Region of the Great Bend of the Wabash River Water Quality Summary Report

1.0 Water Quality Assessment

1.1 Water Quality Sampling Methodologies

The Wabash River Enhancement Corporation (WREC), in partnership with Purdue University implemented a three-year baseline pre-implementation and a two-year post-implementation professional water quality monitoring program. Pre-implementation monitoring occurred from 1 April 2009 to 8 April 2012, while post-implementation monitoring occurred from 14 February 2013 through 13 February 2015. The final Quality Assurance Project Plan was approved by IDEM for pre-implementation monitoring on 20 April 2009; the post-implementation QAPP was approved by IDEM on 18 July 2012.

WREC and Purdue University conducted water chemistry sampling as part of a paired watershed monitoring program. Monitoring occurred at the three red and two yellow sites demarcated in Figure 1 and detailed in Table 1. Paired watershed sample sites were selected during the watershed management planning process (pre-implementation) based on land use and watershed drainage size. These same sites were used during post-implementation monitoring as well. The three tributary sites represent two pairs of test watersheds: one urban or urbanizing (Elliot Ditch) paired with the control watershed (Little Pine Creek) and one rural or agricultural (Little Wea Creek) paired with the control watershed (Little Pine Creek). The Wabash River upstream-downstream pair was designed to identify any observable impacts of Greater Lafayette on the Wabash River. Each of the five chemistry sample sites were sampled weekly to establish a baseline of water quality data during the pre-implementation period to which post-implementation water quality data could be compared in hopes of showing a measurable change in water quality. Weekly water chemistry samples were measured during both monitoring periods for nitrate-nitrogen, total phosphorus, total suspended solids, and E. coli. Stream flow gages were installed at the three tributary monitoring sites by the U.S. Geological Survey (USGS) in June 2009 and continued in operations during both monitoring periods. Data sondes were installed to measure dissolved oxygen, temperature, pH, conductivity, turbidity, and conductivity in August 2009 and continued in operation during both monitoring periods. Flow and sonde measurements occurred every fifteen minutes throughout the sampling period.

Table 1. Sample site locations and parameters sampled by site within the Region of the Great Bend of the Wabash River watershed.

Stream Site	Latitude, Longitude
Elliot Ditch Site: Old Romney Road	N40 22.25, W86 54.57
Little Pine Creek Site: CR 350 North	N40 28.03, W87 03.48
Little Wea Creek Site: CR 800 South	N40 18.08, W86 55.40
Wabash River: Davis Ferry Bridge at N. 9th Street	N40 28.52, W86 52.15
Wabash River: Granville Bridge at CR 700 West	N40 25.32, W86 53.82

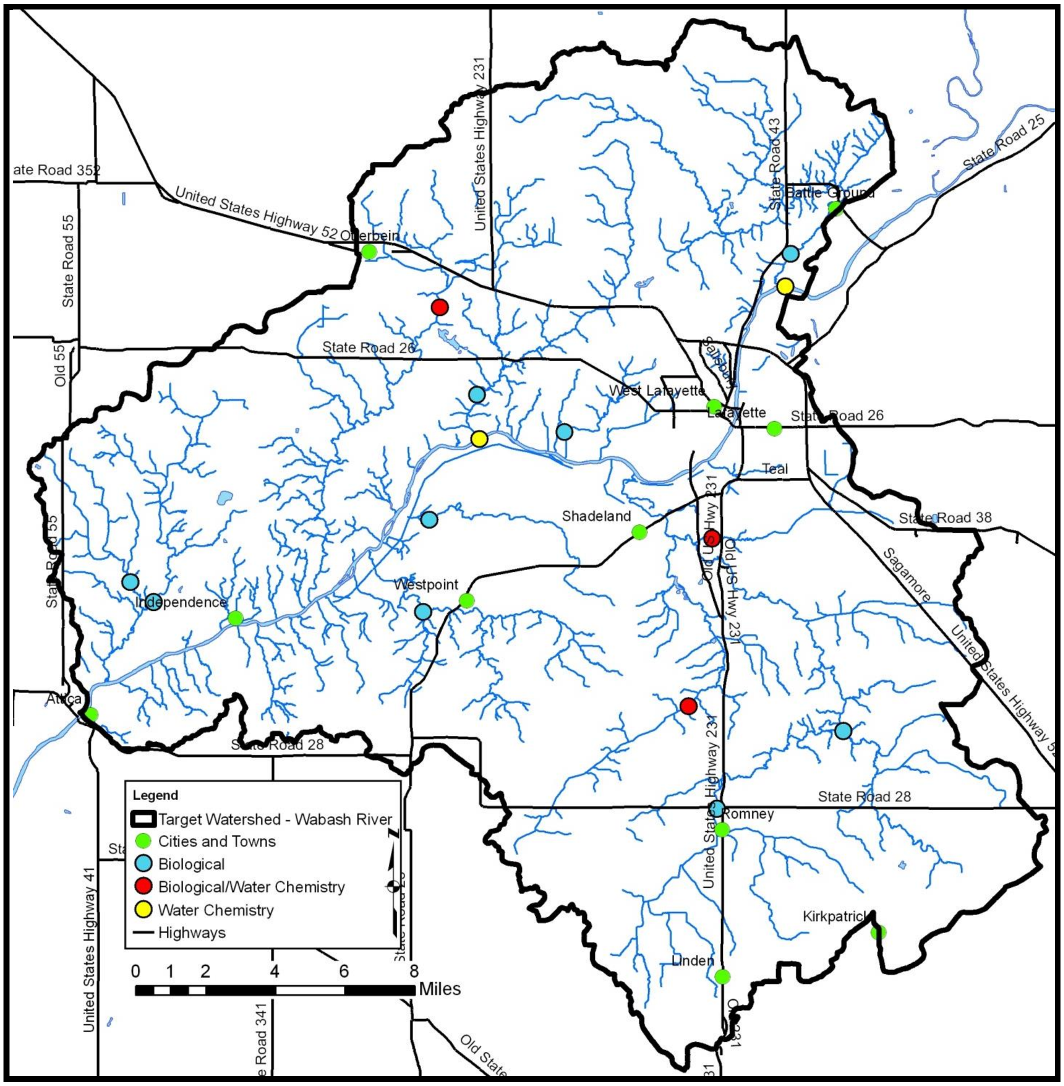


Figure 1. Sites sampled as part of the Region of the Great Bend of the Wabash River Watershed Management Plan.

1.2 <u>Field Chemistry Results</u>

Figure 2 through Figure 6 display results for field chemistry data collected every fifteen minutes at the three tributary sites. At each of the three stream sites, a multi parameter probe is deployed. The probe collects data for temperature, dissolved oxygen, specific conductivity, pH and turbidity at 15 minute intervals. Data shown below are an average of all the values in a given day.

Field data measured continuously at the three subwatershed sites generally measured within ranges identified by state water quality standards and water quality targets. Temperature measurements varied throughout the two sampling periods with the highest temperatures measured during summer months (July-September) and lowest temperatures measured from December through March (Figure 2). Temperatures measured approximately the same at all three sites during the sampling periods. All temperature measurements were below the state temperature standards, which vary monthly by season..

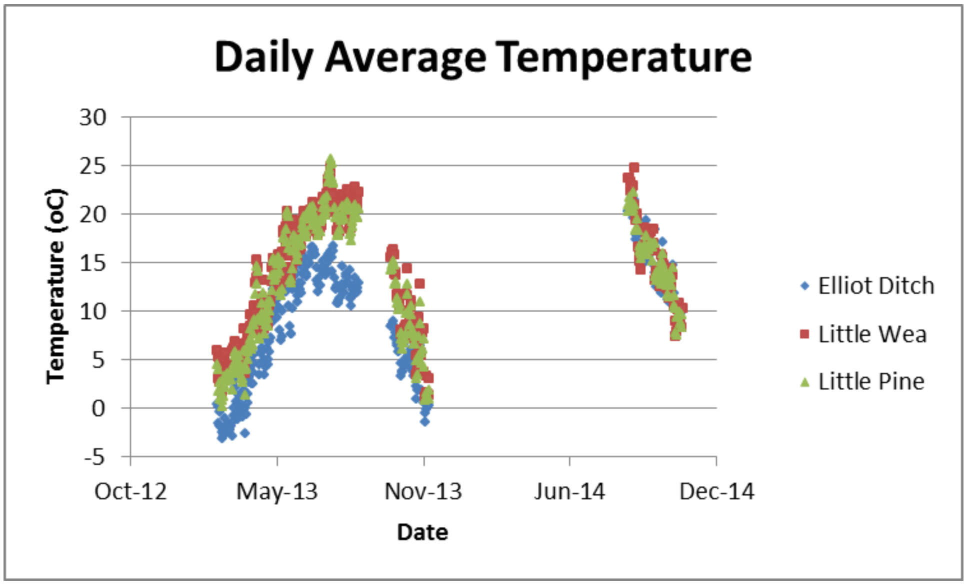


Figure 2. Temperature measured in Elliot Ditch, Little Pine Creek, and Little Wea Creek.

Like temperature, dissolved oxygen concentrations varied seasonally with the highest dissolved oxygen concentrations measured during summer months and the lowest dissolved oxygen concentrations measured during winter months (Figure 3). All three streams displayed diurnal variations in temperature and dissolved oxygen. Dissolved oxygen concentrations were routinely depleted during the winter months and fell to concentrations near 0 mg/L in the winter of 2010 in Little Pine Creek and in the winters of 2013 and 2014 at all three tributary streams. Low dissolved oxygen levels were observed in the winter when all three streams froze limiting oxygen entrainment. Relatively high dissolved oxygen concentrations were measured at all sites during the summers suggesting that algal productivity was elevated during the time period.

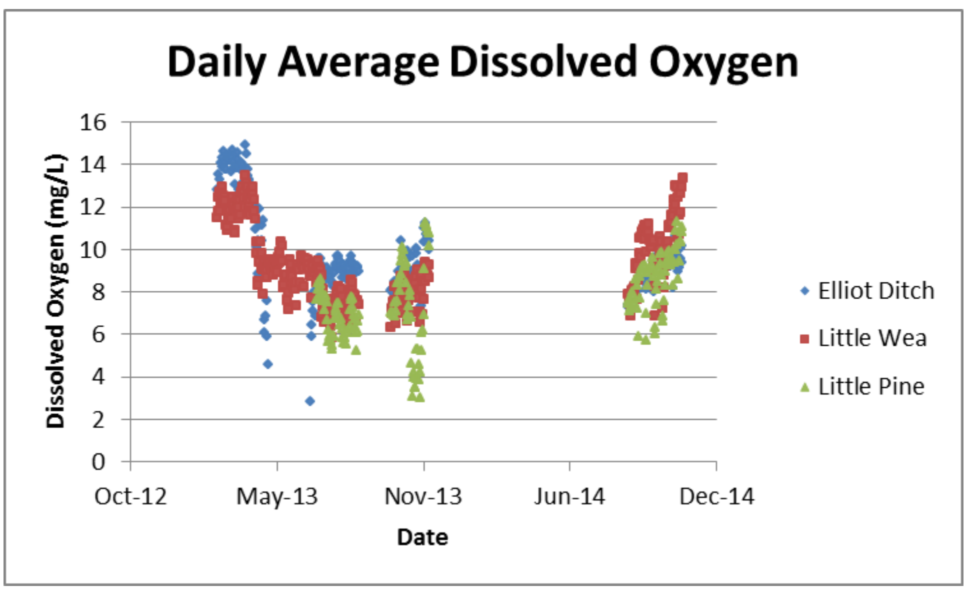


Figure 3. Dissolved oxygen measured in Elliot Ditch, Little Pine Creek, and Little Wea Creek.

Throughout the two sampling periods, pH remained in an acceptable range in all three streams (Figure 4). No discernable pattern can be found in pH levels in any of the three monitored streams. Although fluctuations appear to be wide, pH levels varied within 1 unit of pH. Generally, conductivity concentrations were below the state standard during both sampling periods (Figure 5). However, during December 2009 to March 2010 and again during the winter months of 2010 to 2011, conductivity increased in Elliot Ditch. In total, conductivity daily averages in Elliot Ditch exceeded the state standard (950 mS/cm) 12% of the time (59 of 467 days) during the pre-implementation monitoring period but only exceeded state standards 6% of the time (8 of 116 days) during the post-implementation period. The area around Elliot Ditch is urban and the increase in conductivity may be due to salts put down on the roads to melt snow and ice or result from industrial inputs during low flow stream conditions. The sustained high conductivity concentrations could be detrimental to biological communities present in Elliot Ditch.

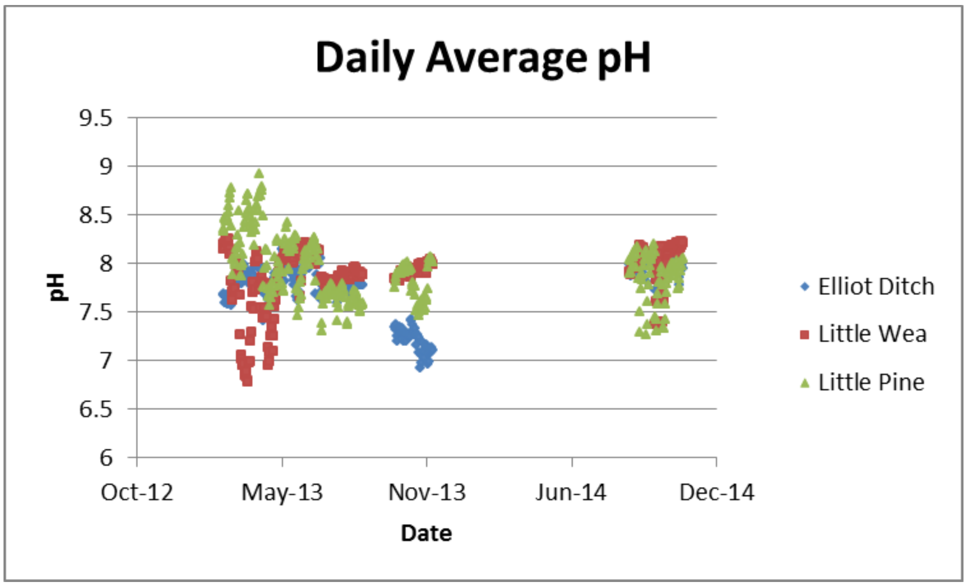


Figure 4. pH measured in Elliot Ditch, Little Pine Creek, and Little Wea Creek.

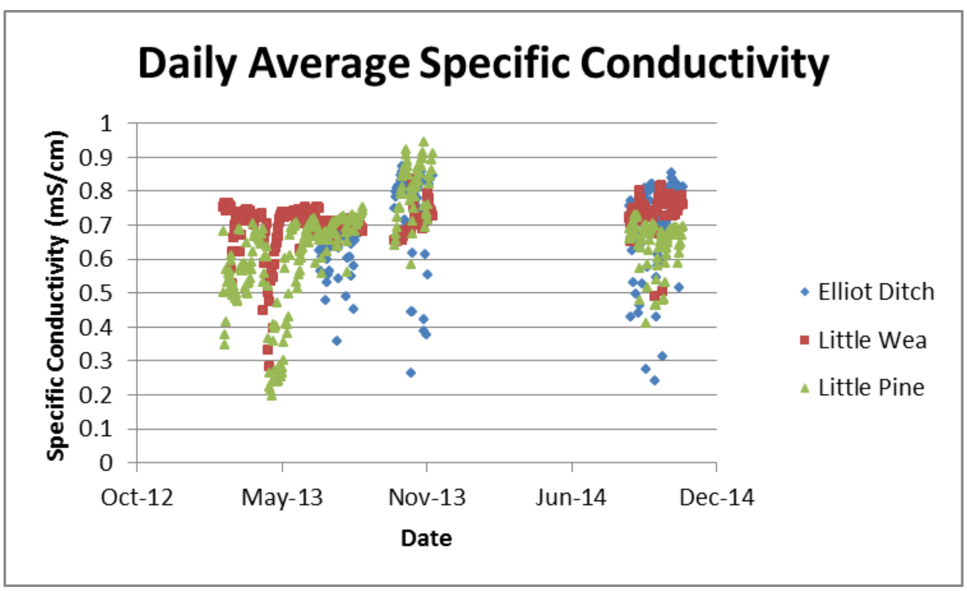


Figure 5. Conductivity measured in Elliot Ditch, Little Pine Creek, and Little Wea Creek.

Turbidity concentrations exceeded the 15 NTU target 35% of the time in Elliot Ditch during pre-implementation and 82% of the time during post-implementation, 23% of the time in Little Pine Creek during pre-implementation and 91% of the time during post-implementation, and 17% of the time in Little Wea Creek during pre-implementation and 58% of the time during post-implementation (Figure 6). Turbidity tends to spike during high flow events when elevated runoff washes particulates from streets and open agricultural fields into adjacent streams. Random turbidity peaks occurred throughout both sampling periods. It is unclear neither why the values in Elliot Ditch exceed the target nor why Little Pine Creek turbidity peaks are so high. It could be due to cows or wildlife accessing the stream, wildlife making a home in the guard surrounding the probes, or a malfunction. Herons have been observed near the probes in all three of the streams. When removing the sondes, fish, crawdads, and other macroinvertebrates have been pulled up with the instrument. All of these may cause spikes in turbidity.

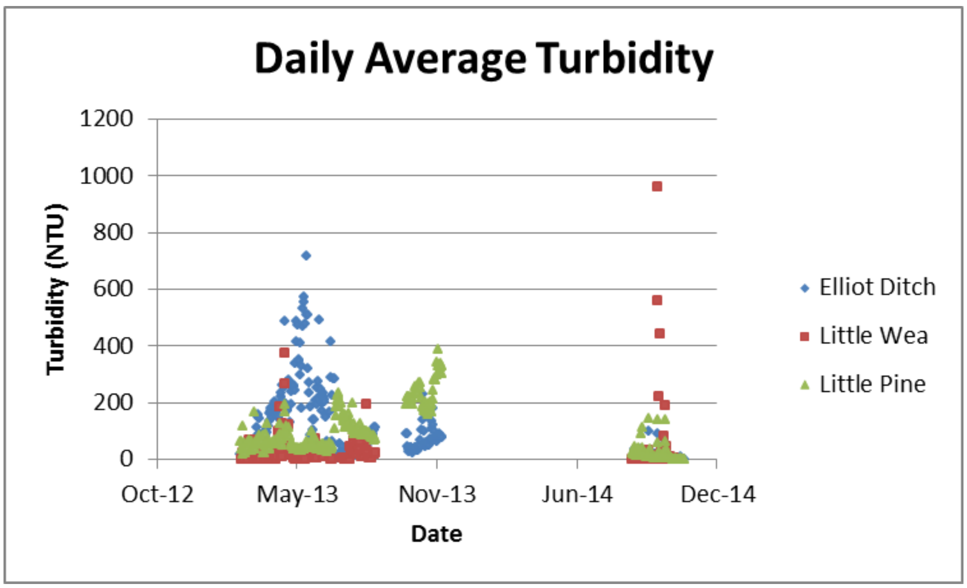


Figure 6. Turbidity measured in Elliot Ditch, Little Pine Creek, and Little Wea Creek.

