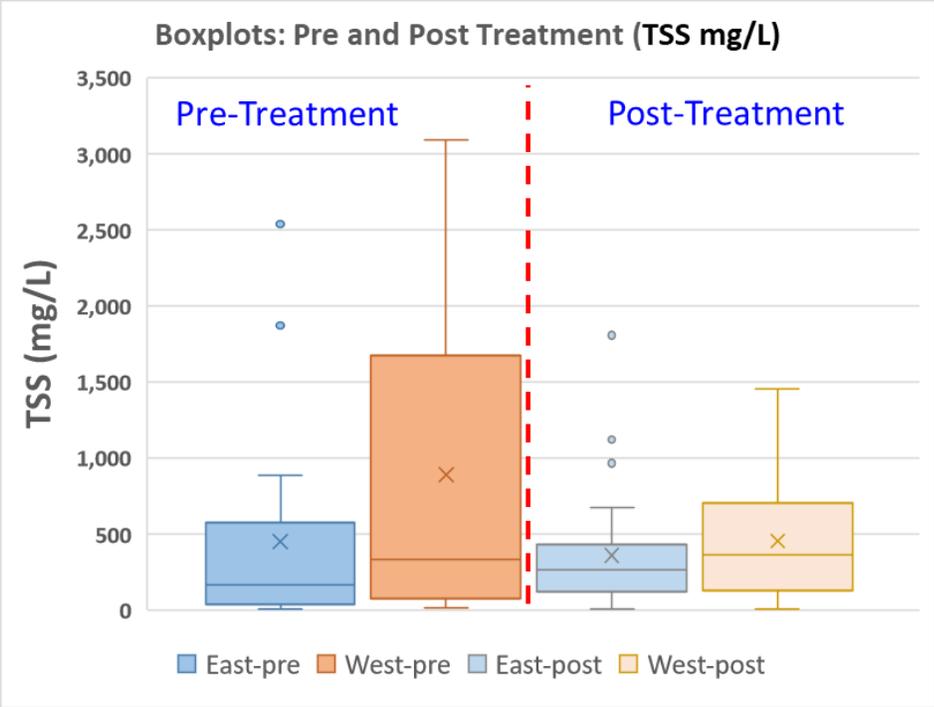


Figure 11. Boxplots of pre and post treatment TSS event-mean concentrations at the east and west catchments. West catchment had grassed strip in flow channel during post-treatment period.

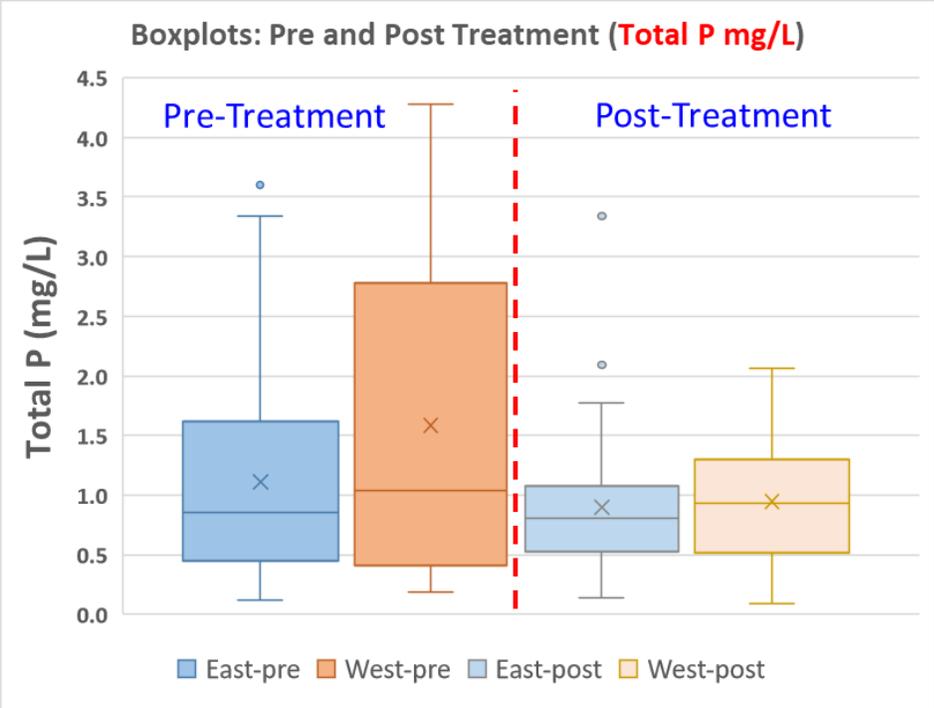


Figure 12. Boxplots of pre and post treatment TP event-mean concentrations at the east and west catchments. West catchment had grassed strip in flow channel during post-treatment period.

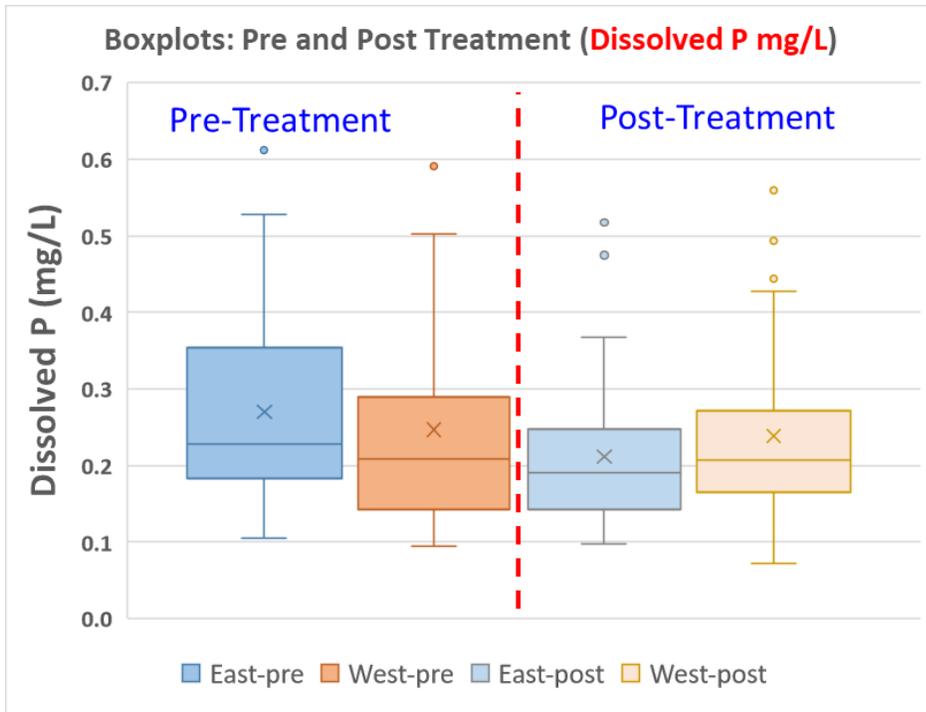


Figure 13. Boxplots of pre and post treatment DP event-mean concentrations at the east and west catchments. West catchment had grassed strip in flow channel during post-treatment period.