1.3 Water Chemistry Results

Figure 7 through Figure 10 display results for nitrate-nitrogen (NO_3), total phosphorus (TP), total suspended solids (TSS), and E. coli collected weekly from five locations in the Region of the Great Bend of the Wabash River watershed. Data are displayed over stream discharge (blue) measured at USGS-maintained gaging stations during the sample period. Appendix I details individual measurements collected throughout the sampling period.

1.3.1 Nitrate-Nitrogen Concentrations

Nitrate levels varied greatly throughout the year but seem to follow a seasonal pattern with the lowest levels occurring during low flow periods during the late summer and fall months during both sampling periods (Figure 7). Concentrations measured in Elliot Ditch generally measure below the target concentration (2 mg/L) with nitrate-nitrogen concentrations exceeding the target in 8% of samples during the pre-implementation and 6% of samples during the post-implementation sampling periods. Changes in nitrate-nitrogen concentration in Little Pine Creek appeared to generally follow flow conditions with lower concentrations occurring during lower flow conditions; however, spikes in nitrate-nitrogen concentration did not always coincide with spikes in streamflow. In Little Pine Creek, nitrate-nitrogen concentrations exceeded target concentrations in 86% of samples during the pre-implementation and in 93% of samples during the post-implementation sampling period. In Little Pine Creek, seasonal variations in nitrate-nitrogen varied widely with concentrations annually from April to September exceeding both the target and the state standard for drinking water (10 mg/L). Nitrate-nitrogen concentrations measured during the sampling period mimic concentrations observed during historic water quality assessments within Little Pine Creek. This suggests that nitrate-nitrogen concentrations may be due to background conditions or that land use has changed little over time and that high volume application of manure or other fertilizers within Little Pine Creek may inflate nitrate-nitrogen concentrations within this watershed.

Similarly, nitrate-nitrogen concentrations in Little Wea Creek mimic flow conditions with higher concentrations typically occurring during periods of higher flow. Nitrate-nitrogen concentrations in Little Wea Creek do not measure as high as those measured in Little Pine Creek (15 mg/L); however, like Little Pine Creek, concentrations in Little Wea Creek more often measured above the target concentration during the pre-implementation monitoring period where concentrations exceeded targets in 83% of samples. During post-implementation monitoring, only 64% of samples exceeded the target concentration. Nitrate-nitrogen concentrations in the Wabash River mimic flow conditions with lower nitrate concentrations typically occurring during low flow conditions.

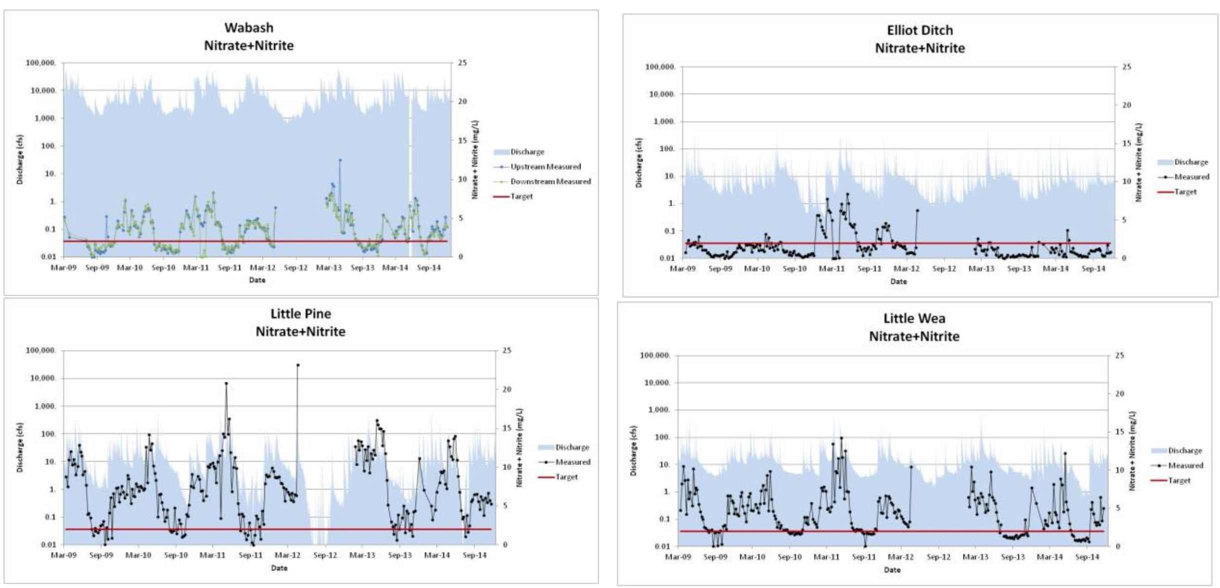


Figure 7. Nitrate-nitrogen concentrations overlain on discharge in the Wabash River (upstream and downstream), Elliot Ditch, Little Pine Creek, and Little Wea Creek. The red line indicates the target concentration (2 mg/L).

1.3.2 Total Phosphorus Concentrations

Total phosphorus concentrations do not follow a seasonal pattern (Figure 8). In Elliot Ditch, total phosphorus concentrations exceed target concentrations in 23% of samples during the pre-implementation and in 41% of samples during the post-implementation sampling period. Concentrations spiked during high flow events in both sampling periods. In Little Pine Creek, total phosphorus concentrations do not follow flow patterns. In Little Pine Creek, total phosphorus concentrations exceeded target concentrations 55% of samples during the pre-implementation and in 54% of samples during the post-implementation sampling period. In general, total phosphorus concentration in Little Wea Creek and the Wabash River measured lower than those observed in Elliot

Ditch and Little Pine Creek. In Little Wea Creek, total phosphorus concentrations exceeded target concentrations in 13% of samples during the pre-implementation and in 14% of post-implementation samples. Peaks in total phosphorus concentration generally do not coincide with peak flow conditions within Little Wea Creek. As in Little Wea Creek, peak total phosphorus concentrations measured in the Wabash River do not coincide with peak flow conditions. In the Wabash River, total phosphorus concentrations exceeded target concentrations for 34% of samples upstream of Greater Lafayette and for 73% of samples downstream of Greater Lafayette during the pre-implementation sampling period; in 54% of upstream samples and 80% of downstream samples during the post-implementation sampling period. Concentrations measured downstream of Greater Lafayette typically exceeded concentrations measured upstream of Greater Lafayette with average concentrations measuring 0.1 mg/L and 0.06 mg/L, respectively. This suggests that Greater Lafayette contributed total phosphorus to the Wabash River with larger contributions occurring under high water conditions.

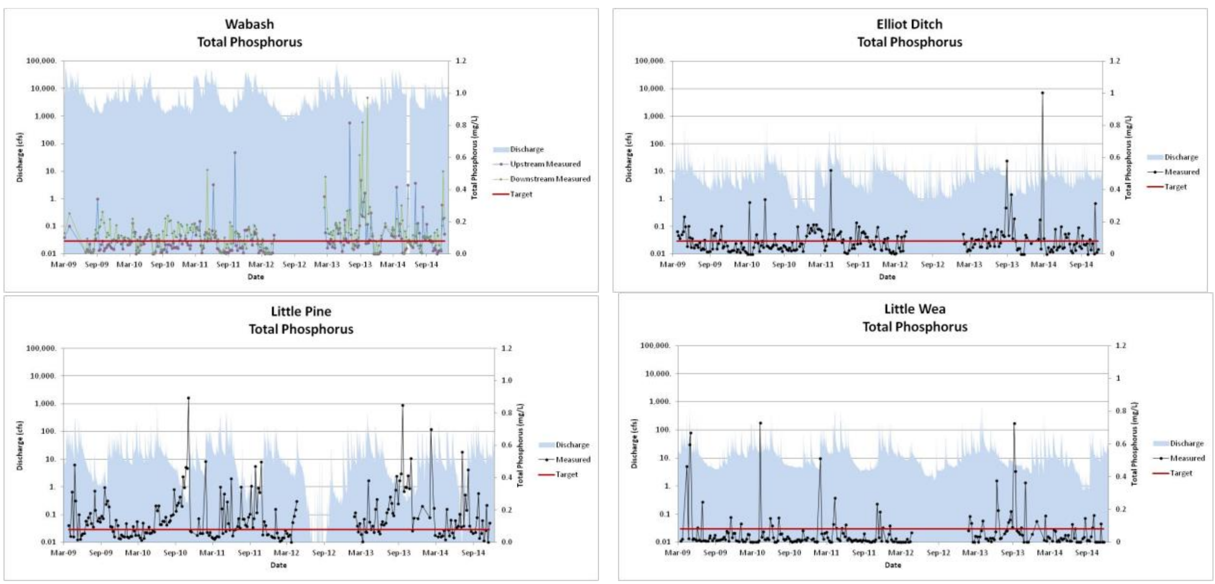


Figure 8. Total phosphorus concentrations overlain on discharge in the Wabash River (upstream and downstream), Elliot Ditch, Little Pine Creek, and Little Wea Creek. The red line indicates the target concentration (0.08 mg/L).

1.3.3 Total Suspended Solids Concentrations

In general, total suspended solids increased during high flow events due to sediment runoff and soil erosion (Figure 9). In Elliot Ditch, total suspended solids concentrations typically measured below the target concentration (15 mg/L). Target concentrations were exceeded in 15% of samples during the pre-implementation sampling period and in 8% of samples during the post-implementation sampling period. Although peak TSS concentrations do not coincide with peak flows, TSS concentrations typically increased during higher flow conditions. This is to be expected as increases in TSS following storm events suggests that stormwater carries larger amounts of dissolved and suspended solids than are present during base flow conditions. Higher overland flow velocities typically resulted in an increase in sediment particles in runoff. Additionally, greater streambank and streambed erosion typically occurred during high flow. Therefore, higher TSS concentrations are typically measured in storm flow samples. In Little Pine Creek, TSS concentrations exceed target concentrations 46% of samples during the pre-implementation sampling period and in 21% of sample collected during the post-implementation sample period.

Total suspended solids concentrations typically measured below the target concentration within Little Wea Creek. In Little Wea Creek, total suspended solids concentrations exceeded target concentrations for 20% of samples during the pre-implementation sampling period and in 21% of samples during the post-implementation sample period. Like Elliot Ditch, the TSS concentrations typically measured in Little Wea Creek suggest that targeting sediment moved during high flow or storm conditions will result in decreased TSS concentrations. In the Wabash River, total suspended solids concentrations typically mimicked flow conditions with the highest concentrations occurring from March to September. In the Wabash River, upstream and downstream total suspended solids concentrations exceed target concentrations for 54% and 60% of pre-implementation samples, respectively. During post-implementation sampling, TSS concentrations exceed targets in 54% and 60% of upstream and downstream samples, respectively. No obvious pattern is observable between upstream and downstream samples. This suggests that during some conditions Greater Lafayette contributes suspended sediments to the Wabash River but under different conditions, no contribution occurs.

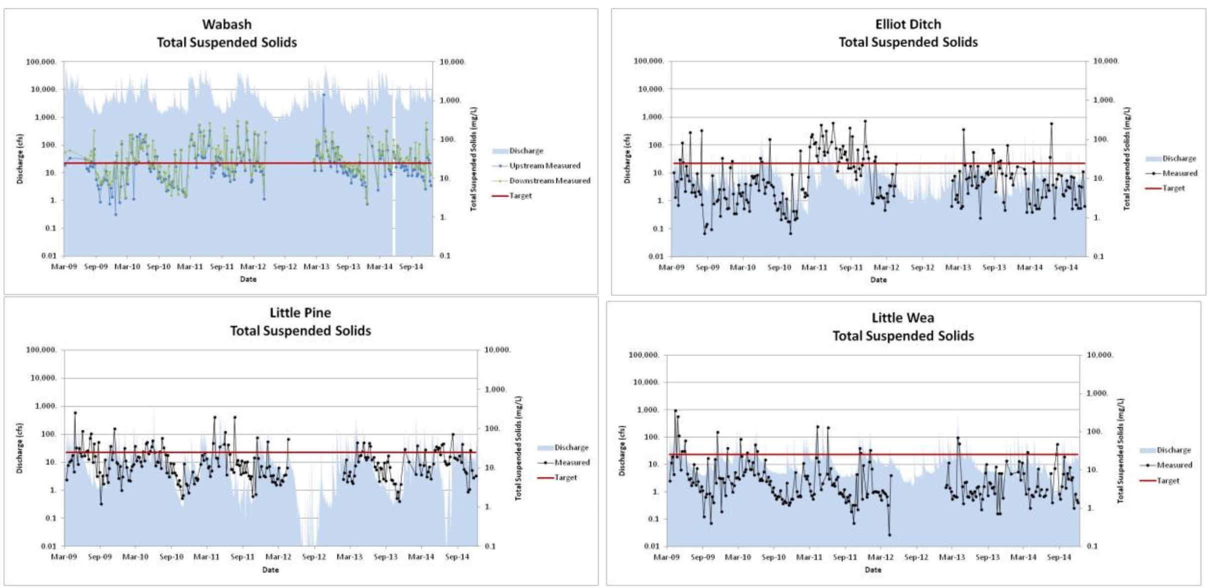


Figure 9. Total suspended solids concentrations overlain on discharge in the Wabash River (upstream and downstream), Elliot Ditch, Little Pine Creek, and Little Wea Creek. The red line indicates the target concentration (15 mg/L).

1.3.4 E. coli Concentrations

As shown in Figure 10, *E. coli* concentrations within Little Pine Creek, Little Wea Creek, and Elliot Ditch typically exceed the state standard (235 colonies/100 mL). In Elliot Ditch, *E. coli* concentrations exceeded target concentrations in 49% of samples during the pre-implementation sampling period and in 65% of samples collected during the post-implementation sampling period. *E. coli* concentrations mimicked flow conditions which suggest that *E. coli* concentrations increased when streamflows increased. This does not occur in Little Pine Creek where increases in *E. coli* concentrations do not occur during streamflow increases. Rather, *E. coli* concentrations are generally high under any condition. In Little Pine Creek, *E. coli* concentrations exceeded target concentrations for 92% of samples during the pre-implementation sampling period and in 83% of samples during the post-implementation sampling period. *E. coli* concentrations mimicked flow conditions within Little Wea Creek. When flows increase, *E. coli* concentrations typically increased. In Little Wea Creek, *E. coli* concentrations exceeded target concentrations 50% of samples during the pre-implementation sampling period and in 40% of sample collected during the post-implementation sampling period. Peak E. coli concentrations within all three tributaries measured approximately 10,000 CFU/L, suggesting that high *E. coli* concentrations are possible within these systems. *E. coli* concentrations measured in the Wabash River are typically lower than concentrations measured in the tributary streams. Differences in *E. coli* concentrations between tributary and mainstream sites can be attributed to a number of factors including dilution, lack of direct *E. coli* sources and inputs, or more periodic direct inputs of *E. coli* rather than continuous sources as observed in the tributaries. In the Wabash River, upstream and downstream *E. coli* concentrations exceed target concentrations in 33% and 49% of pre-implementation samples, respectively. During post-implementation sampling, E. coli concentrations exceed target concentrations in 18% and 30% of upstream and downstream samples, respectively.

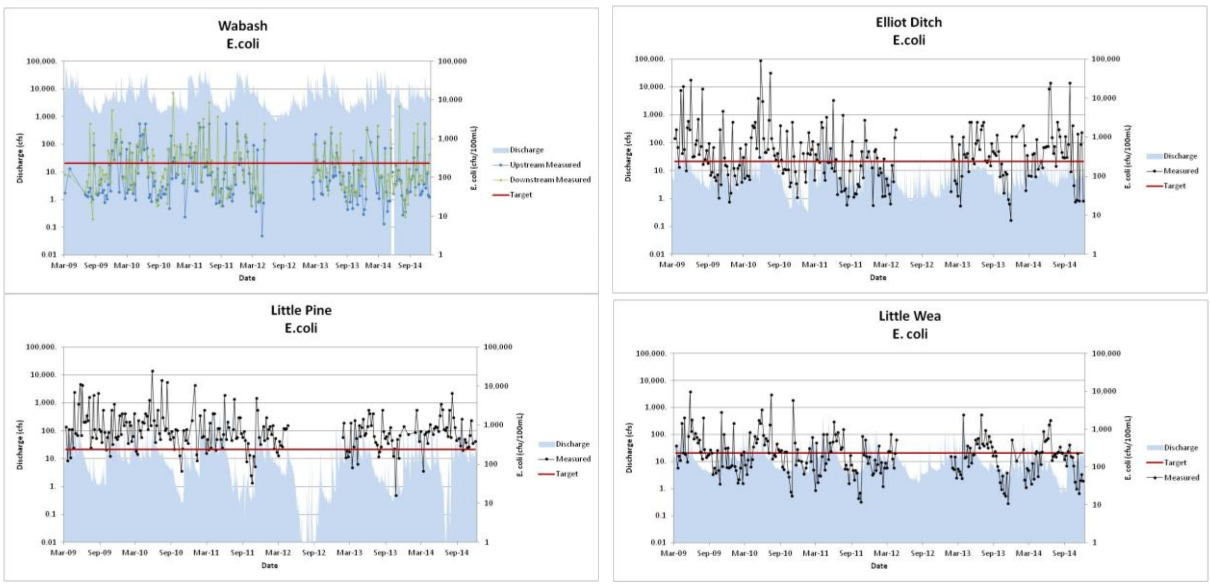


Figure 10. *E. coli* concentrations overlain on discharge in the Wabash River (upstream and downstream), Elliot Ditch, Little Pine Creek, and Little Wea Creek. The red line indicates the target concentration (235 colonies/100 ml) and state standard.

1.4 <u>Flow and Load Duration Curves</u>

Flow duration curves characterize the range of flow conditions present within a particular stream, while load duration curves allows for comparison of instream loading with streamflow so that conditions of concern can be identified.

1.4.1 Flow Duration Curves

Flow duration curves allow characterization of flow conditions within a particular stream. Instead of plotting individual flows as a time series, they are plotted as a percent of time that a given flow occurs within the stream. The resultant curve indicates the percent of time that a given flow is equaled or exceeded within the system. For instance, the median flow (Q_{50}) is the minimum flow observed in the stream 50% of the time. Flows less than the inflection point of the curve (near Q_{50} for these watersheds) indicate baseflow conditions within the stream. If this portion of the curve contains a steep slope, a relatively small contribution from natural storage sources like groundwater is suggested. Other indices can be used to characterize low flow conditions within the stream. The ratio of discharge observed 90% of the time compared to that observed 50% of the time (Q_{90}/ Q_{50}) is commonly used to determine the portion of flow that is contributed due to groundwater storage. Of additional importance is calculation of the percentage of time that zero-flow conditions occur.

Flow duration curves were developed based on the flow records for the four stream gages located within the Region of the Great Bend of the Wabash River watershed. For the Wabash River sites USGS Gage 03335500 was used, while Little Pine used USGS Gage 033356786, Little Wea used USGS Gage 03335673, and Elliot Ditch used USGS Gage 033356725. Flow duration curves developed using daily discharge values from each of the three tributary stream gages and the Wabash River gage are shown in Figure 11. The Wabash River contains much higher flows than any of the smaller streams. The Wabash River has a minimum discharge measuring 1000 cfs and this results in a much higher load value; load is a function of concentration multiplied by the volume. Trends in water flow were consistent across both the pre and post implementation sampling. Flow samples across both time periods showed that all three smaller streams display similar patterns with minimum discharges averaging about 1 cfs in each stream. Maximum flows in the three subwatershed streams occur infrequently (correlated with large storm events) and rarely exceed 1000 cfs.

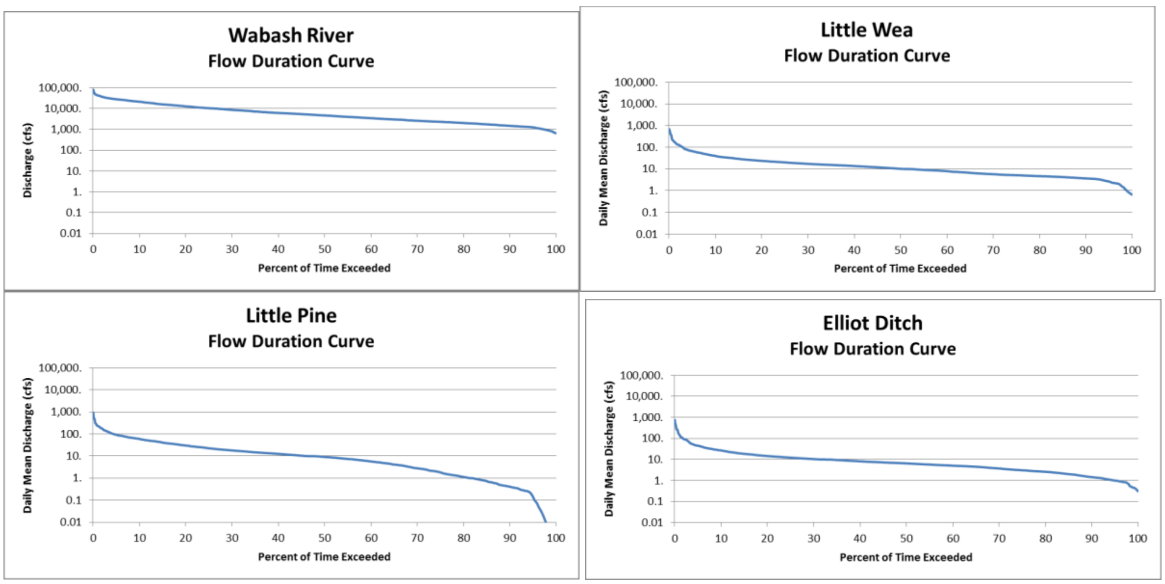


Figure 11. Flow duration curves developed for the Wabash River, Little Wea Creek, Little Pine Creek and Elliot Ditch.

1.4.2 Load Duration Curves

Load duration curves allow for comparison of observed instantaneous instream loading with target loads that area function of streamflow alone so that conditions of concern can be identified. Sample concentrations were plotted as points along these curves. Continuous target load duration curves for each material of concern were developed from the flow duration curves shown in Figure 11 using the following equation:

(observed flow (cfs)) x (conversation factor) x (target concentration or state criteria) = total load /day) (1)

Load duration curves suggest that high flow conditions deliver high volumes of nutrients, sediment, and pathogens to the receiving waters. Nitrate-nitrogen loads tend to exceed suggested target levels only during high flow events regardless of pre- or post-implementation (Figure 12). We see exceedance of the 2 mg/L level only about 50% of the time. This is especially apparent in Little Wea Creek (agricultural) and Little Pine Creek (control) where load patterns in the Little Pine Creek were unchanged post implementation. In all cases, the nitrogen loads associated with a 10 mg/L nitrate-nitrogen standard were only exceeded in Little Pine Creek at high flow levels under pre or post conditions. If we set a lower target standard (2 mg/L), post implementation samples in the Little Wea Creek exceeded the threshold less than 50% of the time. Using this approach, these data indicate a trend towards loads associated with the lower standard levels for the Little Wea Creek suggesting implementation of practices may be impacting water quality. In the Elliot Ditch (urban), nitrate-nitrogen rarely exceeds targets suggesting that this is not a parameter of concern within this subwatershed. Upstream and downstream the Wabash River curves display little difference pre or post, suggesting that while Greater Lafayette contributes low levels of nitrate-nitrogen to the Wabash River, implementation of practices did not affect the overall load in the river. This is not surprising in that Nitrogen loads in the Wabash River are an order of magnitude higher than in the smaller streams and most of the input occurs upstream and outside of the watershed.

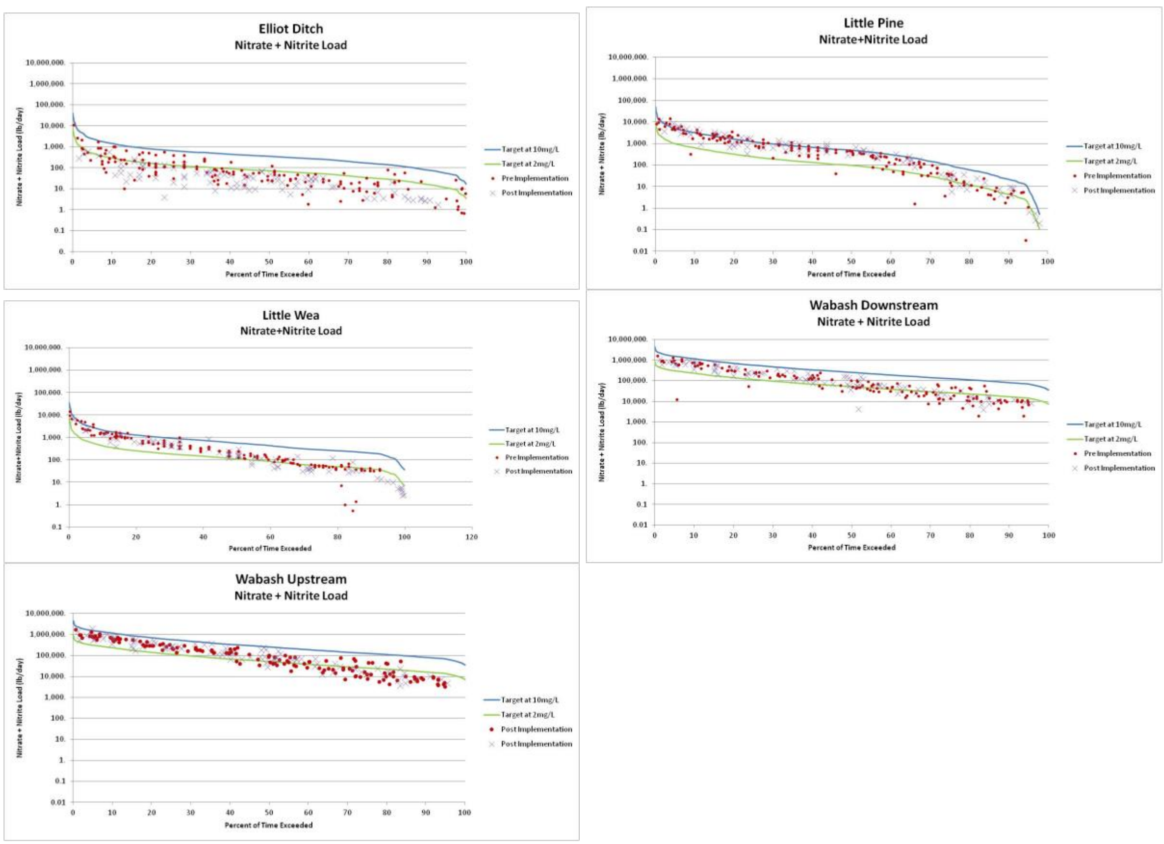


Figure 12. Nitrate-nitrogen load duration curves for the Wabash River (upstream and downstream), Little Wea Creek, Little Pine Creek and Elliot Ditch.

Total phosphorus and total suspended solids load duration curves indicate that loads are generally greater than targets during high flow conditions and lower than target under low flow conditions (Figure 13 and Figure 14). However, the data also indicates that phosphorus loads tend to exceed 1 kg per day in the Little Pine Creek more than 70% of the time. This suggests more phosphorus retention oriented practices should be implemented. Most of the suspended sediment enters the stream during periods of heavy overland flow and this is correlated with phosphorus inputs. While phosphorus can enter the stream attached to sediment particles, it also occurs in the dissolved fraction contributing to the higher loading during storm events. Wabash River phosphorus and sediment samples collected upstream of Greater Lafayette typically exceed target loads suggesting that although Greater

Lafayette contributes sediment and phosphorus to the Wabash River, other sources of these pollutants are also present outside of the watershed.

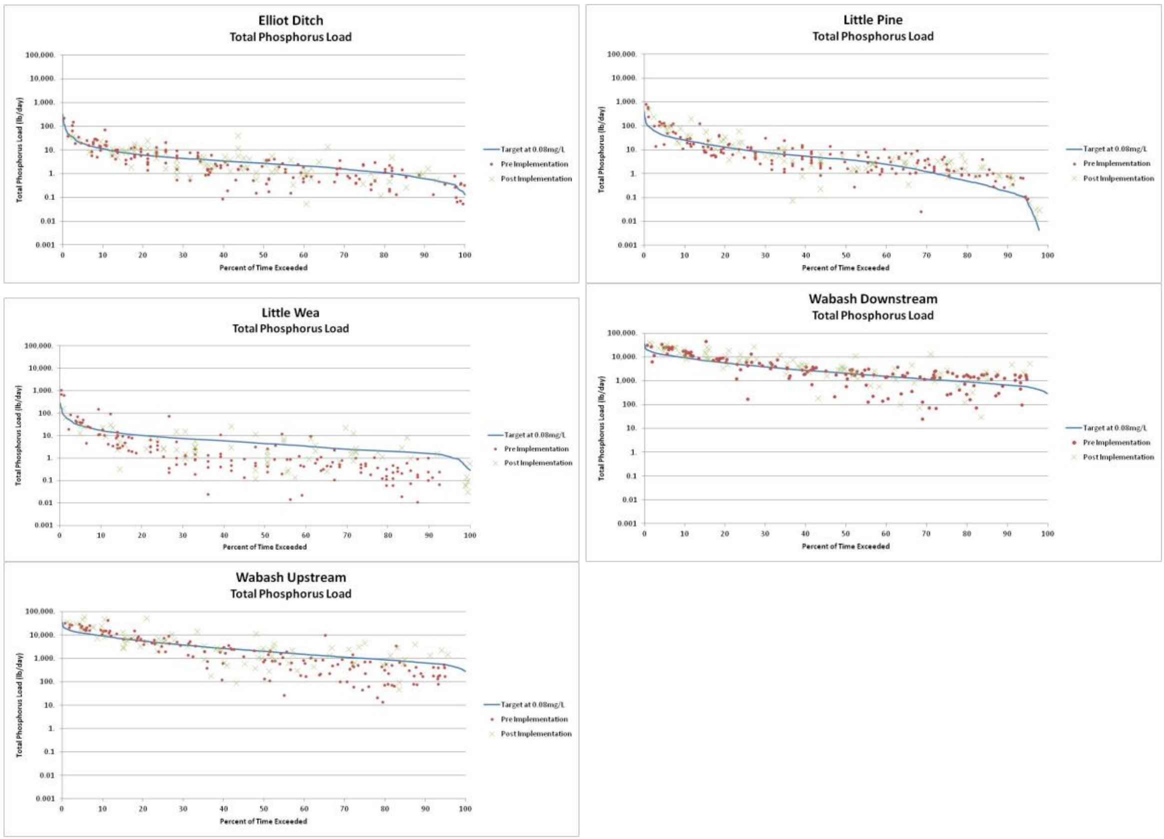


Figure 13. Total phosphorus load duration curves for the Wabash River (upstream and downstream), Little Wea Creek, Little Pine Creek and Elliot Ditch.

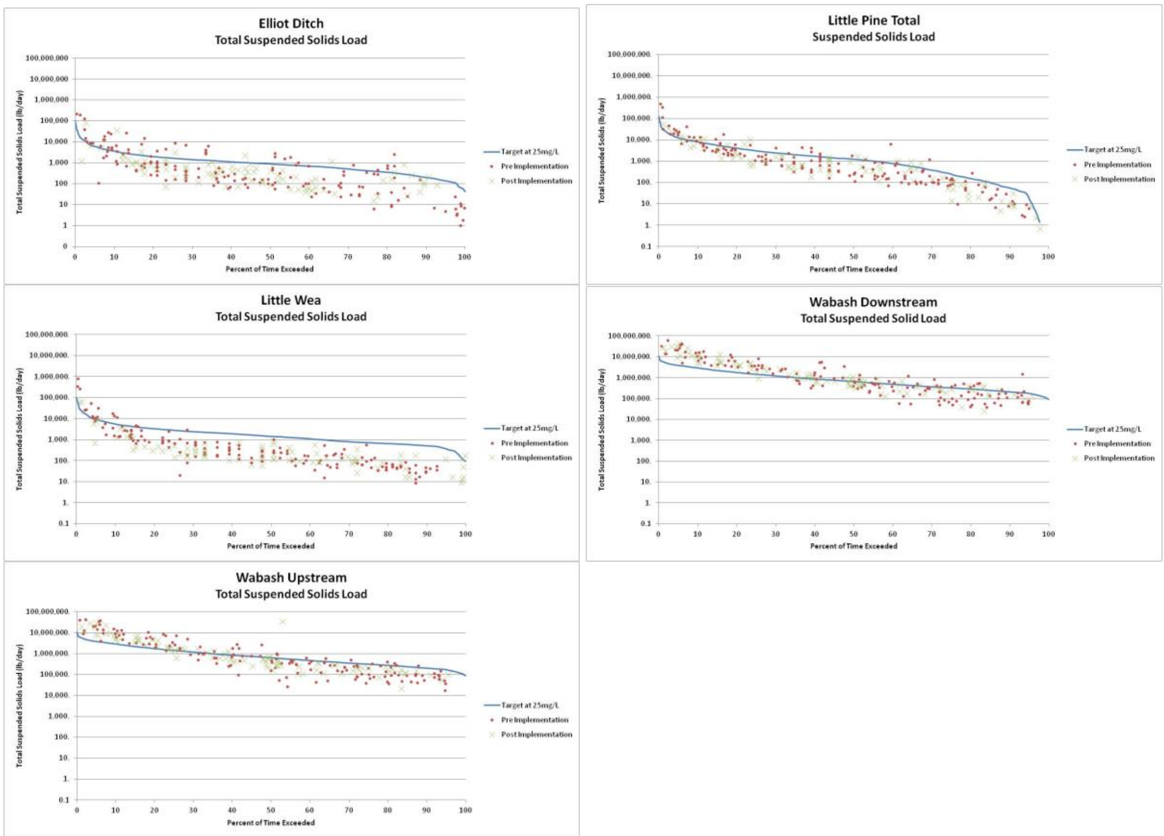


Figure 14. Total suspended solids load duration curves for the Wabash River (upstream and downstream), Little Wea Creek, Little Pine Creek and Elliot Ditch.

Pathogen load duration curves mimic those calculated for total suspended solids (Figure 15). *E. coli* in all three of the subwatershed sites with exceedances typically occurred under high flow conditions. In the Wabash River, *E. coli* loads occurred in excess of target loads in 28% of upstream and 44% of downstream samples. These data indicate that E. coli exceedances typically occur during storm conditions and also suggest that Greater Lafayette contributes *E. coli* to the Wabash River under moderate and high flow conditions. The data also suggest the implementation of practices in the watershed had little impact on the *E. coli*.

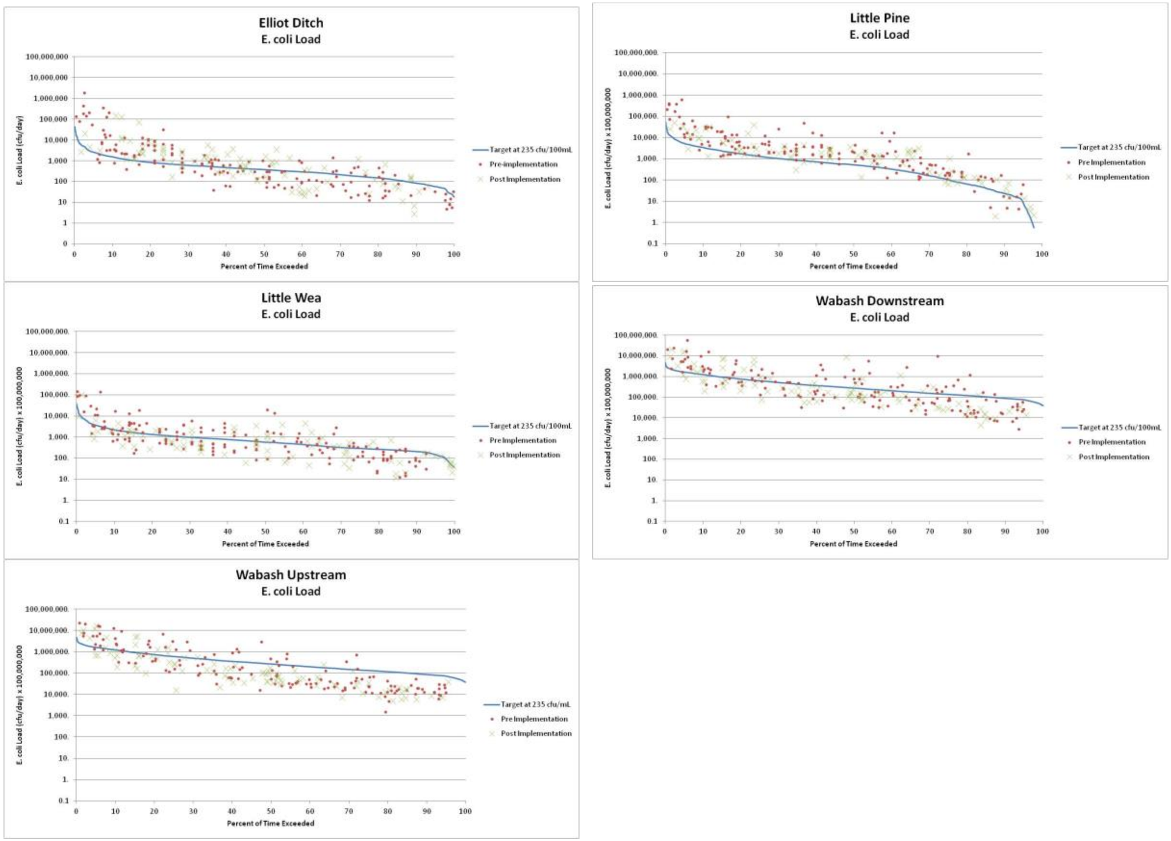


Figure 15. ***E. coli*** **load duration curves.**

2.0 <u>Statistical Analysis</u>

2.1 <u>Data Preparation</u>

Three different types of data were analyzed for the three sub-watersheds, including the weekly concentrations from grab samples, monthly load and storm discharge. Cleaning and pre-processing was conducted on all data prior to analyses. Not Available (NA) and -999 data were removed, and outliers were identified and removed. Data designated as -999 resulted from a software or hardware error during sampling, while missing data (categorized as NA) was the result of a failure to collect or analyze the samples. Outlier thresholds were visually identified by preparing histograms, time series, and box plots. To preserve the data for paired analysis, a missing value or outlier from one watershed resulted in the removal of that same point from all watersheds.