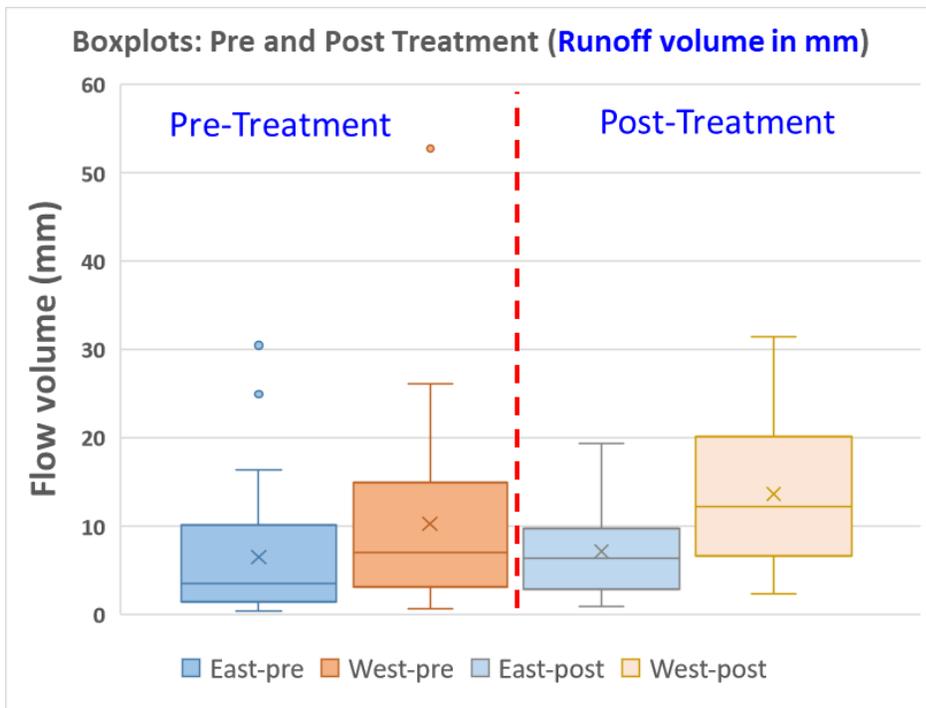


Figure 14. Boxplots of pre and post treatment runoff volume at the east and west catchments. West catchment had grassed strip in flow channel during post-treatment period.

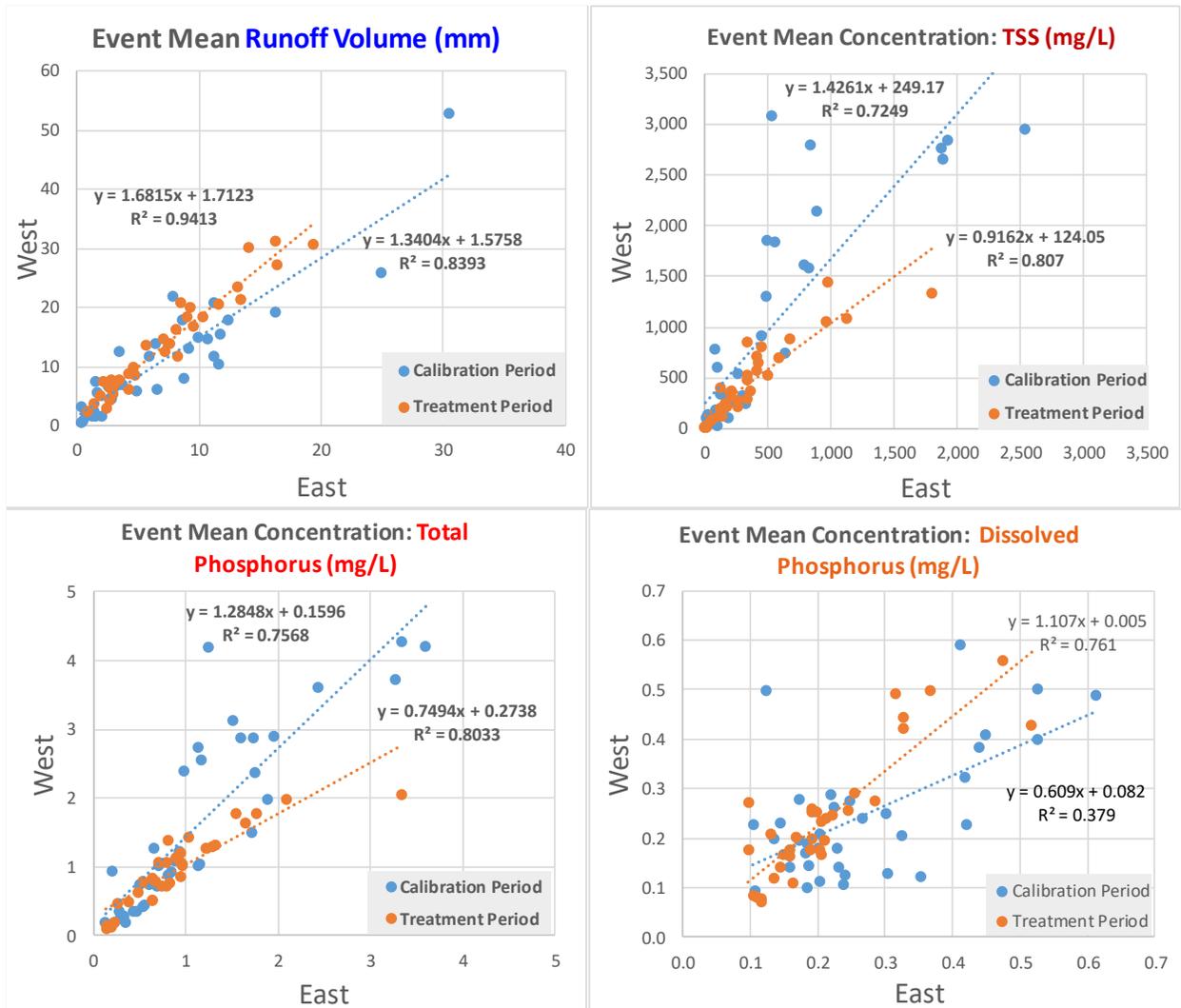


Figure 15. Regression plots of East and West paired EOF catchments for runoff volume (n=72), and event-mean concentrations of TSS (n=72), TP (n=72) and DP (n=69) during 38 calibration (pre-treatment) period events and 34 treatment period events.

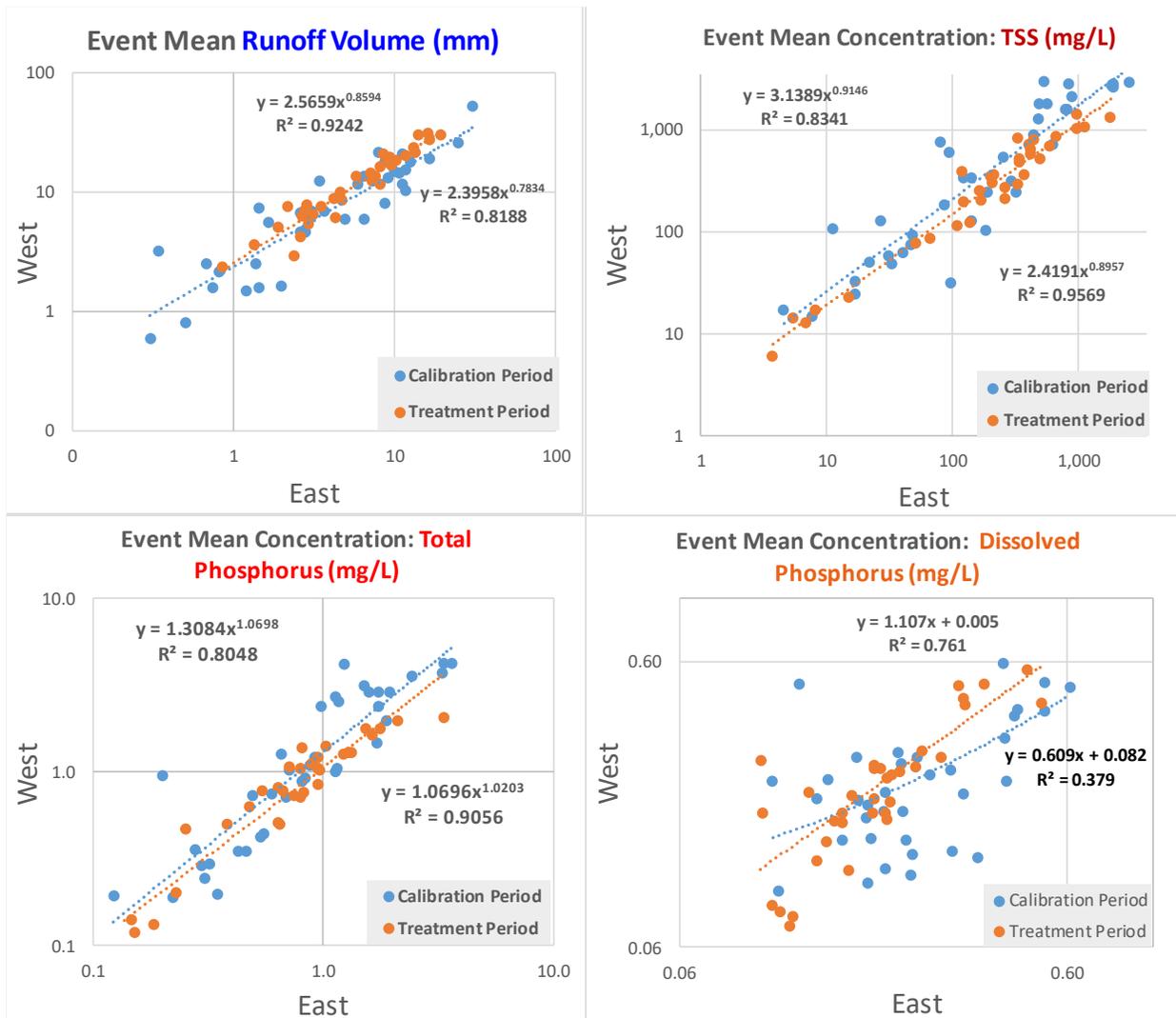


Figure 16. Log-space regression plots of East and West paired EOF catchments for runoff volume (n=72), and event-mean concentrations of TSS (n=72), TP (n=72) and DP (n=69) during 38 calibration (pre-treatment) period events and 34 treatment period events.

Statistical analyses with ANCOVA were performed on the data sets to test for significant differences between pre- and post-treatment periods for TSS, TP, DP and runoff volume, using the methods described by Clausen and Spooner (1993) for a paired watershed study design. Natural log transformed event-mean concentrations of TSS ($p = 0.007$) and TP ($p = 0.038$) were found to be significantly lower in the West catchment during the grassed strip treatment period, as compared to the pre-treatment calibration period (Table 3). No significant difference was detected for natural log-transformed DP concentrations ($p > 0.05$). This finding was expected because installation of a grassed strip in a concentrated flow path should have little effect on dissolved constituents. ANCOVA plots with 95% confidence limits are shown in Figures 17 for TSS and in Figure 18 for TP.

Table 3. ANCOVA and regression ANOVA statistics for comparisons of natural log event mean concentrations and mass between pre and post treatment periods (all data natural log transformed).

ANCOVA	TP-conc	TSS-conc	dP-conc	Flow	TP-mass	TSS-mass	dP-mass
p-Values (<i>bold italics are significant differences at the 0.05 level</i>)							
Type I	<i>0.038</i>	<i>0.007</i>	0.167	<i>0.027</i>	0.631	0.296	<i>0.009</i>
Type III	<i>0.036</i>	0.544	<i>0.008</i>	0.594	0.132	0.068	0.099
interaction (with period)	0.649	0.821	<i>0.018</i>	0.416	0.055	0.131	0.301
LS-means	<i>0.037</i>	<i>0.007</i>	0.147	<i>0.035</i>	0.600	0.283	<i>0.010</i>
	Post < Pre	Post < Pre	none	Post > Pre	none	none	Post>Pre

Regression ANOVA's	adjusted R-squared						
Pre-Treatment	0.799	0.830	0.225	0.814	0.744	0.760	0.764
Treatment	0.903	0.955	0.670	0.922	0.928	0.955	0.856
	F-value						
Pre-Treatment	147.7	181.1	10.9	163.1	108.7	118.2	111.3
Treatment	306.8	708.4	68.0	390.5	427.7	694.6	196.4
P-Value (slope)	< 0.0001	< 0.0001	0.0024	< 0.0001	< 0.0001	< 0.0001	< 0.0001

Event TSS and TP mass were also evaluated with ANCOVA to determine if there was a treatment effect due to installation of the grassed strip in the west catchment. No significant differences were determined for natural log-transformed event mass of these two parameters ($p > 0.05$). However, there was a significant increase in the natural log of event runoff volume in the West catchment during the treatment period, relative to the relationship to the East catchment (ANCOVA: $p = 0.027$), as shown in Table 3, and Figure 16a. This difference is readily visualized in the runoff boxplots illustrated in Figure 14. Therefore, although there was a significant decrease in TSS and TP log-transformed event mean concentrations, there may not have been an accompanying decrease in loads because of the increase in runoff volume from the west catchment during the treatment period (relative to the east catchment). In addition, there was a significant increase (ANCOVA: $p = 0.009$) in the log-transformed event mass of DP at the west catchment during the treatment period (relative to the east catchment), despite no significant change in DP concentration. These findings suggests that the detected difference in flow

during the treatment period was the likely cause for: a) not detecting a decrease in TSS and TP loads in the west catchment from the grassed strip treatment; and b) detecting a significant increase in DP loads in the west catchment during the treatment phase.

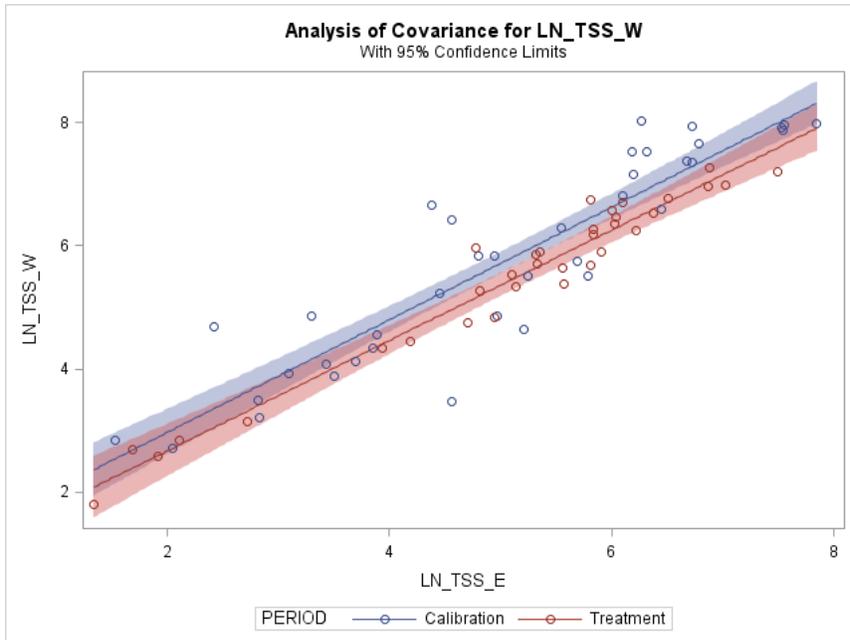


Figure 17. ANCOVA plot of the natural log of event-mean TSS concentrations (mg/L) during the calibration and treatment periods: east catchment (LN_TSS_E) versus west treatment catchment (LN_TSS_W), with 95% confidence limits.

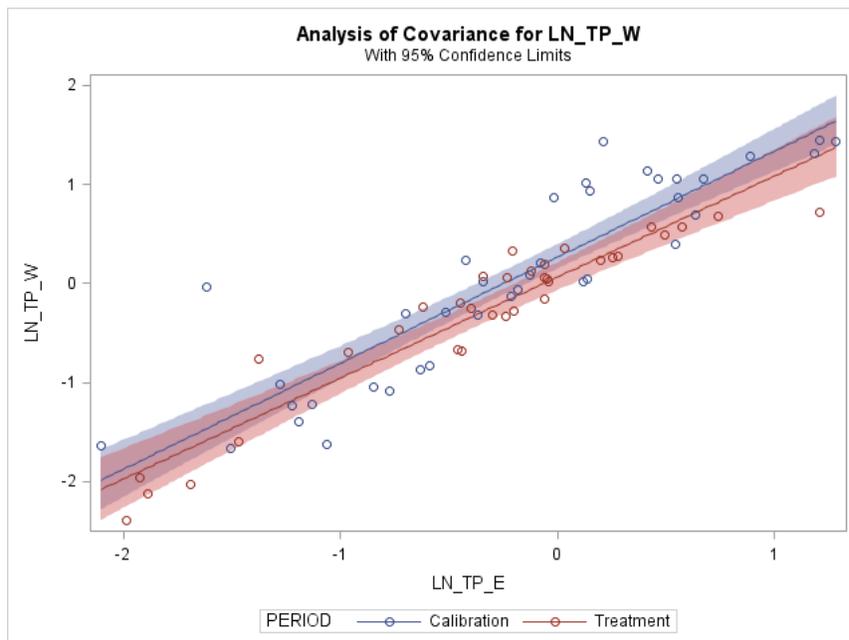


Figure 18. ANCOVA plot of the natural log of event-mean TP concentrations (mg/L) during the calibration and treatment periods: east catchment (LN_TP_E) versus west treatment catchment (LN_TP_W), with 95% confidence limits.