Key assumptions of the statistical tests performed include that the data is independently, identically and normally distributed (t-test), that the model residuals are independently, identically and normally distributed (ANCOVA) and that the data are independent (Mann-Kendall). Dependence in the data may refer to correlation between the pre-treatment and post-treatment datasets or serial (auto-correlation) of a time series with itself. This is a common occurrence in water quality time series because of natural memory in the system (the tendency for streamflow values today to reflect streamflow values yesterday due to the slow draining of soils or groundwater in the watershed.) If the sampling frequency is high enough, serial correlation is virtually certain to exist. The impact of such dependence is that sample variances will underestimate the true population variance and as a result, the null hypothesis is rejected too frequently (there is a "false positive" detection of change).

One method for addressing serial correlation is to create a new discrete time series that will not be correlated from a continuous dataset, such as extracting peak discharge values from the 15-minute discharge dataset, or analyzing monthly load rather than daily. When this is not a good alternative, a lag-one auto-regressive term (AR) was removed from the dataset prior to analysis to address the serial correlation. This is a method used to correct for dependence in time series data. The general equation for an AR process is given in equation 2:

$$x_t = \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \cdots + \varphi_p x_{t-p} + w_t \qquad (2)$$

where p is the order of the model, x is the data series and ϕ are constants, with w_t representing white noise with a mean of 0 and a variance of s_w^2. The order is determined by examining the autocorrelation and partial correlation functions of the data (Shumway and Stoffer 2011). If an observed time series was found to follow a first-order autoregressive process (AR1), a new time series calculated by removing this term will be independent:

$$y_t = x_t - \varphi_1 x_{t-1} = w_t \qquad (3)$$

Where ϕ1 is the lag-one correlation. Analysis then proceeds on the "pre-whitened" time series y_t after removal of this AR1 term.

In each case, the data were evaluated to see if the assumptions were met and appropriate adjustments were made if not. Any adjustments made are documented in Sections 2.1.1, 2.1.2 and 2.1.3. The results are summarized in traditional four plots for each transformed dataset; these are included in Appendix 2.

The exact same procedures used to treat the pre-implementation data for dependency and seasonality were imposed on the post-implementation data for consistency.

2.1.1 Concentration Data

Nitrate: The strong seasonal cycle in nitrate data creates dependency between observations. Nitrate was de-seasonalized by subtracting monthly means (calculated from May 01, 2009 to January 24, 2012) from the data, followed by the removal of a first-order AR term. No outliers were identified, but 7 missing or invalid data points were removed from the pre-treatment data and 9 missing or invalid data points from the post-treatment data.

Total Phosphorus: TP data underwent a log transform and the removal of an AR1 term to correct for non-normality and dependence evidenced in the autocorrelation plots. Outlier thresholds are listed in Table 2.

Table 2. Outlier thresholds for Total Phosphorus Concentration.

Pre- Implementation	Post-Implementation
LP > 0.6 mg/L	LW < 0.3 mg/L
ED > 0.4 mg/L	ED > 0.6 mg/L
Outliers removed: 2	Outliers removed: 7

Total Suspended Solids: A log transform followed by removal of an AR1 parameter was performed to adjust the data for non-normality and dependence in the data. Outliers are listed in

Table 3.

Table 3. Outlier thresholds for total suspended solids concentrations.

Pre- Implementation	Post-Implementation
LW > 300 mg/L	LW > 50 mg/L
LP > 200 mg/L	ED > 100 mg/L
	LP > 50 mg/L
Outliers removed: 7	Outliers removed: 2

E. coli: A log transform followed by removal of an AR1 parameter was performed to adjust the data for non-normality and dependence in the data. Outliers are listed in Table 4.

Table 4. Outlier thresholds for *E. Coli* concentrations.

Pre- Implementation	Post-Implementation
LP > 20000 mg/L	LP > 4000 mg/L
ED > 4000 mg/L	ED > 5000 mg/L
LW > 6000 mg/L	
Outliers removed: 9	Outliers removed: 2

2.1.2 Load Data

Since concentration sampling was done on a weekly basis, it is necessary to estimate the values on the days not sampled. Each parameter was first tested for correlation with observed discharge on the day of sampling. Overall correlation values were low, and so it was determined that there was little value in using a discharge versus concentration (Q versus C) relationship for estimating concentration on days without observations. Instead, a linear interpolation was used to calculate daily concentration. Once the daily concentration was calculated, a daily load was determined for each watershed in both the pre- and post-treatment periods. The daily load was summed to capture a monthly load total (in kg/month and No. of CFUs/month for E. coli) for analysis. As with concentration, a series of plots were created to examine the distribution and independence of each dataset, as summarized below. No outliers were removed from the monthly datasets.

Total Phosphorus: In the case of total phosphorus, a natural log transform was applied to both the pre- and post monthly load data.

E. coli: In the case of E. coli, a natural log transform was applied to both the pre- and post monthly load data.

Nitrate-nitrogen: Nitrate was de-seasonalized using two seasonal means. The first season was defined as March through August, while the second season was September through January. These seasons were determined through visual observation of the plots to identify high and low periods in load.

Total Suspended Solids: TSS was split into two seasons to de-seasonalize the data based on visual observations of high and low periods. Season one was defined as May through July, and season two covered August to April.

2.1.3 Discharge Data

In addition to the water quality datasets, two datasets were extracted from the 15 minute continuous record of discharge available from the gauging stations on Little Pine Creek, Little Wea Creek and Elliot

Ditch established through collaboration with the USGS in order to test of Best Management Practices have had a detectable influence on storm runoff.

Peaks-Over-Thresholds: The Peaks-over-threshold (POT) or Partial Duration series is a common method of hydrologic analysis that relies on extracting the peak observed discharge rate for all flow events that exceed an established discharge threshold (in cubic feet per second). For each watershed, a discharge threshold was established that resulted in approximately 10 – 15 storm events per year. This represents a relatively low threshold, so these storm events should not be considered large flood events, but include moderate storms as well. The extracted peaks were further screened to only retain peaks that were at least one day a part, so that multiple peaks from the same storm event were not included. After extracting the peaks from all watersheds, the datasets were compared between watersheds to only keep storm events that resulted in peak flows for all three watersheds within one day of each other. Prior to analysis, flow rates were converted to liters/s (L/s) a log transform was performed on the dataset to create a normal distribution.

Storm Volume: The on-line Web-based Hydrograph Analysis Tool (WHAT; https://engineering.purdue.edu/~what/) was used to separate storm flow and baseflow hydrographs for each watershed using a one parameter digital filter with filter parameter = 0.999. This parameter was selected to create distinct storm runoff events with a moderately responsive baseflow value. Total stormflow volume from multiple events was summed to produce a monthly dataset for each watershed of stormflow volume in liters/month. Prior to analysis, a log transform was performed on the dataset to create a normal distribution.

2.2 Statistical Tests

Three different statistical analyses were completed for each dataset. These include the two-sample t-test, analysis of covariance (ANCOVA) and the Mann-Kendall Seasonal test for trend. For each test, the magnitude of the difference was also calculated using the untransformed data. In addition, an analysis of the minimum detectable difference possible for the t-test and ANCOVA, given the observed variability in the data was also conducted. All tests used a significance level, $\alpha = 0.05$ and $\alpha = 0.10$.

2.2.1 Two Sample t Tests

The two sample t-test tests whether the mean of the pre-implementation data is different than the mean of the post-implementation data. An F-test was first conducted to determine if the pre and post datasets had equal variance. This was performed using the var.test function in the statistical computing environment R (R Development 2008).

The magnitude of the difference for the t-test is estimated as the difference in the means of the pre- and post-implementation datasets. Although the test was performed on transformed data, the difference in the mean of the untransformed data is reported for improved interpretation.

2.2.2 Analysis of Covariance (ANCOVA)

The Analysis of Covariance (ANCOVA) tests the effect of a categorical variable (pre- or post-treatment period) on the relationship between two continuous variables (control and treatment watershed response). The first step involves completion of two one-way ANOVA tests to evaluate if a significant relationship exists between the treatment and control watersheds. This uses the treatment and control watershed pairs for just the pre-implementation data, and then for just the post-implementation data. These analyses were done with the transformed datasets, but were repeated with the un-transformed

data in order to estimate a magnitude of change in real units. The built-in function aov was used to calculate both the ancova and anova results.
Table 5 details the linear model used in all ANOVA and ANCOVA calculations, respectively, where the t indicates the post-treatment data. The format is response variable ~ independent variable + interactions (optional).

Table 5. Linear model specifications for the ANOVA and ANCOVA analyses

<u>ANOVA</u>
Ecoli.lmu <-aov(ED ~ LP, data=Ecoli.df)
Ecoli.lma <-aov(LW ~ LP, data=Ecoli.df)
Ecoli.lmut <-aov(tED ~ tLP, data=Ecoli.df)
Ecoli.lmat <-aov(tLW ~ tLP, data=Ecoli.df)
<u>ANCOVA</u>
Ecoli.cov1<-aov(ED ~ LP + MGMT + LP:MGMT, data=ecolic.df)
Ecoli.cov2<-aov(LW ~ LP + MGMT + LP:MGMT, data=ecolic.df)

2.2.3 Seasonal Mann-Kendall Test

The two-sample t-test and ANCOVA both test for shifts in environmental data, that is, a step change in environmental response caused by an abrupt change in the system. Since the implementation phase has already continued over years, a test for trend, a continuous change over time, is more appropriate. However, it is generally recommended that tests for trend use at least ten years of data in order to have sufficient power to detect change. Given the seasonal nature of most environmental data, conducting trend analysis on data with a shorter time step (i.e. monthly) in order to increase sample size is problematic.

For this analysis, the Seasonal Mann-Kendall test was used to check for monotonic trend in seasonally varying data, while accounting for dependence between seasons, using the approach of Hirsch and Slack (1984). In this case months were used as seasons. This means that the trend test makes use of monthly data in order to account for differences in change at different times of the year, while calculating the significance of an overall annual trend slope. This is a non-parametric test, so the key assumption is that data are independent between months, therefore the untransformed data was used. For datasets with more than one observation per month (such as the weekly concentration) the multiple observations are treated as replicates and the median value is selected. Analysis was completed using the Matlab functions sktt.m developed by Jeff Burkey.

2.2.4 Minimum Detectable Difference

The minimum detectable difference (MDD) is a statement of the minimum difference in mean parameter value that could be detected for a given statistical test for a specified significance level and power. Power is the probability that a test will correctly reject the null hypothesis when the null hypothesis is false. There are several factors that affect the power of a statistical test. First there is a trade-off between power and the significance level, ⍰, such that power will increase with increasing probability of rejecting the null hypothesis when it is true (false positive). In other words, power increases when significance level, ⍰, increases. Power also increases with sample size and effect size (MDD) and is usually increased with the introduction of a covariate that controls for other sources of variation (paired watershed). A power analysis for a statistical test quantifies these trade-offs between sample size and effect size in order to quantify how many samples are needed to "detect" a given effect size (that is to correctly reject the null hypothesis of no change when it is in fact false).

Although the R environment has several built-in power functions, there is no built-in function to represent the power of an ANCOVA analysis, therefore Power Analysis was performed by simulation. This approach essentially uses a Monte Carlo analysis to generate 1000 datasets that are statistically similar to the observed data (normally distributed, same mean, variance and covariance between the treatment and control watersheds). Datasets of varying size are simulated for the post-treatment period, with the mean increased by different levels of MDD. The hypothesis test is then applied to each randomly-generated dataset, and the number of times that the null hypothesis is rejected is counted. Since the null hypothesis is known to be false (the post-treatment mean was intentionally increased by MDD in the simulation), the fraction of rejects represents the power of the test. Power was compiled across a range of samples sizes and MDD, and then the MDD and required sample size needed to attain a power of 0.8 was extracted via linear interpolation.

In order to check the new functions, the existing power function for the t-test in the R Environment was used to determine the effect size. The power function takes up to six arguments, with up to two arguments passed as NULL. This NULL parameter is then determined using the others. The function can solve for the power, sample size, or effect size. The method follows that of Jacob Cohen's power analysis (1988).

R returns the detectable effect size for the Student t-distribution, which can be interpreted as the critical difference in means (normalized) that we can detect within the data:

$$effect\ size = d = \frac{|\mu_1 - \mu_2|}{\sigma}; \qquad (4)$$

In order to compare with the results for the power.ttest() the effect size was multiplied by the standard deviation of the pre-treatment period to get MDD as a difference in means:

$$MDD = 100 * \sigma * d \qquad (5)$$

Using this method, an array of effect sizes was calculated based on 10 years of post-treatment sampling. We chose 0.8 as the power after reading several documents on statistical power. The untransformed data was used in this analysis so that the MDD would be calculated in the true units of the data (i.e. mg/L or CFU/ml) and represents a difference in the absolute value of the post-treatment mean relative to the pre-treatment mean. MDD analysis was not completed for the trend analysis.

2.3 <u>Results</u>

2.3.1 <u>Concentration Data</u>

<u>Time Series:</u> Time series plots with both pre- and post-treatment periods were created, and the mean of each period was calculated to examine initial increases or decreases in concentration. Figures 16-2 show the results; all watersheds saw a decrease in *E. coli* over time, while all watersheds increased concentrations for TP. For TSS, LP and LW decreased in concentrations, while ED increased. LW and ED decreased in concentration for NO_3, and LP increased.

<u>Two-sample t-test:</u> The results of the two-sample t-test on weekly concentrations are summarized in Table 6. The pre/post variance for TP in Little Pine and Little Wea were significantly different, as was NO_3 in Elliot Ditch and Little Wea for the F-test. All other results had equal variance; however, this only impacted the var.equal call in the t-test function.

Statistically significant differences between the pre- and post-treatment means were shown across all watersheds for nitrate-nitrogen (Figure 16). Little Pine Creek showed an increase in nitrate-nitrogen concentrations over time, while Elliot Ditch and Little Wea Creek showed decreases in nitrate-nitrogen. TSS concentrations increased significantly in Elliot Ditch (Figure 17), and total phosphorus concentrations increased in Little Wea Creek (Figure 18). There were no statistically significant changes in *E. coli* (Figure 19).

Table 6. Two sample T-Test concentration results.

Watershed and Pollutant	p-value	Test statistic	Difference in Mean[1] untransformed
NO3			
Little Pine**	0.03	-2.16	0.98 mg/l
Elliot Ditch**	0.00	3.49	-0.28 mg/l
Little Wea**	0.00	8.41	-0.67 mg/l
Total Suspended Solids			
Little Pine	0.85	0.20	-1.92 mg/l
Elliot Ditch**	0.01	-2.83	3.69 mg/l
Little Wea	0.17	1.37	-4.97 mg/l
Total Phosphorus			
Little Pine	0.57	-0.58	0.01 mg/l
Elliot Ditch	0.55	-0.60	0.02 mg/l
Little Wea**	0.00	-3.90	0.01 mg/l
E. coli			
Little Pine	0.28	1.08	-240.67 CFU/l
Elliot Ditch	0.97	0.04	-366.12 CFU/l
Little Wea	0.12	1.56	-90.46 CFU/l

[1]All differences are presented as post-treatment – pre-treatment

**Significant difference at = 0.05

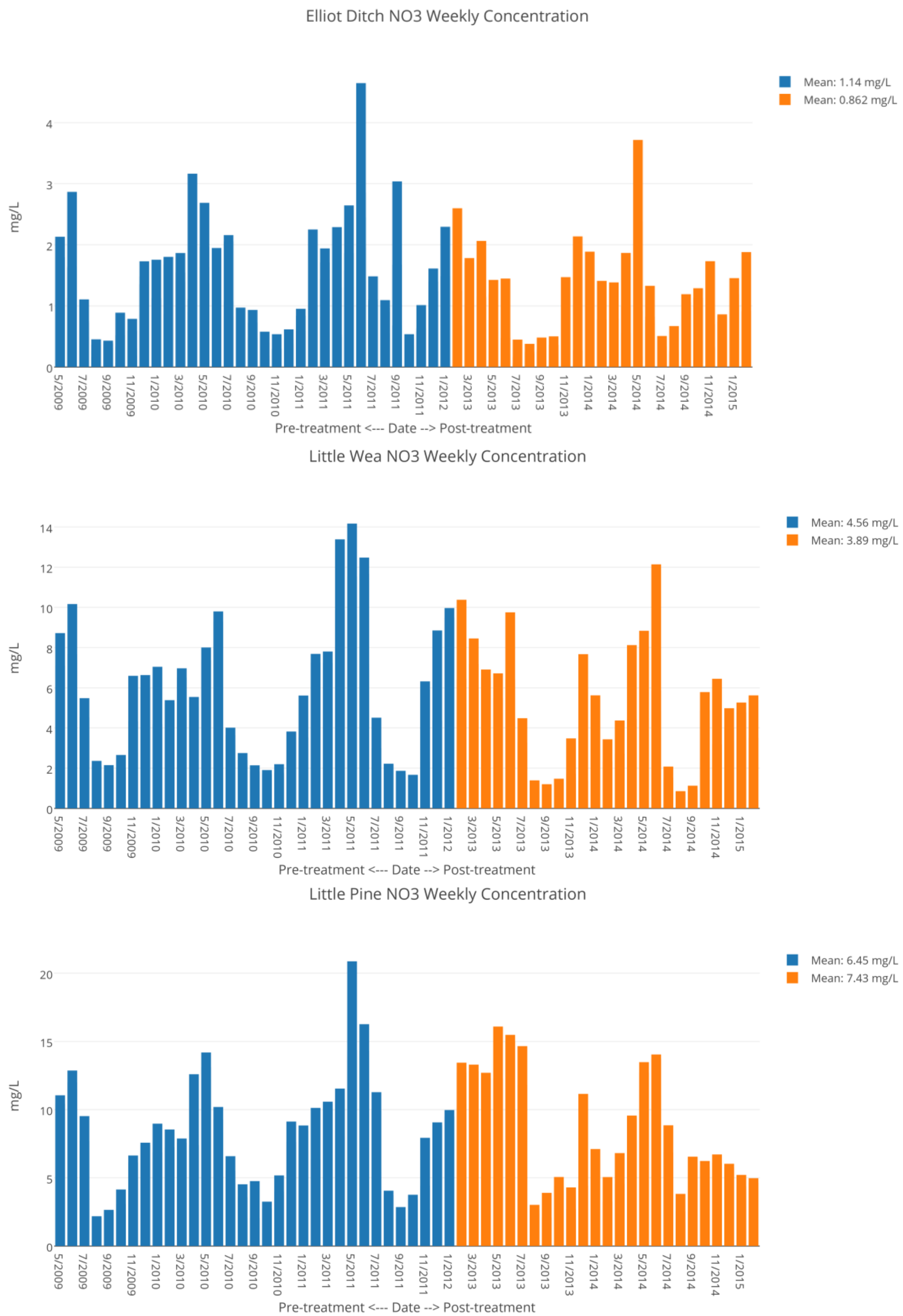


Figure 16. Pre-treatment and post-treatment time series for weekly nitrate-nitrogen concentration between Elliot Ditch, Little Wea Creek and Little Pine Creek.