Non-parametric Wilcoxon rank-sum statistical analyses were also used to test for significant differences between pre- and post-treatment periods for TSS, TP and DP (Helsel and Hirsch 1992). In this test, the concentration of the East catchment constituent was subtracted from the West catchment constituent for each event, and then the Wilcoxon statistical analysis test was performed on these differences to compare the pre- and post-treatment periods. Results with the Wilcoxon test were comparable to the ANCOVA test results. Both TSS ($p = 0.0126$) and TP ($p = 0.0226$) event-mean concentrations were again significantly lower in the West catchment during the treatment period. These probability statistics are based on the hypothesis that we expected that these constituents would be reduced after treatment with the grassed strip. If no prior expectation had been anticipated, then the statistics would be doubled, but still significant at the 95% confidence level. In contrast, DP concentrations in the West catchment increased during the treatment period ($p = 0.0045$). Boxplots of the distribution of Wilcoxon scores for these constituents are shown in Figures 19 to 21.



Figure 19. Boxplots of Wilcoxon scores for TSS during calibration (pre-treatment) and post-treatment periods. A significant decline in TSS during treatment phase is indicated ($p < 0.05$).

The Wilcoxon non-parametric test indicates a significant increase in flow from the West catchment during the treatment period, relative to the relationship to the East catchment ($p < 0.001$), thereby confirming the ANCOVA results. This finding also implies that the detected increase in flow from the West catchment during the treatment period was a potential reason significant declines in TSS and TP loads were not detected from the west catchment under the grassed strip treatment.

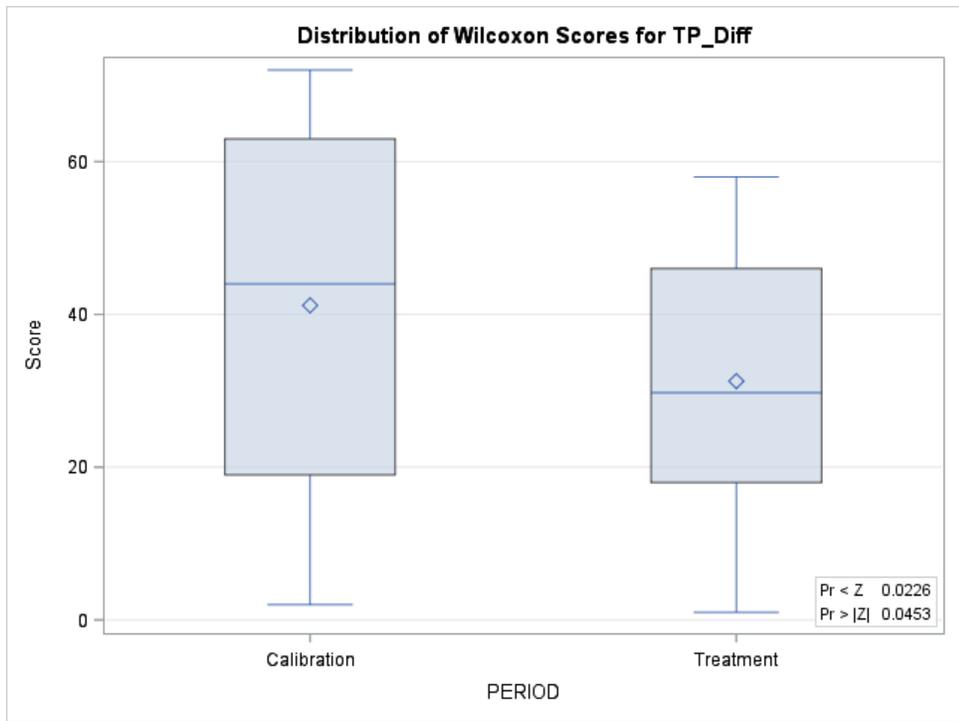


Figure 20. Boxplots of Wilcoxon scores for total phosphorus during calibration (pre-treatment) and post-treatment periods. A significant decline in TP during treatment phase is indicated ($p < 0.05$).

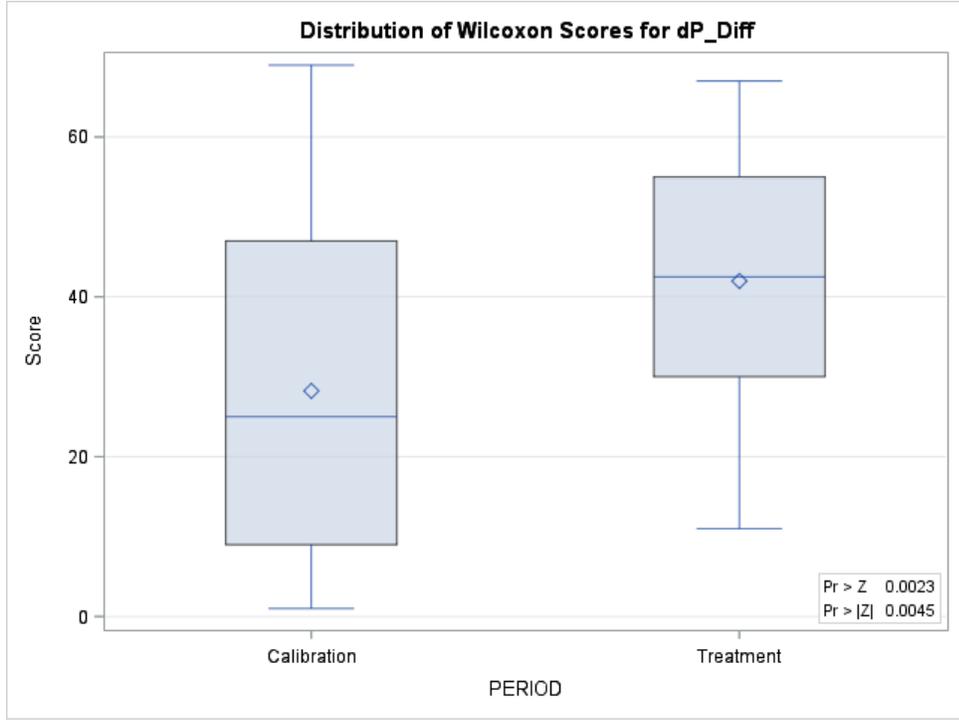


Figure 21. Boxplots of Wilcoxon scores for dissolved phosphorus during calibration (pre-treatment) and post-treatment periods.

An ANCOVA was conducted which included the natural log of event mean runoff volume as a covariate, to see if natural log-transformed event TSS and TP mass declined during the post-treatment period. Preliminary results indicate that there was a significant decline in TSS and TP mass in the West catchment during the post-treatment period ($p < 0.05$). However, these results have not been verified with a statistician to ensure the ANCOVA was correctly applied. A similar ANCOVA was also conducted, to see if natural log-transformed event DP mass changed during the post-treatment period, but with the natural log of event mean runoff volume as a covariate. Preliminary results indicate that there was not a significant decline in DP mass in the West catchment during the post-treatment period ($p > 0.05$); which is in contrast to Wilcoxon Rank Sum test and ANCOVA (without flow as a covariate) which both showed an increase in natural log-transformed DP event mean concentrations. These TSS, TP and DP results combined seem to indicate the mass of TSS and TP decreased, while the mass of DP did not change during the treatment phase. These preliminary results have not been verified with a statistician to ensure the ANCOVA was correctly applied.

Overall, the data illustrated in Figures 7 to 20 indicate that TSS and TP event-mean concentrations in the West plot were lower during the treatment period, relative to the pre-treatment period and the relationship to the control East plot. This apparent reduction in concentrations is likely due to the installation of the grassed strip in the concentrated flow channel.

Reduction estimate: The relationship between the East and West catchments during Control/Calibration period, was applied to East EMC's during the Treatment period to predict what the West values would have been without the grassed strip treatment (Clausen and Spooner 1993). The estimated effect of the grassed strip treatment on the West catchment was a 31% and 20% reduction in mean TSS and TP concentrations (mg/L), based on the difference between the mean predicted and observed concentrations listed in Table 4. These predictions are valid because the weight of evidence, including the ANCOVA and Wilcoxon statistical analyses, indicated a significant decline in EMC in the West catchment during the treatment period (relative to the East catchment). It is likely that the reductions would have been greater if the grassed strip had been better established during the treatment period.

Table 4. Summary of observed average TSS and TP event-mean concentrations (left) and the estimated reductions due to grassed strip treatment in the concentrated flow channel of the West catchment (right). West catchment concentrations listed on the right are predictions of what the concentrations would have been without the treatment, which are the basis for the estimated reductions.

	East control	West Treatment	West Predicted	West Reduction
mean TSS concentrations (mg/L)				
Calibration Period	448	888		
Treatment Period	358	452	655	31.0%
mean TP concentrations (mg/L)				
Calibration Period	1.11	1.59		
Treatment Period	0.90	0.95	1.19	20.1%

Turbidity Results: The relationships between backscatter in-situ turbidity (commercial probe) and concentrations of TSS (or SSC) and TP at the East station are illustrated in Figure 22. Statistical comparisons are summarized in Table 5 for the commercial and low-cost probes at both stations. Results from side-scatter turbidity are not included in the table because in-situ turbidity at the stations often exceeded the upper limit of the commercial probe (~1200 SSTU). In general, the commercial probe did not perform as well as desired for either TSS ($R^2 = 0.75$) or TP ($R^2 = 0.78$) at the East station, for unknown reasons. Plus, the low-cost probe seemed to perform better.

The relationships with constituent concentrations and turbidity were much better at the East station compared to the West station. This difference is likely due to incomplete mixing at the point where the West station turbidity probes were mounted. This explanation is supported by graphical analysis (not shown here), and the fact that correlations were much better at the east site than the west site for both the commercial and low-cost probes (Table 5). The turbidity probes were mounted on the opposite side of the inlet end of the flume (left), relative to the sample line intake location (near flume outlet, and right side). This mounting location was chosen to ensure that the probes did not interfere with stage height, and measurement. However, it is likely that water entering the West station flume was not fully mixed at that point compared to the sample intake line, during some events. Therefore, in-situ turbidity and concurrent water samples at the West station were not always representing/sampling the same water.

The relationship between measured and turbidity-estimated event-mean concentrations of SSC or TSS are shown in Figure 23 (commercial probe). The plotted data include 30 events from the pre-treatment period, during non-winter conditions at the East station only. However, an outlier from the 5/2/18 event was removed from the original data set after reviewing discrepancies, when it was found that backscatter turbidity was highly variable during this event, perhaps because of debris near the sensor. After removing this obvious outlier, R^2 improved from 0.71 to 0.78. There is a distinct overestimation of some turbidity-estimated EMC's at the lower end of the range of measured EMC's. This pattern is not apparent in the discrete data shown in Figure 22. Further review of the turbidity data set is required before a recommendation can be made about whether turbidity can be reliably used as an accurate surrogate for TSS (or SSC) at the study site. One potential reason for not having a better correspondence between observed and turbidity-estimated TSS (or SSC) is that suspended sediment grain size distribution varies during an event, and between events.

Table 5. Summary of coefficient of determination statistics (R-squared) between in-situ turbidity and concurrent concentrations of TSS (or SSC) and total phosphorus from discrete water samples collected at station outlets. BS is backscatter turbidity from a commercial probe. LCP stands for low-cost probe.

	--- OBS-500 Turbidity Probe ---				---- Low-Cost Turbidity Probe ---				TSS & SSC R-squared		Total phosphorus R-squared	
	start	End	# TSS	# TP	start	end	# TSS	# TP	BS	LCP	BS	LCP
East	6/14/2016	10/1/2020	101	72	4/3/2019	12/8/2019	86	59	0.75	0.90	0.78	0.91
West	6/10/2016	10/1/2020	129		4/3/2019	12/8/2019	113		0.57	0.67		

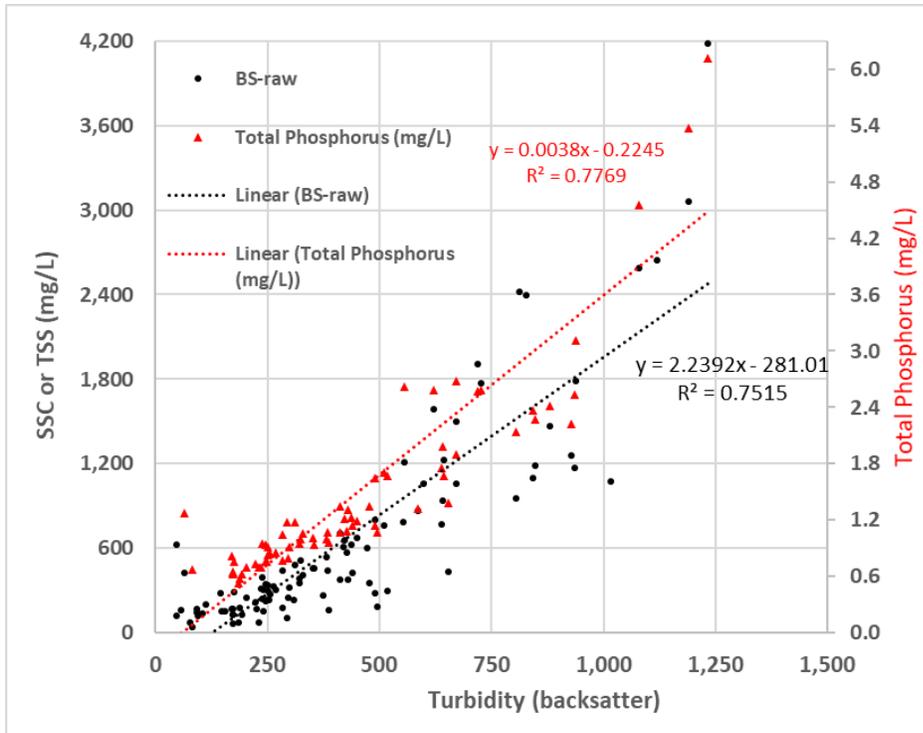


Figure 22. Turbidity versus measured Suspended Sediment Concentration (SSC, n= 86) or Total Suspended Solids (TSS, n=15), and total phosphorus (n=57): discrete samples from East site only.

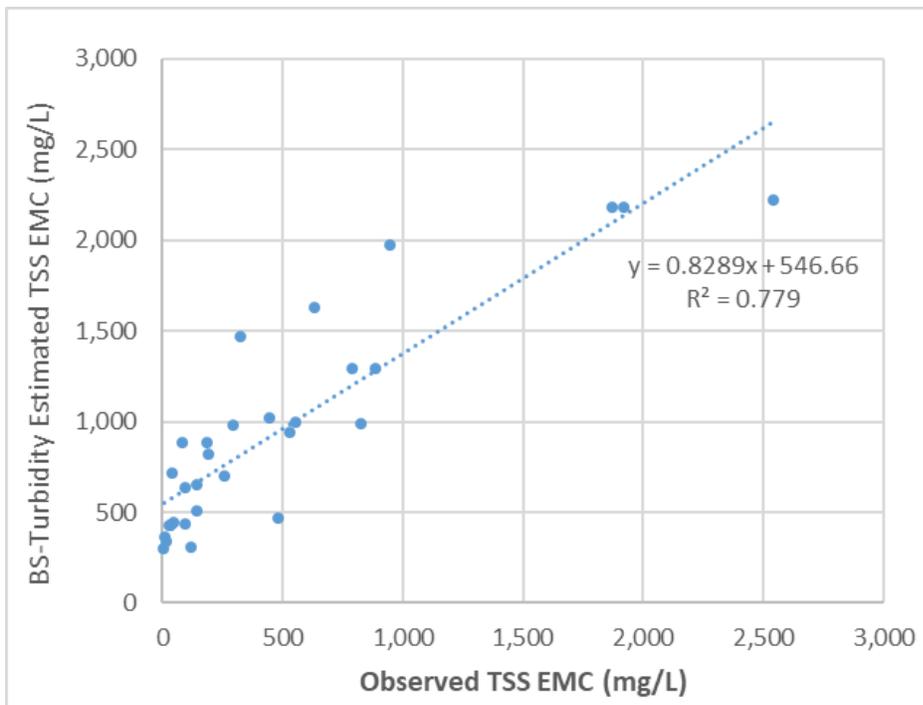


Figure 23. Observed versus backscatter turbidity-estimated event mean concentrations (EMC) of SSC or TSS during 31 pre-treatment non-winter events when turbidity probe was deployed: East site only.