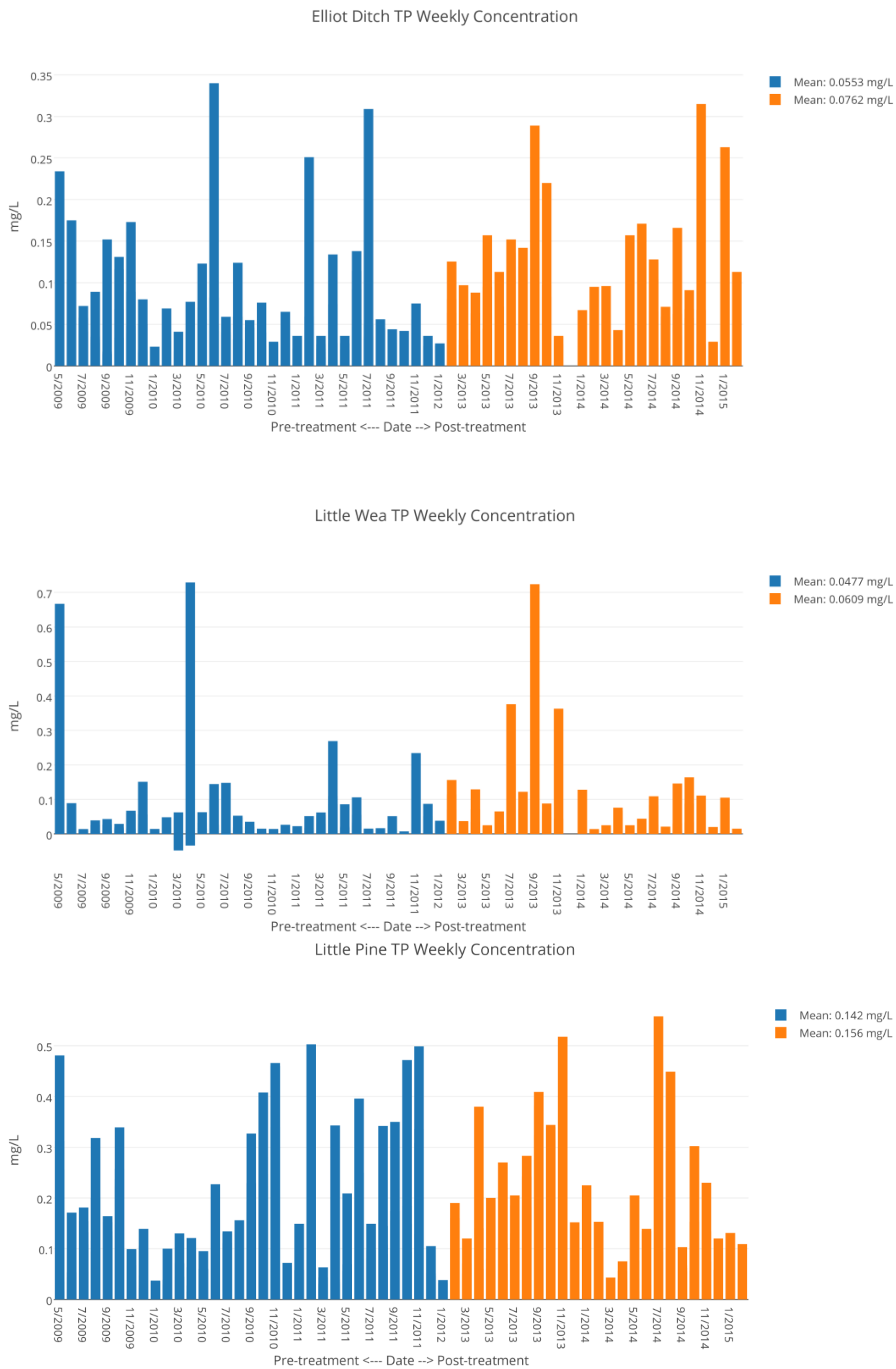


Figure 17. Pre-treatment and post-treatment time series for weekly Total Phosphorus concentration between Elliot Ditch, Little Wea Creek and Little Pine Creek.

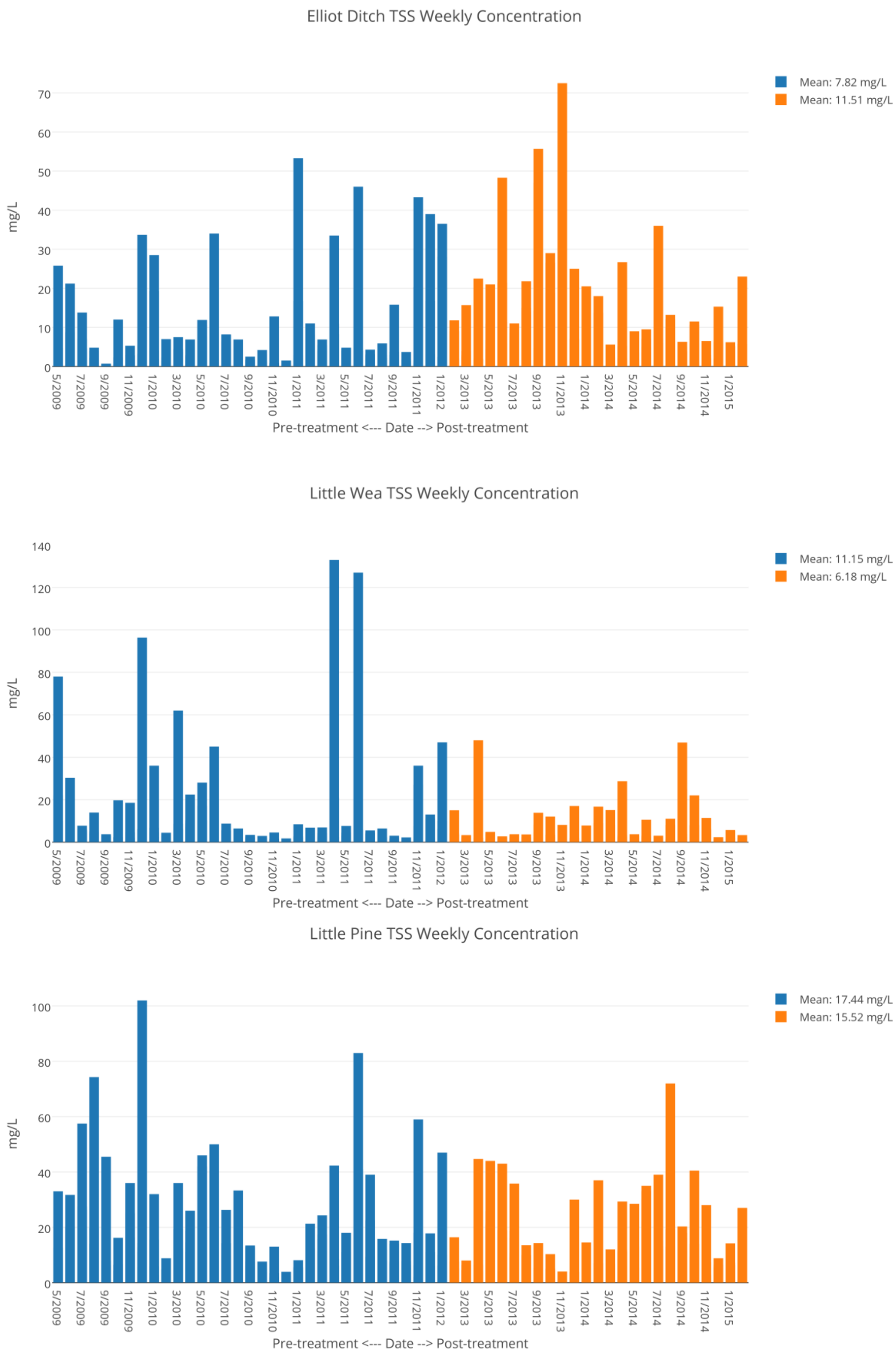


Figure 18. Pre-treatment and post-treatment time series for weekly Total Suspended Solids concentration between Elliot Ditch, Little Wea Creek and Little Pine Creek.

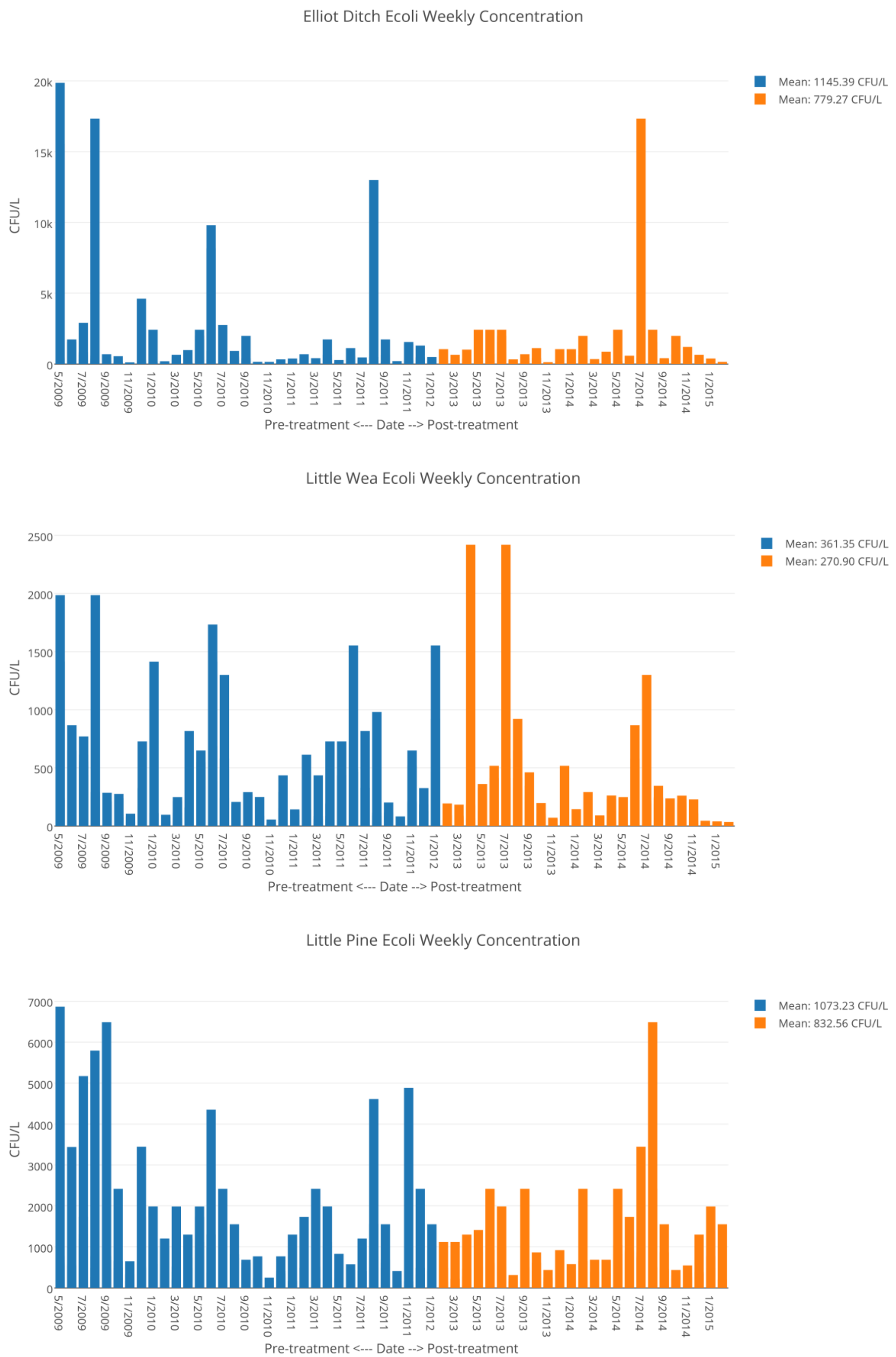
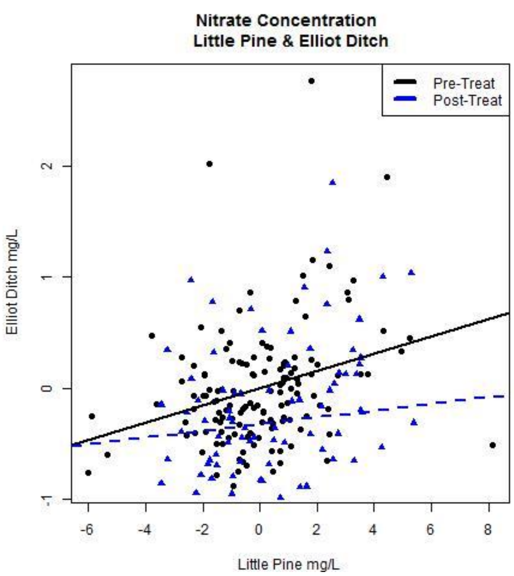


Figure 19. Pre-treatment and post-treatment time series for weekly *E. coli* concentration between Elliot Ditch, Little Wea Creek and Little Pine Creek.

ANCOVA: The ANCOVA analysis first involves a one-way ANOVA to test for correlation between the paired watersheds for the pre- and post-treatments. If the ANOVA results yielded significant relationships between the watersheds, then an ANCOVA was conducted to test for a change in the relationship following treatment. To determine significance in the ANCOVA results, a p-value of less than 0.05 or 0.10 was required for either the intercept or the slope. This indicates a significant change in the regression relationship due to the management practice. Prior to conducting the ANOVA and ANCOVA tests, four plots were run on the residual data to determine the distribution for the actual versus predicted values (not shown).

The pre- and post-treatment period regression relationships that result from the ANOVA analysis (using the transformed data) are shown in Figure 20 through Figure 23 for the Elliot Ditch/Little Pine pair and the Little Wea /Little Wea pairs. As summarized in Table 7, total phosphorus for Little Wea Creek and Little Pine Creek were the only pair to show no significant relationship in both the pre- and post-treatment ANOVA test. Although the post-treatment regression for Elliot Ditch/Little Pine Creek nitrate-nitrogen and total phosphorus were also not statistically significant, making the usefulness of ANCOVA for these pairs questionable. The ANCOVA analysis indicates a statistically significant difference in nitrate-nitrogen and E. coli in Little Wea Creek and total suspended solids in Elliot Ditch. These two changes are consistent with those detected in the two-sample t-test, and suggest that they may not be due to weather alone.



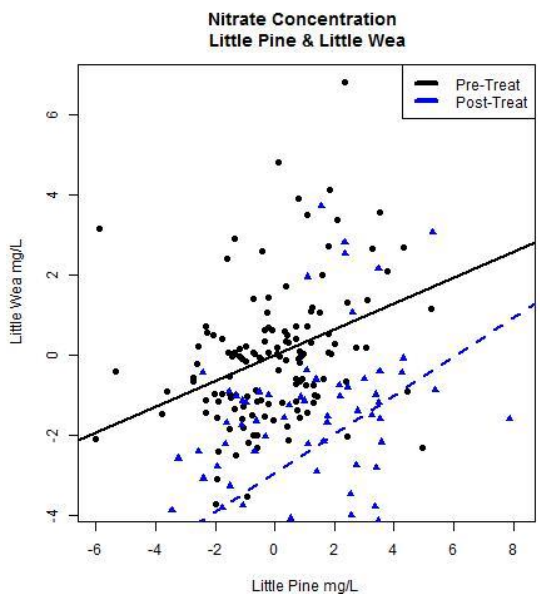


Figure 20. Pre-treatment and post-treatment regression relationships for weekly nitrate-nitrogen concentration between Elliot Ditch and Little Pine Creek and Little Wea Creek and Little Pine Creek.

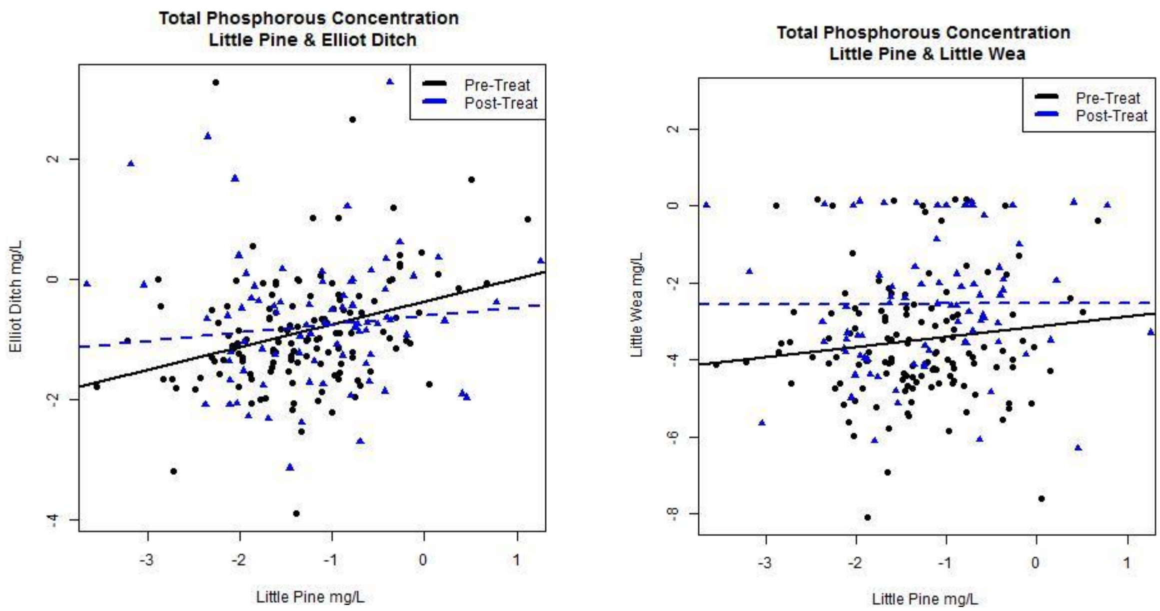


Figure 21. Pre-treatment and post-treatment regression relationships for weekly Total Phosphorus concentration between Elliot Ditch and Little Pine Creek and Little Wea Creek and Little Pine Creek.