CONCLUSIONS

The weight-of-evidence indicates that the treatment practice of installing a grassed strip in the west catchment was effective at reducing TSS and TP event-mean concentrations. Both ANCOVA and non-parametric Wilcoxon Ranked sum difference test showed significant declines in both parameters from the west catchment during the treatment period ($p < 0.05$). Five types of comparison plots also supported this conclusion. The estimated effect of the grassed strip treatment on the West catchment was a 31% and 20% reduction in event-mean TSS and TP concentrations (mg/L). Preliminary statistical analysis suggest that It is likely that there was a similar reduction in event-mean mass when runoff volume is included as a covariate in the analysis. An annual record amount of precipitation was observed in 2018 at the National Weather Service station in Green Bay, Wisconsin. This record was super-ceded by 24% in 2019. The excessive precipitation that occurred in these two years contributed to saturated soil conditions and frequent rainfall and runoff events. These conditions helped increase the number of monitored events in the study, but also made it difficult to establish a fully protective grassed strip in the concentrated flow path of the west treatment catchment. Therefore, the observed reductions in TSS and TP event mean concentrations would likely have been greater if the grassed strip had been better established during the treatment period.

REFERENCES

- Arnold, J.G., Srinivasan, R., Muttiah, R.S. and Williams, J.R. 1998. Large area hydrologic modeling and assessment part I: model development 1, *Journal of the American Water Resources Association*, 34(1), 73-89.
- Clausen, J.C. and J. Spooner. 1993. Paired watershed study design. 841-F-93-009. Prepared for S. Dressing, Office of Water, U.S. Environmental Protection Agency, Washington, DC. EPA 841-F-93-009. 8 p.
- Grabow, G.L., J. Spooner, L.A. Lombardo, and D.E. Line. 1999. Detecting water quality changes before and after bmp implementation: use of SAS for statistical analysis. NWQEP Notes No. 93. NCSU Water Quality Group, Biological and Agricultural Engineering Department NC State University, Raleigh, NC.
- Kalk F. 2018. Evaluation of the APEX model to simulate runoff, sediment, and phosphorus loss from agricultural fields in northeast Wisconsin. MS Thesis, University of Wisconsin Green Bay.
- Laboski, C.A.M. and Peters, J.B. 2012. Nutrient application guidelines for field, vegetable, and fruit crops in Wisconsin. UWEX Publ. A2809. 88 p.
- QAPP 2015. Quality Assurance Project Plan (QAPP) for Targeting outcome-based sediment reduction in the lower fox watershed. Prepared by Fox-Wolf Watershed Alliance, Inc. for the United States Environmental Protection Agency, Great Lakes National Program Office. Grant No. GL00E01451.
- Schmitz A.R. 2020. Low-cost turbidity sensors as a method for watershed monitoring. MS Thesis, University of Wisconsin Green Bay.
- Stuntebeck, T.D., Komiskey, M.J., Owens, D.W., and Hall, D.W. 2008. Methods of data collection, sample processing, and data analysis for edge-of-field, stream gaging, subsurface-tile, and meteorological stations at Discovery Farms and Pioneer Farm in Wisconsin, 2001–7: U.S. Geological Survey Open-File Report 2008–1015, 51 p.
- Stuntebeck, T.D., Komiskey, M.J., Peppler, M.C., Owens, D.W., and Frame, D.R. 2011. Precipitation-runoff relations and water-quality characteristics at edge-of-field stations, Discovery Farms and Pioneer Farm, Wisconsin, 2003–8: U.S. Geological Survey Scientific Investigations Report 2011–5008, 46 p.