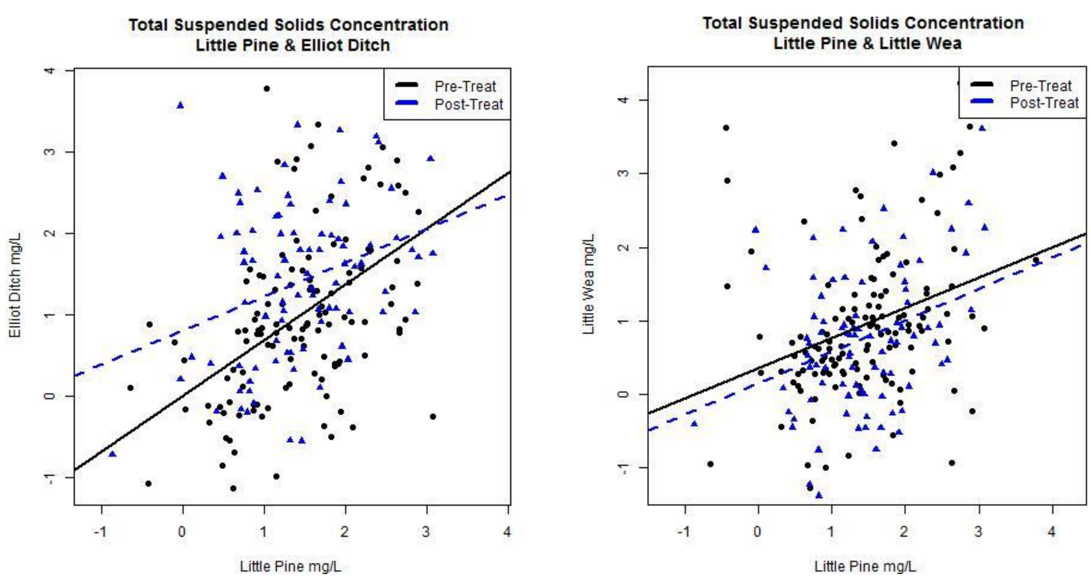


Figure 22. Pre-treatment and post-treatment regression relationships for weekly TSS concentration between a) Elliot Ditch and Little Pine Creek and b) Little Wea Creek and Little Pine Creek

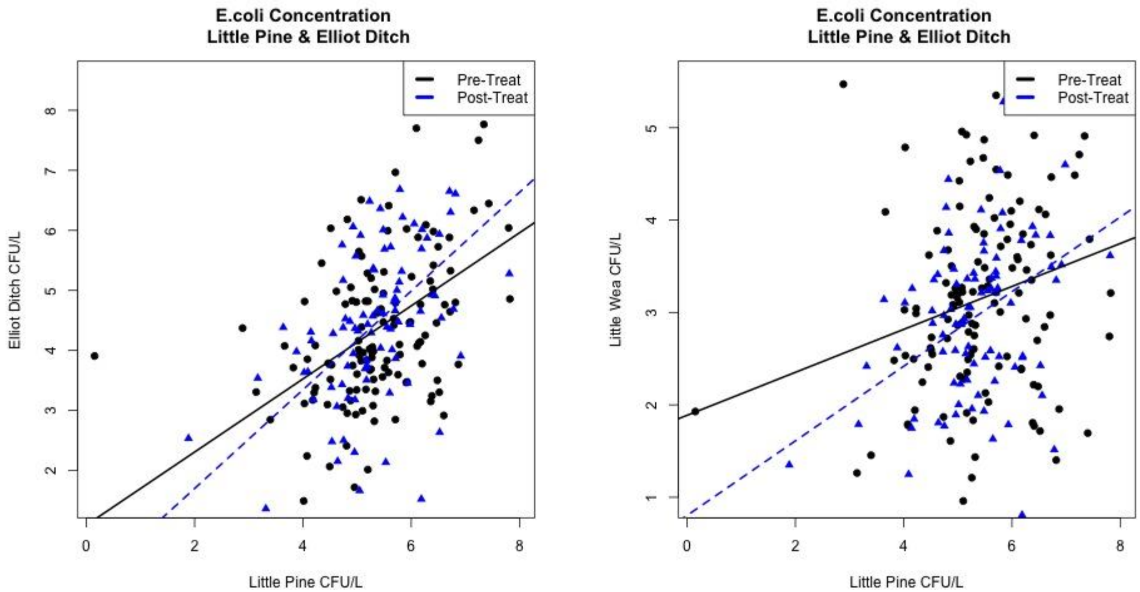


Figure 23. Pre-treatment and post-treatment regression relationships for weekly e. coli concentration between a) Elliot Ditch and Little Pine Creek and b) Little Wea Creek and Little Pine Creek

<u>Mann-Kendall Test:</u> Time series graphs of the concentration data sets with a smoothed line are shown in Figure 24 to Figure 27. The line represents a Locally-Weighted Scatter Plot Smoothing (LOWESS) with a span of 20. This is done rather than a linear regression line to visualize any potential trends in the dataset, since Mann-Kendall is a test for non-linear trend.

The pre- and post-treatment time series were combined for the trend test. In addition to tests on the time series for each watershed, two difference series were also tested: Elliot Ditch – Little Pine Creek and Little Wea Creek– Little Pine Creek. This is an alternative method of paired catchment analysis, in which the difference series is less responsive to weather related changes in the response variables that affect both watersheds. As summarized in Table 8, there was a statistically significant decreasing trend in TSS concentration over time in Little Wea Creek as well as a decreasing trend in E. coli in Little Wea Creek.

Table 7. ANOVA and ANCOVA test results for weekly concentration.

Watershed & Pollutant	ANOVA p-value	ANCOVA p-value	Significant?	Percent Magnitude of Change[1]
NO3				
Elliot Ditch & Little Pine	Pre: 0.00** Post: 0.33	Slope: 0.19 Intercept: 0.00**	No sig. relation. Sig. change	< 0.001
Little Wea & Little Pine	Pre: 0.00** Post: 0.00**	Slope: 0.17 Intercept: 0.00**	Sig. relation. Sig. change.	77.8%
Total Phosphorus				
Elliot Ditch & Little Pine	Pre: 0.00** Post: 0.29	Slope: 0.13 Intercept: 0.34	No sig. relation. No sig. change	-0.002
Little Wea & Little Pine	Pre: 0.16 Post: 0.97	Slope: 0.33 Intercept: 0.15	No sig. relation. No sig. change	0.004
Total Suspended Solids				
Elliot Ditch & Little Pine	Pre: 0.00** Post: 0.00**	Slope: 0.11 Intercept: 0.00**	Sig. relation. Sig. change	-0.001
Little Wea & Little Pine	Pre: 0.00** Post: 0.00**	Slope: 0.93 Intercept: 0.47	Sig. relation. No sig. change	0.001
E. coli				
Elliot Ditch & Little Pine	Pre: 0.00** Post: 0.00**	Slope: 0.15 Intercept: 0.20	Sig. relation. No sig. change	0.005
Little Wea & Little Pine	Pre: 0.00** Post: 0.00**	Slope: 0.16 Intercept: 0.10*	Sig. relation. No sig. change	<-0.001

[1]Calculated by predicting the treatment watershed response for the post-treatment period using the pre-treatment regression equation for the untransformed data and post-treatment observation from the controlled watershed. Difference is 100% *(observed – predicted)/observed, so a positive value indicates an increase in the treatment watershed response.
**Statistically significant at α = 0.05
*Statistically significant at α = 0.1

Table 8. Seasonal Mann-Kendall test for trend results for concentration.

Watershed & Pollutant	p-value	Trend Slope
Nitrate-Nitrogen		
Elliot Ditch	0.16	-0.04 mg/l/yr
Little Pine	0.16	0.11 mg/l/yr
Little Wea	0.00**	-0.21 mg/l/yr
Elliot Ditch - Little Pine	0.51	-0.13 mg/l/yr
Little Wea - Little Pine	0.02**	-0.19 mg/l/yr
Total Phosphorus		
Elliot Ditch	0.21	0.0038 mg/l/yr
Little Pine	0.94	0.0013 mg/l/yr
Little Wea	0.94	0.0002 mg/l/yr
Elliot Ditch - Little Pine	0.94	0.008 mg/l/yr
Little Wea - Little Pine	0.77	-0.0017 mg/l/yr
Total Suspended Solids		
Elliot Ditch	0.16	0.5 mg/l/yr
Little Pine	0.94	0.06 mg/l/yr
Little Wea	0.30	-0.28 mg/l/yr
Elliot Ditch - Little Pine	0.21	0.46 mg/l/yr
Little Wea - Little Pine	0.61	-0.51 mg/l/yr
E. coli		
Elliot Ditch	0.42	12.83 CFU/l/yr
Little Pine	0.83	-9.4 CFU/l/yr
Little Wea	0.05**	-13.1 CFU/l/yr
Elliot Ditch - Little Pine	0.34	28.0 CFU/l/yr
Little Wea - Little Pine	0.94	2.3 CFU/l/yr

**Statistically significant at ⍰ = 0.05

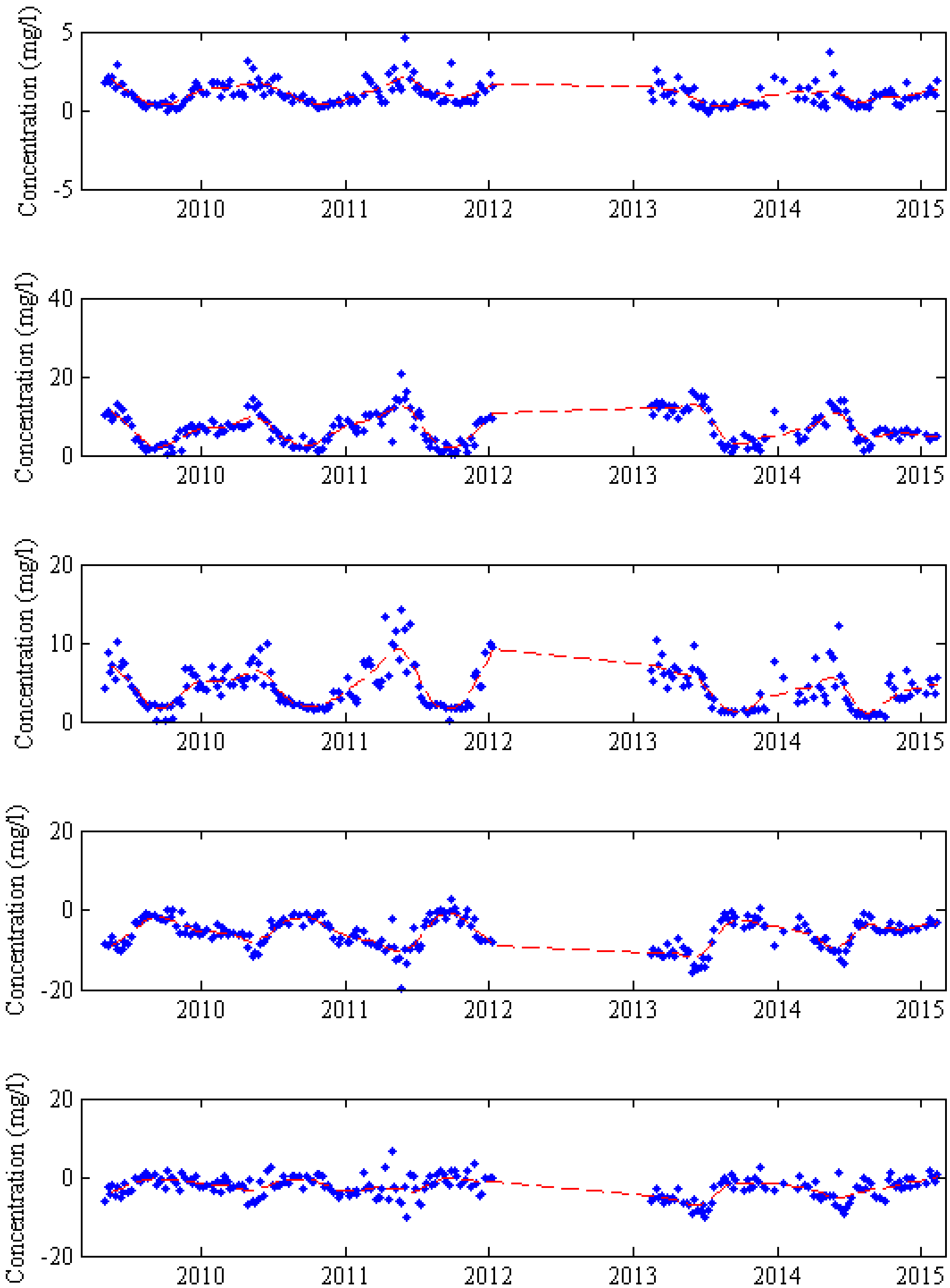


Figure 24. Trend analysis for weekly nitrate concentrations in: a) Elliot Ditch, b) Little Pine Creek, c) Little Wea Creek, d) Elliot Ditch – Little Pine Creek and e) Little Wea – Little Pine Creek.

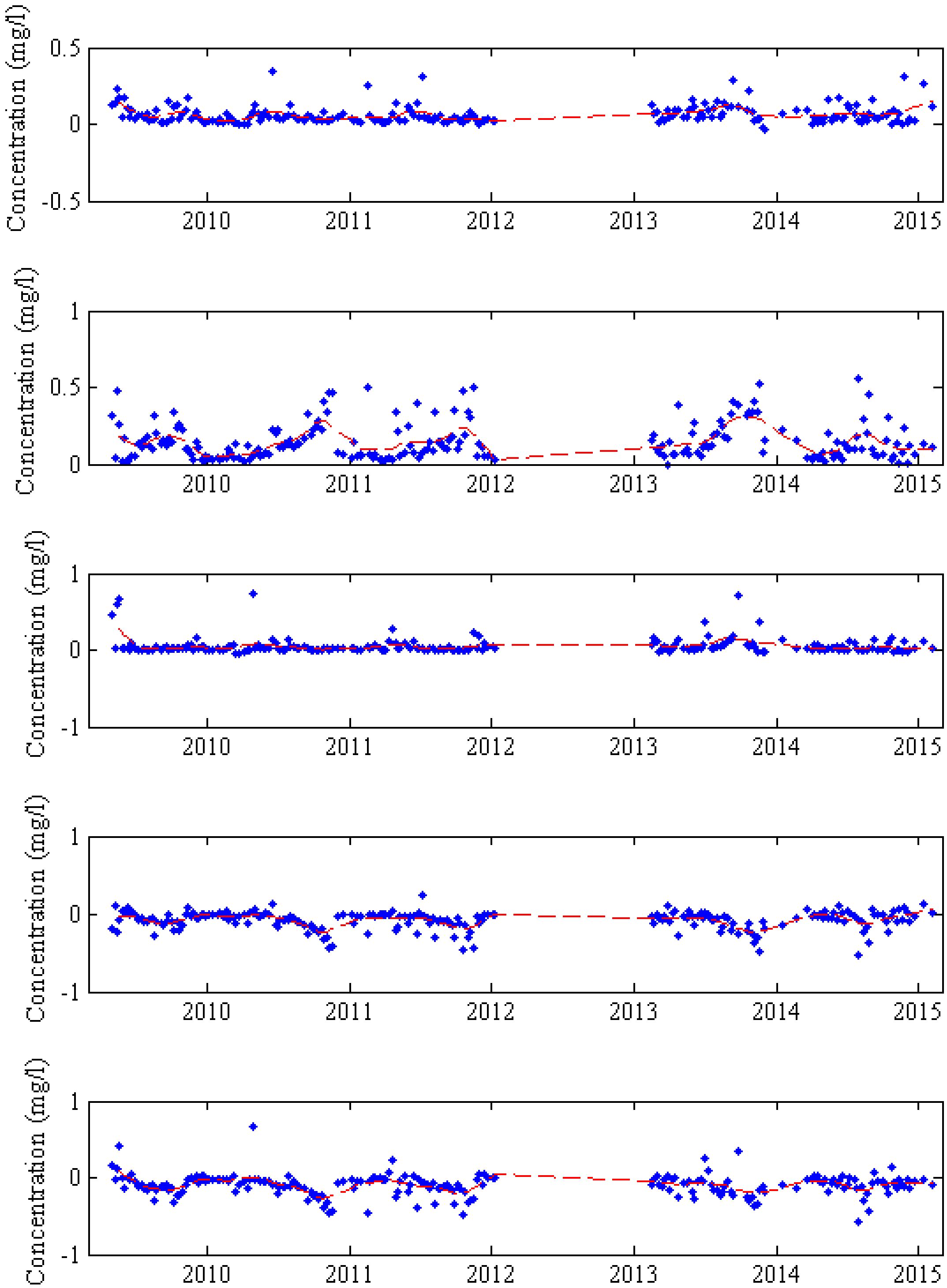


Figure 25. Trend analysis for weekly total phosphorus concentrations in: a) Elliot Ditch, b) Little Pine Creek, c) Little Wea Creek, d) Elliot Ditch – Little Pine Creek and e) Little Wea – Little Pine Creek.

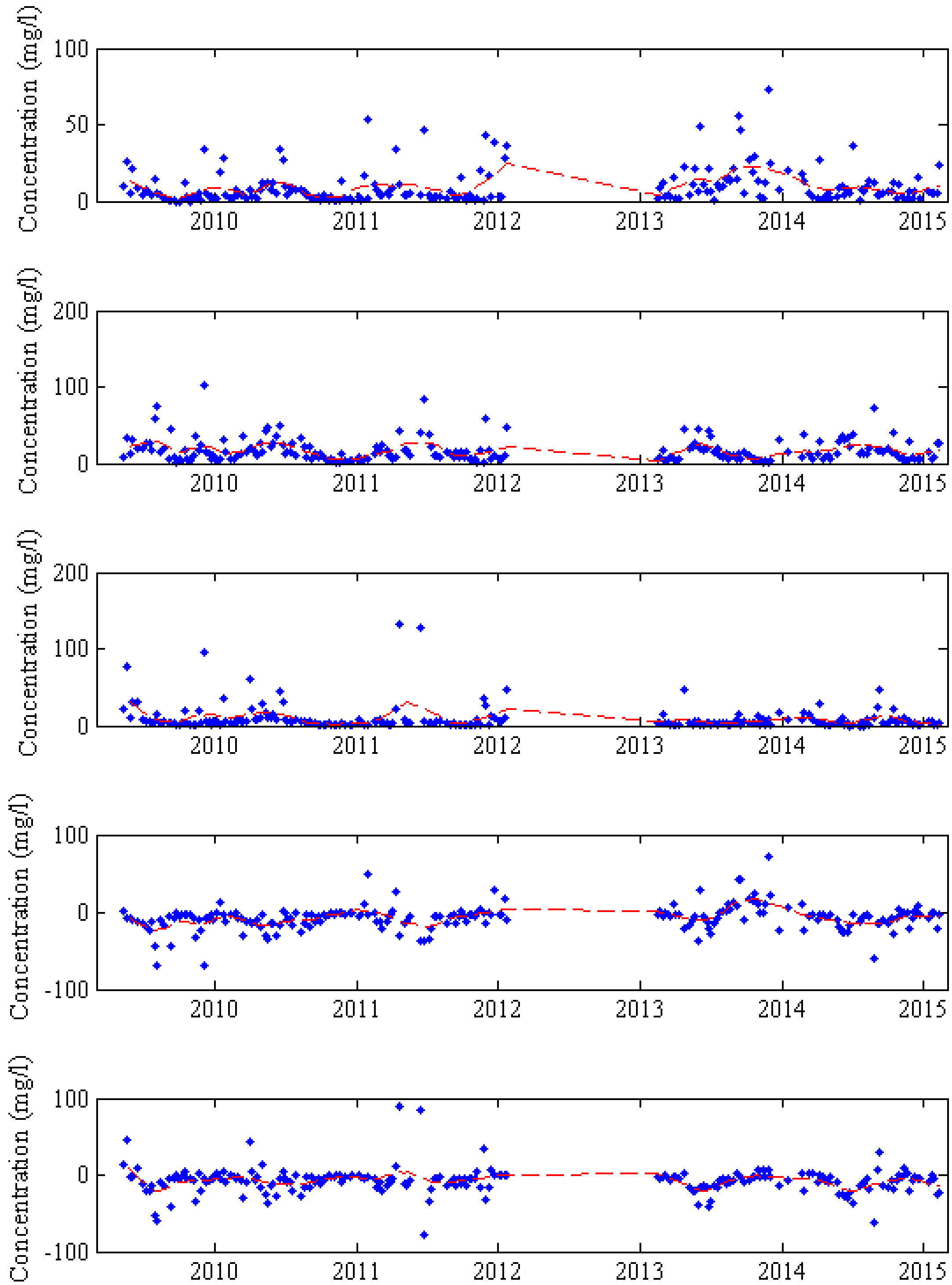


Figure 26. Trend analysis for weekly total suspended solid concentrations in: a) Elliot Ditch, b) Little Pine Creek, c) Little Wea Creek, d) Elliot Ditch – Little Pine Creek and e) Little Wea – Little Pine Creek.

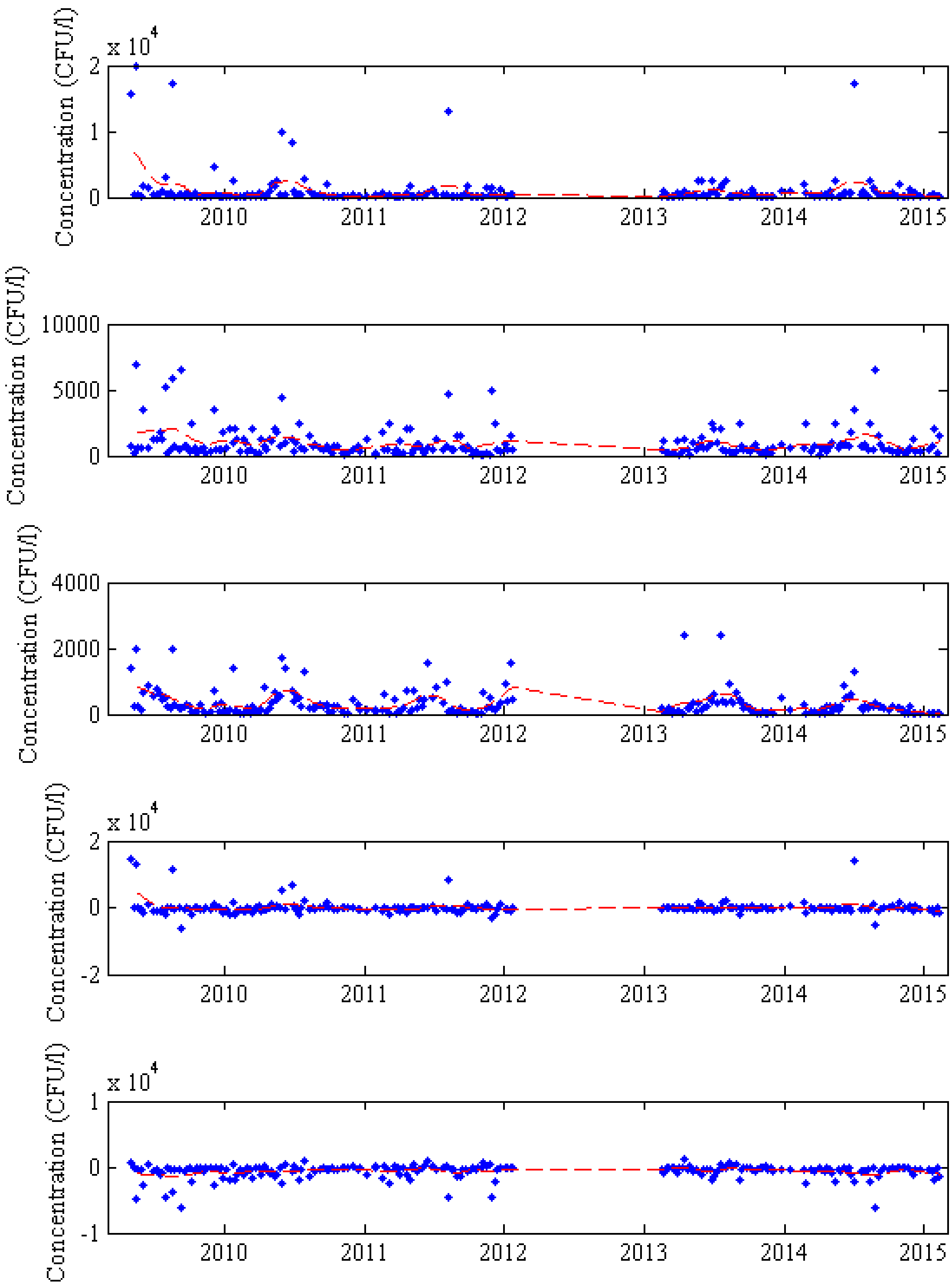


Figure 27. Trend analysis for weekly e. coli concentrations in: a) Elliot Ditch, b) Little Pine Creek, c) Little Wea Creek, d) Elliot Ditch – Little Pine Creek and e) Little Wea – Little Pine Creek.

2.3.2 **Load Data**

Two-sample T-Test: Figure 28 through Figure 31 shoe tiem series developed for all load data in the three watersheds with their corresponding means for the pre/post treatment periods. All three watersheds showed decreases in mean loads for NO_3, TSS, and E. coli. Little Wea and Elliot Ditch had load decreases for TP, while Little Pine Creek had an increase. The results of the F-test for equal variance did not show any significant results. As summarized in Table 9, the only significant results for monthly load were for nitrate-nitrogen in Elliot Ditch and Little Pine Creek. Elliot Ditch saw a decrease in load over time, as did Little Pine Creek. All other results were not statistically significant.

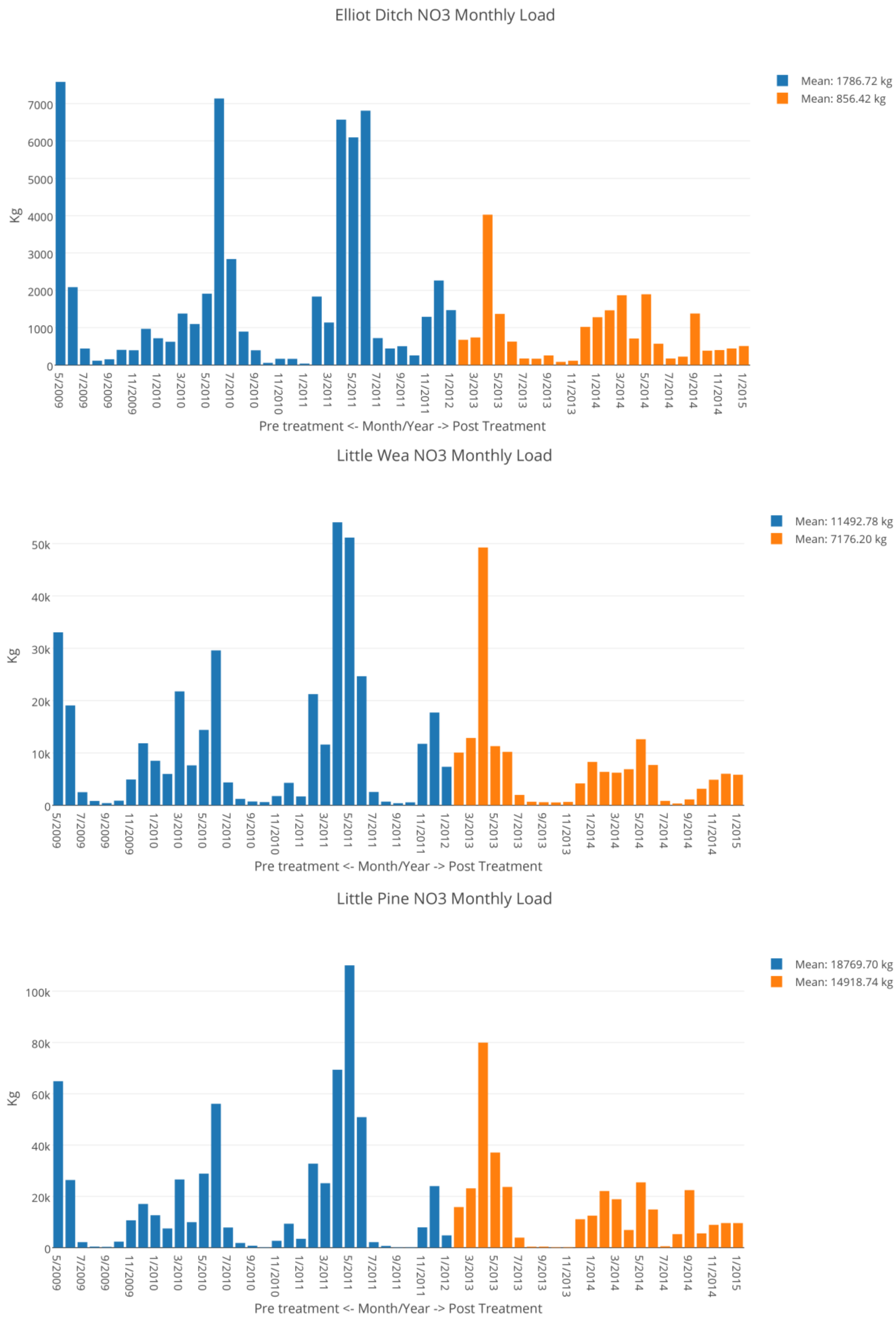


Figure 28. Pre-treatment and post-treatment time series for monthly nitrate-nitrogen load between Elliot Ditch, Little Wea Creek and Little Pine Creek.

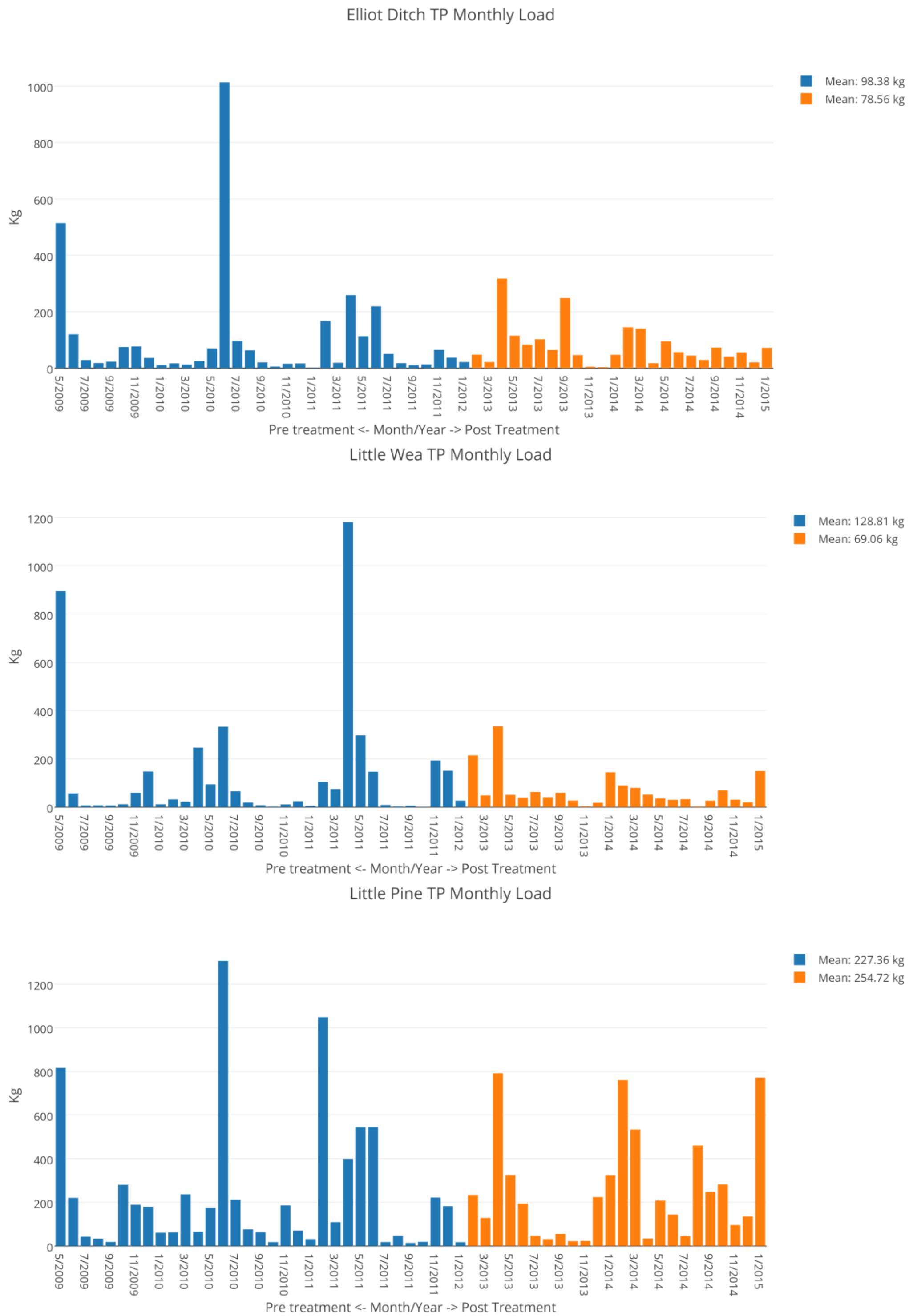


Figure 29. Pre-treatment and post-treatment time series for monthly Total Phosphorus load between Elliot Ditch, Little Wea Creek and Little Pine Creek.