APPENDIX A. Full Event Data Summary

EVENT	Plum East Edge-of-Field												Plum West Edge-of-Field													
	rain mm	Runoff			Sediment			area = 2.57 ha			Dissolved P			Runoff			Sediment			area = 3.45 ha			Dissolved P			
		Liters	mm	mg/L	kg	kg/ha	mg/L	kg	kg/ha	mg/L	kg	kg/ha	Liters	mm	mg/L	kg	kg/ha	mg/L	kg	kg/ha	Liters	mm	mg/L	kg	kg/ha	mg/L
#1: 6-16-16	47.8	225,023	8.74	2,540	571.6	22.20	3.27	0.736	0.286	0.225	0.051	0.020	280,180	8.13	2,960	829.3	240.6	3.725	1.044	0.303	0.264	0.074	0.021			
#2: 9-22-16	47.5	66,655	2.59	122	8.1	3.2	0.924	0.062	0.024	0.449	0.030	0.012	230,745	6.69	344	79.4	23.0	1.224	0.282	0.082	0.408	0.094	0.027			
#3: 10-12-16	11.9	21,103	0.82	183	3.9	1.5	1.148	0.024	0.009	0.612	0.013	0.005	73,713	2.14	104	7.7	2.2	1.052	0.078	0.022	0.489	0.036	0.010			
#4: 10-26-16	33.3	235,721	9.15	632	149.0	57.9	1.888	0.445	0.173	0.527	0.124	0.048	453,749	13.16	736	334.0	96.9	1.987	0.902	0.262	0.502	0.228	0.066			
#5: 11-28-16	15.7	71,595	2.78	144	10.3	4.0	0.88	0.063	0.024	0.418	0.030	0.012	160,678	4.66	128	20.6	6.0	1.088	0.175	0.051	0.324	0.052	0.015			
#6: 11-29-16	18.8	289,419	11.24	96	27.8	10.8	0.597	0.173	0.067	0.174	0.050	0.020	402,222	11.67	32	12.9	3.7	0.745	0.300	0.087	0.195	0.078	0.023			
#7: 12-06-16	5.8	30,576	1.19	5	0.1	0.1	0.295	0.009	0.004	0.239	0.007	0.003	52,348	1.52	17	0.9	0.3	0.290	0.015	0.004	0.108	0.006	0.002			
#8: 2-18-17	melt	42,630	1.66	22	0.9	0.4	0.429	0.018	0.007	0.324	0.014	0.005	193,243	5.61	51	9.8	2.8	0.351	0.068	0.020	0.206	0.040	0.012			
#9: 2-21-17	5.6	71,068	2.76	31	2.2	0.9	0.46	0.033	0.013	0.337	0.024	0.009	161,274	4.68	59	9.5	2.8	0.353	0.057	0.017	0.065	0.010	0.003			
#10: 2-28-17	11.7	317,965	12.35	17	5.3	2.1	0.304	0.097	0.038	0.204	0.065	0.025	617,467	17.91	33	20.2	5.9	0.247	0.153	0.044	0.113	0.070	0.020			
#11: 3-06-17	melt	419,004	16.27	8	3.2	1.3	0.222	0.093	0.036	0.184	0.077	0.030	665,693	19.31	15	10.1	2.9	0.189	0.126	0.036	0.101	0.067	0.020			
#12: 3-07-17	4.1	67,331	2.61	49	3.3	1.3	0.53	0.036	0.014	0.305	0.021	0.008	158,977	4.61	95	15.1	4.4	0.420	0.067	0.019	0.130	0.021	0.006			
#13: 3-25-17	13.7	37,186	1.44	27	1.0	0.4	0.554	0.021	0.008	0.354	0.013	0.005	257,946	7.48	130	33.5	9.7	0.438	0.113	0.033	0.123	0.032	0.009			
#14: 3-27-17	13.7	221,203	8.59	33	7.3	2.8	0.278	0.061	0.024	0.188	0.042	0.016	617,921	17.92	49	30.3	8.8	0.359	0.222	0.064	0.145	0.090	0.026			
#15: 3-31-17	7.1	78,707	3.06	17	1.3	0.5	0.346	0.027	0.011	0.240	0.019	0.007	242,186	7.03	25	6.1	1.8	0.198	0.048	0.014	0.127	0.031	0.009			
#16: 4-04-17	15.0	166,326	6.46	47	7.8	3.0	0.322	0.054	0.021	0.159	0.026	0.010	475,594	13.80	76	36.1	10.5	0.294	0.140	0.041	0.142	0.068	0.020			
#17: 4-15-17	18.3	8,809	0.34	86	0.8	0.3	0.834	0.007	0.003	0.421	0.004	0.001	112,774	3.27	188	21.2	6.1	0.934	0.105	0.031	0.229	0.026	0.007			
#18: 4-16-17	10.2	93,859	3.65	190	17.8	6.9	0.807	0.076	0.029	0.231	0.022	0.008	239,908	6.96	246	59.0	17.1	0.883	0.212	0.061	0.142	0.034	0.010			
#19: 4-20-17	22.6	255,189	9.91	326	83.2	32.3	1.718	0.438	0.170	0.247	0.063	0.024	519,803	15.08	250	130.0	37.7	1.488	0.773	0.224	0.275	0.143	0.041			
#20: 4-27-17	23.9	152,381	5.92	40	6.1	2.4	0.693	0.106	0.041	0.266	0.041	0.016	403,359	11.70	62	25.0	7.3	0.725	0.292	0.085	0.240	0.097	0.028			
#21: 5-01-17	20.6	273,801	10.63	296	81.0	31.5	1.122	0.307	0.119	0.183	0.050	0.019	509,009	14.76	316	160.8	46.7	1.014	0.516	0.150	0.171	0.087	0.025			
#22: 6-03-17	27.9	19,130	0.74	1,887	36.1	14.0	3.344	0.064	0.025	0.527	0.010	0.004	55,540	1.61	2,660	147.5	42.8	4.275	0.237	0.069	0.399	0.022	0.006			
#23: 6-04-17	12.4	125,368	4.87	1,873	234.8	91.2	3.599	0.451	0.175	0.439	0.055	0.021	204,395	5.93	2,770	566.2	164.2	4.204	0.859	0.249	0.384	0.078	0.023			
#24: 6-15-17	20.6	35,076	1.36	828	29.0	11.3	1.744	0.061	0.024	0.302	0.011	0.004	87,800	2.55	1,580	138.7	40.2	2.370	0.208	0.060	0.249	0.022	0.006			
#25: 6-22-17	15.0	7,786	0.30	884	6.9	2.7	1.727	0.013	0.005	0.202	0.002	0.001	20,268	0.59	2,140	43.4	12.6	2.874	0.058	0.017	0.180	0.004	0.001			
#26: 6-22-17	11.2	167,368	6.50	1,920	321.3	124.8	2.423	0.406	0.157	0.185	0.031	0.012	208,962	6.06	2,850	595.5	172.7	3.599	0.752	0.218	0.189	0.039	0.011			
#27: 6-23-17	31.0	642,547	24.95	256	164.5	63.9	0.711	0.457	0.177	0.172	0.111	0.043	897,620	26.04	540	484.7	140.6	1.020	0.916	0.266	0.278	0.250	0.072			
#28: 6-29-17	16.5	37,011	1.44	792	29.3	11.4	1.954	0.072	0.028	0.204	0.008	0.003	55,249	1.60	1,610	89.0	25.8	2.889	0.160	0.046	0.209	0.012	0.003			
#29: 7-02-17	15.5	50,403	1.96	832	41.9	16.3	1.59	0.080	0.031	0.146	0.007	0.003	55,881	1.62	2,800	156.5	45.4	2.880	0.161	0.047	0.230	0.013	0.004			
#30: 7-16-17	36.6	298,164	11.58	448	133.6	51.9	0.984	0.293	0.114	0.305	0.091	0.035	361,040	10.47	916	330.7	95.9	2.388	0.862	0.250	1.200	0.433	0.126			
#31: 7-18-17	8.9	12,846	0.50	484	6.2	2.4	1.137	0.015	0.006	0.220	0.003	0.001	27,956	0.81	1,867	52.2	15.1	2.742	0.077	0.022	0.289	0.008	0.002			
#32: 2-25-18	10.2	87,764	3.41	490	43.0	16.7	1.51	0.133	0.051	0.160	0.014	0.005	433,029	12.56	1,300	562.9	163.3	3.130	1.355	0.393	0.863	0.374	0.108			
#33: 4-22-18	melt	301,380	11.70	11	3.4	1.3	0.122	0.037	0.014	0.108	0.033	0.013	538,680	15.63	108	58.2	16.9	0.194	0.105	0.030	0.095	0.051	0.015			
#34: 4-23-18	melt	201,834	7.84	96	19.3	7.5	0.198	0.040	0.016	0.104	0.021	0.008	751,691	21.80	610	458.5	133.0	0.960	0.722	0.209	0.228	0.171	0.050			
#35: 5-02-18	31.5	17,465	0.68	80	1.4	0.5	0.654	0.011	0.004	0.228	0.004	0.002	85,936	2.49	785	67.5	19.6	1.260	0.108	0.031	0.179	0.015	0.004			
#36: 5-04-18	51.6	783,386	30.42	555	434.8	168.9	1.16	0.909	0.353	0.123	0.096	0.037	1,818,445	52.75	1,850	3,364.1	975.8	2.550	4.637	1.345	0.498	0.906	0.263			
#37: 5-10-18	25.4	74,948	2.91	141	10.6	4.1	0.496	0.037	0.014	0.136	0.010	0.004	212,006	6.15	344	72.9	21.2	0.738	0.156	0.045	0.199	0.042	0.012			
#38: 6-18-18	76.7	289,339	11.24	530	153.3	59.6	1.23	0.356	0.138	0.411	0.119	0.046	723,001	20.97	3,090	2,234.1	648.0	4.210	3.044	0.883	0.591	0.427	0.124			
#39: 8-27-18	140.2	344,171	13.37	206	70.9	27.5	0.639	0.220	0.085	0.327	0.112	0.044	741,525	21.51	304	225.4	65.4	0.823	0.610	0.177	0.444	0.329	0.096			
#40: 9-04-18	75.2	497,315	19.31	212	105.4	40.9	0.792	0.394	0.153	0.328	0.163	0.063	1,061,793	30.80	368	390.7	113.3	1.060	1.126	0.326	0.423	0.449	0.130			
#41: 9-05-18	10.7	70,841	2.75	66	4.7	1.8	0.67	0.047	0.018	0.474	0.034	0.013	251,084	7.28	86	21.6	6.3	0.781	0.196	0.057	0.559	0.140	0.041			
#42: 10-04-18	10.7	21,904	0.85	340	7.4	2.9	1.29	0.028	0.011	0.518	0.011	0.004	81,122	2.35	482	39.1	11.3	1.300	0.105	0.031	0.428	0.035	0.010			
#43: 10-08-18	24.4	220,011	8.54	500	110.0	42.7	1.32	0.290	0.113	0.367	0.081	0.031	722,162	20.95	520	375.5	108.9	1.310	0.946	0.274	0.499	0.361	0.105			
#44: 10-10-18	23.4	146,699	5.70	204	29.9	11.6	0.711	0.104	0.041	0.317	0.046	0.018	470,372	13.64	350	164.6	47.8	1.070	0.503	0.146	0.493	0.232	0.067			
#45: 11-04-18	15.7	109,965	4.27	110	12.1	4.7	0.632	0.069	0.027	0.285	0.031	0.012	214,420	6.22	115	24.7	7.2	0.514	0.110	0.032	0.275	0.059	0.017			
#46: 11-06-18	15.7	182,030	7.07	332	60.4	23.5	0.942	0.171	0.067	0.206	0.037	0.015	507,528	14.72	296	150.2	43.6	0.850	0.431	0.125	0.167	0.085	0.025			
#47: 4-13-19	9.7	121,189	4.71	4	0.5	0.2	0.137	0.017	0.006	0.110	0.013	0.005	294,475	8.54	6	1.8	0.5	0.092	0.027	0.008	0.080	0.024	0.007			
#48: 4-18-19	12.4	79,404	3.08	5	0.4	0.2	0.151	0.012	0.005	0.118	0.009	0.004	228,229	6.62	15	3.3	1.0	0.119	0.027	0.008	0.076	0.017	0.005			
#49: 4-23-19	9.7	34,836	1.35	8	0.3	0.1	0.184	0.006	0.002	0.116	0.004	0.002	127,848	3.71	17	2.2	0.6	0.132	0.017	0.005	0.072	0.009	0.003			

APPENDIX B. Quality Control

Field blanks were collected through the ISCO sampling system chain by connecting one end of a silicone tube to the sample inlet in the flume, placing the other end into a four liter container of UWGB derived di-ionized or ultra-pure water, and then forcing a manual ISCO sample to be pumped into a standard 1 liter ISCO sampler wedge-shaped polyethylene bottle. This bottle was processed at the UWGB lab using the same techniques as done with the composite samples (i.e., cone-splitter, preservation for phosphorus, and filter for DP), before being delivered to the NEW Water lab for analysis. The collection dates and lab analysis results are summarized in Table B1. Results were reported as below the lab LOD, or nearly so (4.5 and 3.3 mg/L TSS vs 2.2 mg/L LOD for two samples), with one exception in 2018 at the west station (19 mg/L TSS). The East station 6/29/17 field blank was not analyzed by the NEW Water for an unknown reason. However, this oversight was deemed un-important given the low phosphorus field blank results. Trip blanks were not performed because the field blanks were run through the same processing method at the UWGB lab that would have been performed for the trip blanks.

Table B1. Field blank analytical results in mg/L (di-ionized or ultra-pure water pumped from flume sample inlet through ISCO sampler system to field bottle in sampler).

Sample ID	Date-Time	TSS	Phosphorus	Dissolved phosphorus
PF-W-QC-1	10/14/16 15:00	< 2.2	< 0.03	< 0.03
PF-E-QC-1	10/14/16 15:40	< 2.5	< 0.03	< 0.03
PF-W-QC-2	6/29/17 18:30	4.5	< 0.028	< 0.028
PF-E-QC-2	6/29/17 18:50	Missing	0.034	< 0.028
PF-W-QC-2018	6/4/18 11:15	18.9	< 0.057	< 0.057
PF-E-QC-2018	6/4/18 11:30	3.6	< 0.057	< 0.057
PF-E-QC-2019	8/4/19 18:30	< 2.2	< 0.023	< 0.023
PF-W-QC-2019	8/4/19 19:10	< 2	< 0.023	< 0.023
PF-E-QC-2020	7/20/20 10:36	< 6.7	0.028	0.028
PF-W-QC-2020	7/20/20 10:36	< 6.7	< 0.023	0.025

Duplicate analyses of samples were performed by analyzing some samples at UWGB lab for Suspended Sediment Concentration (SSC) and comparing the results to TSS results from the NEW Water lab. These results are summarized in Table B2. This process was also done to confirm lab results that at first may have seemed questionable because the west catchment sample was much higher than the east catchment. SSC analysis is performed by using the entire sample, or a sub-sample that is obtained through an accurate splitting process (e.g., cone-splitter), compared to TSS which is processed by using a pipette to retrieve an aliquot for analysis. Therefore, results from these two analyses are not exactly the same, in part, because it is sometimes difficult to siphon a representative sample with a pipette when the sample is composed of larger particles that quickly settle or won't easily be drawn through the pipette orifice. The different methods of obtaining a sample for analysis has less of an impact on TP, and very little on DP, so the primary means of validating the data was through comparisons of TSS and SSC.

The duplicate sample SSC confirmed that the original TSS results were acceptable in nearly all cases, particularly given that SSC is generally greater than TSS due to reasons previously stated. In addition,

large differences between the two catchments during the pre-treatment period were later attributed to the formation of a deep rill or ephemeral gully in the west catchment. However, the duplicate analyses caused some changes in the finalized dataset. For example, the 6/3/17 event #23 East station sample was retested at NEW Water for TSS after the SSC (1,925 mg/L) at UWGB lab was found to be substantially higher than the original NEW Water result (796 mg/L TSS). Therefore, the resulting NEW Water retest of 1,873 mg/L TSS was substituted for the original value. Nearly all retests or duplicate analysis of reserved samples at UWGB confirmed the original results from NEW Water. One exception occurred with event #52 (7/19/19) when the original TSS results of 2,280 mg/L and 804 mg/L from the east and west catchments, respectively, were replaced with the UWGB-analyzed values of 1,806 and 1,332 mg/L SSC, respectively. The only other exception occurred with event #69 (11/20/19) when the original NEW Water TSS analysis of 104 mg/L from the east catchment was replaced with the UWGB-analyzed values of 165 mg/L SSC values, because settled sediment in the reserved east and west bottles did not appear to be as different as the NEW Water TSS results: 104 vs 252 mg/L, respectively. The relatively similar concentrations of TP from the east (0.48 mg/L) and west (0.63 mg/L) catchments also seemed to support this change. Both substitutions did not affect the statistical analysis because there was still a significant decrease in TSS EMC (and substituted SSC) in the west catchment after the grassed strip was added, as indicated by ANCOVA. The first substitution slightly reduced the detected difference; whereas, the second substitution slightly enhanced the difference.

Table B2. Duplicate analytical results (mg/L). Lab analysis by NEW Water (NEWW) or UWGB.

Composite Sample ID	NEWW TSS	UWGB SSC	Difference	NEWW TSS retest	NEW Water TP	NEW Water TP retest	Difference
PF-E-123	796	1,925	141.9%	1,873	3.60		
PF-W-123	2,770	3,014	8.8%		4.20		
PF-E-124	828	854	3.2%		1.74		
PF-W-124	1,580	1,664	5.3%		2.37		
PF-E-128	792	961	21.4%		1.95		
PF-W-128	1,610	1,833	13.8%		2.89		
PF-E-125	884	851	-3.7%		1.73		
PF-W-125	2,140	2,031	-5.1%		2.87		
PF-E-129	832	949	14.1%		1.59		
PF-W-129	2,800	2,234	-20.2%		2.88		
PF-W-129	2,800	2,329	-16.8%		2.88		
PF-W-129	2,800	2,425	-13.4%		2.88		
PF-W-130	916	1,360	48.5%	927	2.39	2.45	2.8%
PF-E-130	448	460	2.6%	476	0.98	1.18	19.7%
PF-E-130	448	495	10.5%		0.98		
PF-E-152*	2,280	1,806	-20.8%		3.34		
PF-W-152*	804	1,333	65.8%		2.06		
PF-E-153	674	691	2.5%		1.64		
PF-W-153	880	860	-2.3%		1.63		
PF-E-154	979	970	-0.9%		1.54		
PF-W-154	1,450	1,513	4.3%		1.77		
PF-E-155	965	1,016	5.3%		1.77		
PF-W-155	1,060	1,491	40.7%		1.78		
PF-E-158	444	456	2.8%		0.94		
PF-W-158	810	864	6.6%		1.21		
PF-E-159	420	381	-9.2%		0.94		
PF-W-159	647	615	-4.9%		1.06		
PF-E-154	979	984	0.5%		1.54		
PF-W-154	1,450	1,500	3.4%		1.77		
PF-E-163	588	610	3.8%		1.22		
PF-W-163	692	722	4.4%		1.26		
PF-E-165	171	225	31.4%		0.82		
PF-W-165	210	265	26.0%		0.76		
PF-E-166	416	419	0.6%		0.96		
PF-W-166	576	602	4.5%		1.04		
PF-E-169*	104	165	58.3%		0.48		
PF-W-169	252	232	-7.9%		0.63		
PF-E-170	407	421	3.5%		1.03		
PF-W-170	720	701	-2.7%		1.42		
PF-E-169	104	170	63.0%		0.48		
PF-W-109					0.35	0.34	-4.8%

* utilized SSC from UWGB for study, instead of NEW Water TSS result