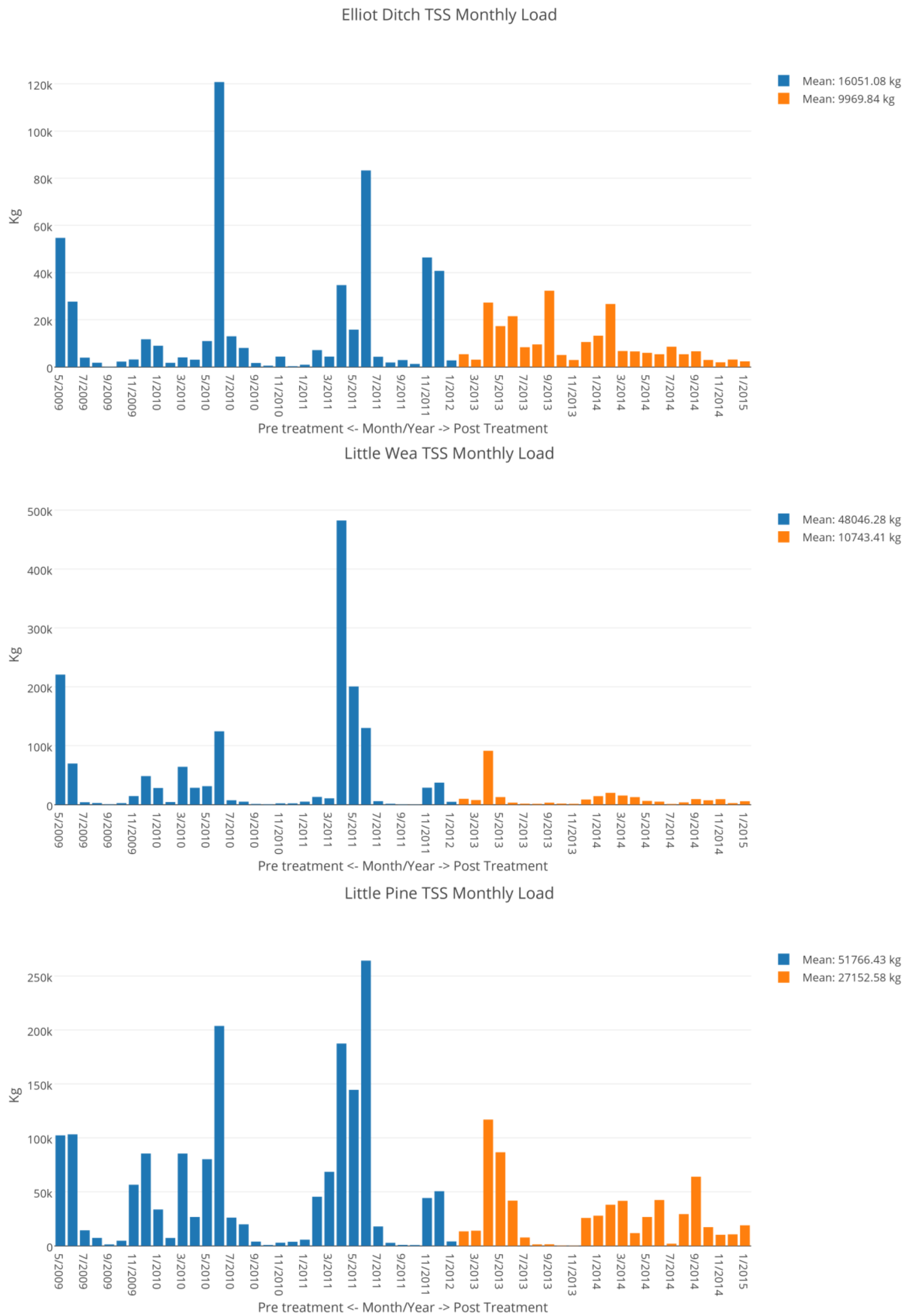
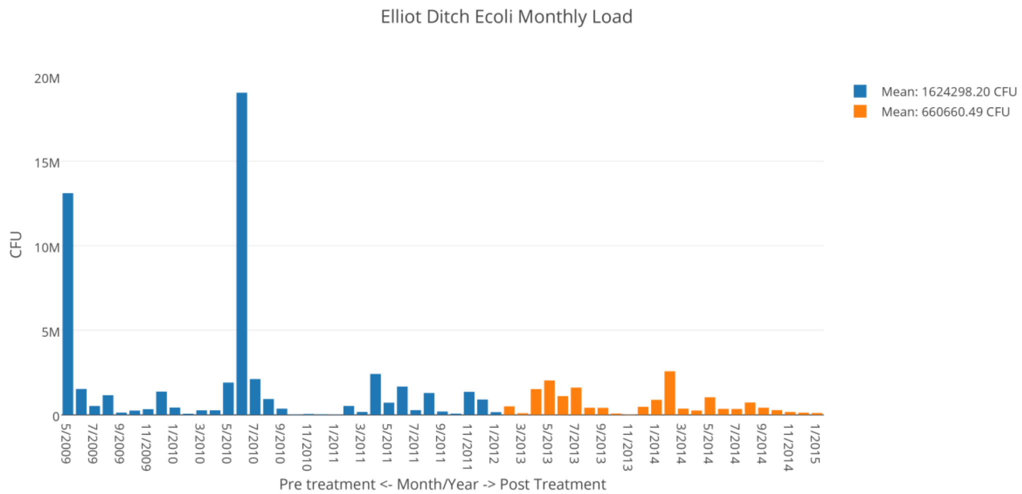
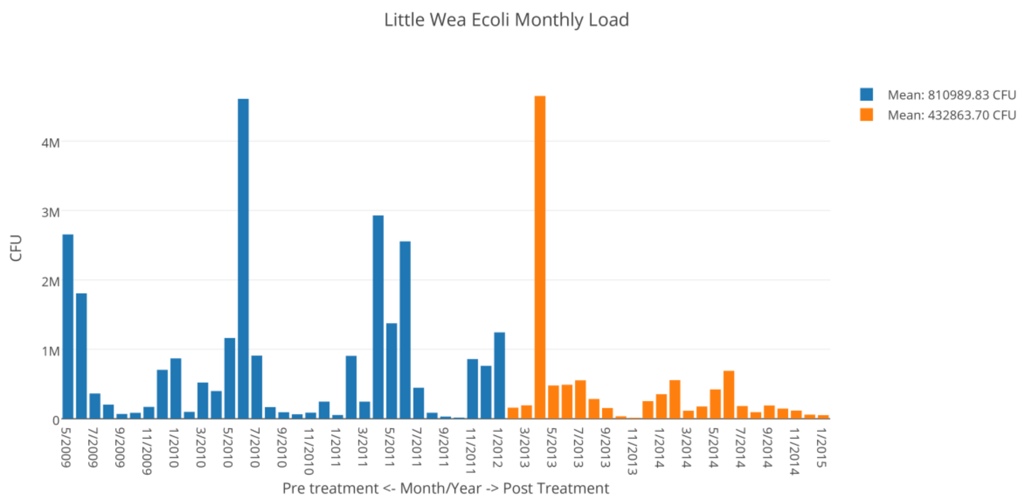


Figure 30. Pre-treatment and post-treatment time series for monthly Total Suspended Solids load between Elliot Ditch, Little Wea Creek and Little Pine Creek.





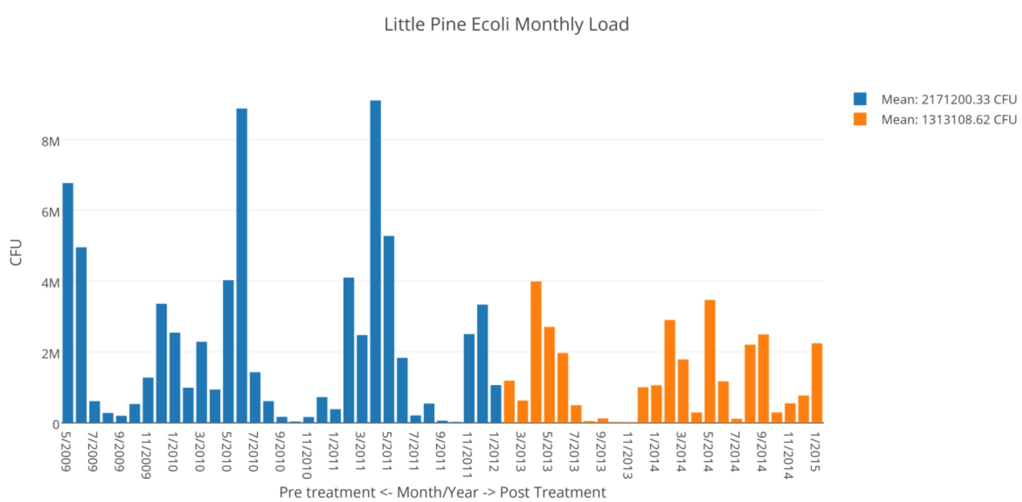


Figure 31. Pre-treatment and post-treatment time series for monthly E. coli load between Elliot Ditch, Little Wea Creek and Little Pine Creek.

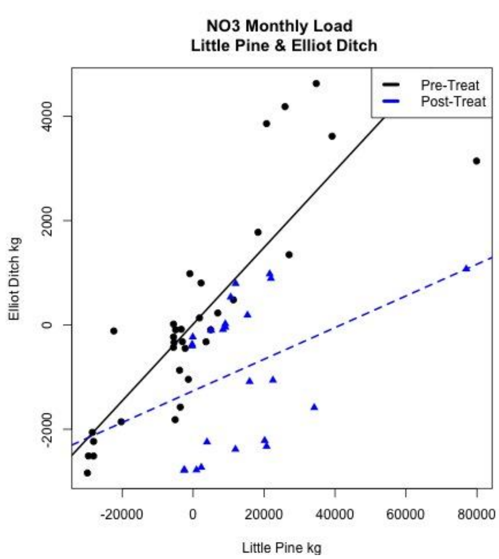
Table 9. Two sample T-Test results for monthly load.

Watershed and Pollutant	p-value	T-value	Difference in Mean[1] untransformed
Nitrate-Nitrogen			
Little Pine**	0.03	-2.18	-3850.96 kg/month
Elliot Ditch**	0.04	2.06	-930.30 kg/month
Little Wea	0.14	-1.48	-4316.59 kg/month
Total Phosphorus			
Little Pine	0.28	-0.73	27.36 kg/month
Elliot Ditch	0.40	1.05	-19.82 kg/month
Little Wea	0.48	-0.71	-59.75 kg/month
Total Suspended Solids			
Little Pine	0.47	-0.73	-24613.85 kg/month
Elliot Ditch	0.30	1.05	-6081.24 kg/month
Little Wea	0.63	0.49	-37302.87 kg/month
E. coli			
Little Pine	0.36	0.92	-858091.71 CFU/month
Elliot Ditch	0.79	0.27	-963637.70 CFU/month
Little Wea	0.13	1.54	-378126.13 CFU/month

[1]All differences are presented as post-treatment – pre-treatment

**Statistically significant at α = 0.05

ANCOVA: All ANOVA values for monthly load returned significant results (at α=0.10), so the ANCOVA was run on the pairs. These regressions are shown in Figure 32 to Figure 35. The results are summarized in Table . Elliot Ditch and Little Pine pairs for Nitrate and Total Phosphorus showed statistically significant changes for both the slope and intercept, while the pair Little Wea and Little Pine had significant changes for E. coli and Total Phosphorus.



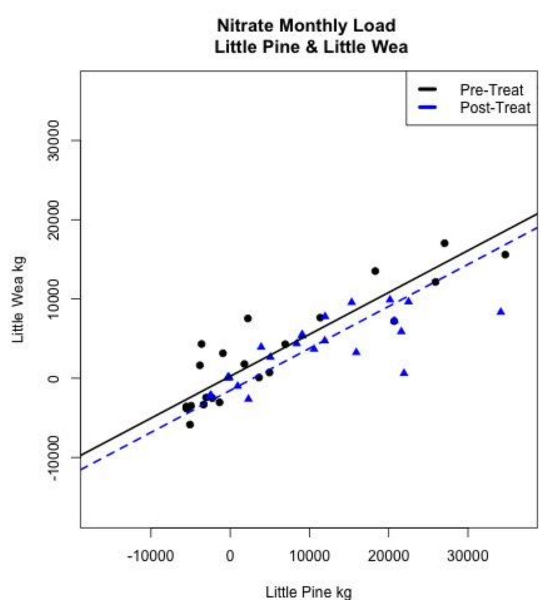


Figure 32. Pre-treatment and post-treatment regression relationships for monthly nitrate load between a) Elliot Ditch and Little Pine Creek and b) Little Wea Creek and Little Pine Creek.

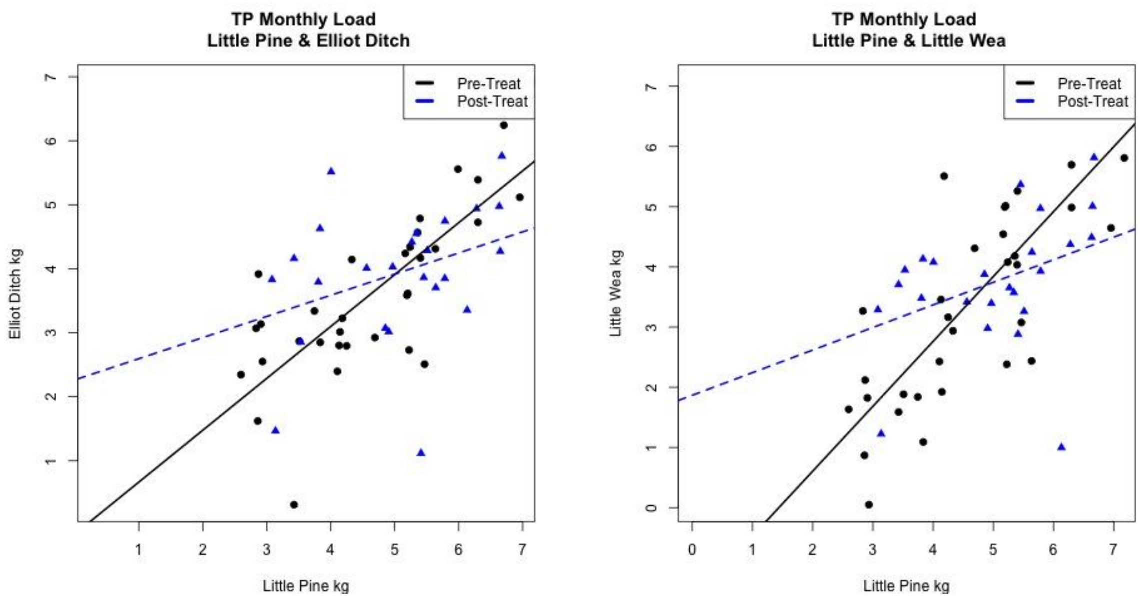


Figure 33. Pre-treatment and post-treatment regression relationships for monthly Total Phosphorus load between a) Elliot Ditch and Little Pine Creek and b) Little Wea Creek and Little Pine Creek.

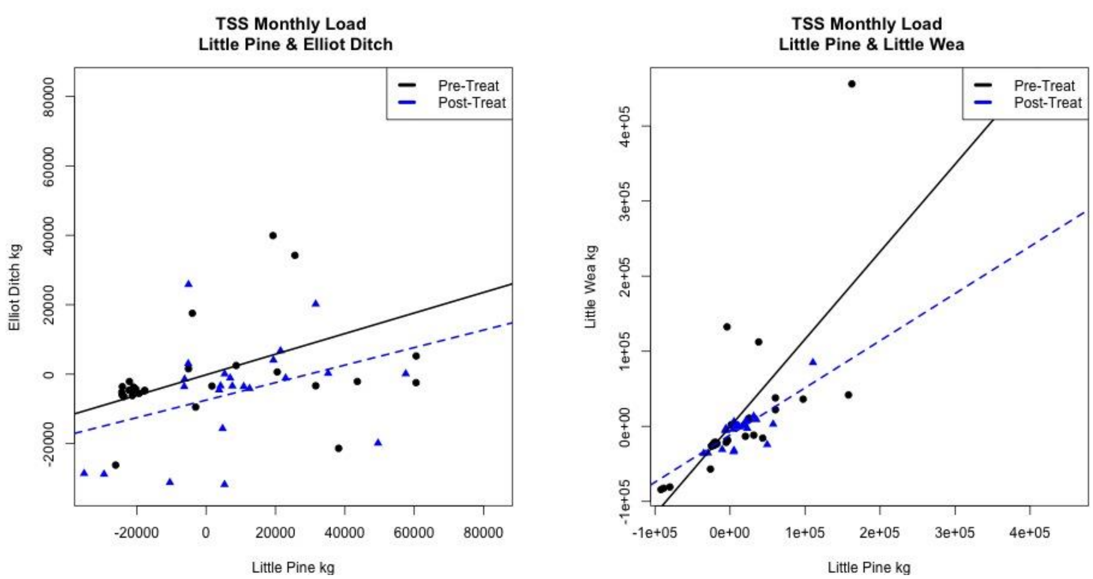


Figure 34. Pre-treatment and post-treatment regression relationships for monthly TSS load between a) Elliot Ditch and Little Pine Creek and b) Little Wea Creek and Little Pine Creek.

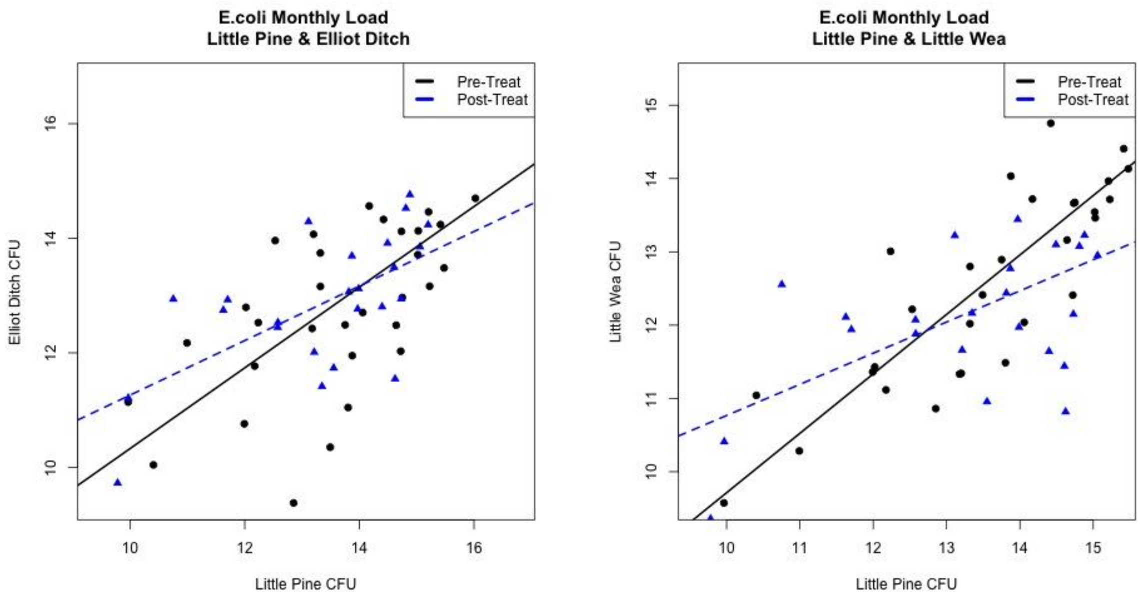


Figure 35. Pre-treatment and post-treatment regression relationships for monthly *E. coli* load between a) Elliot Ditch and Little Pine Creek and b) Little Wea Creek and Little Pine Creek.

Table 10. ANOVA and ANCOVA test results for monthly load.

Watershed & Pollutant	ANOVA p-value	ANCOVA p-value	Is there a detectable change?	Percent Magnitude of Change[1]
Nitrate-Nitrogen				
Elliot Ditch & Little Pine	Pre: 0.00** Post: 0.07*	Slope: 0.01** Intercept: 0.00**	Sig. Relation. Sig. Change	-0.002
Little Wea & Little Pine	Pre: 0.00** Post: 0.00**	Slope: 0.99 Intercept: 0.15	Sig. Relation. No Sig. Change	--27.4
Total Suspended Solids				
Elliot Ditch & Little Pine	Pre: 0.00** Post: 0.02**	Slope: 0.71 Intercept: 0.10*	Sig. Relation. No Sig. Change	32.5
Little Wea & Little Pine	Pre: 0.00** Post: 0.00**	Slope: 0.18 Intercept: 0.45	Sig. Relation. No Sig. Change	-0.007
Total Phosphorus				
Elliot Ditch & Little Pine	Pre: 0.00** Post: 0.10*	Slope: 0.03** Intercept: 0.03**	Sig. Relation. Sig. Change	--0.001
Little Wea & Little Pine	Pre: 0.00** Post: 0.06*	Slope: 0.00** Intercept: 0.00**	Sig. Relation. Sig. Change	--0.009
E. coli				
Elliot Ditch & Little Pine	Pre: 0.00** Post: 0.00**	Slope: 0.25 Intercept: 0.24	Sig. Relation. No Sig. Change	-0.002
Little Wea & Little Pine	Pre: 0.00** Post: 0.00**	Slope: 0.01** Intercept: 0.02**	Sig. Relation. Sig. Change	-27.4

[1]Calculated by predicting the treatment watershed response for the post-treatment period using the pre-treatment regression equation for the untransformed data and post-treatment observation from the controlled watershed. Difference is 100% *(observed – predicted)/observed, so a positive value indicates an increase in the treatment watershed response.

**Statistically significant at α = 0.05

*Statistically significant at α = 0.1

<u>Mann-Kendall Test:</u> Time series graphs of the load data sets with a smoothed line are shown in Figure 36 to Figure 39 and results are summarized in Table 11. There is a statistically significant decreasing trend in nitrate load in Little Wea Creek at a significance level of α = 0.05. In addition, there are decreasing trends in TSS load for Little Pine Creek and E. coli load for Little Pine Creek and Elliot Ditch at a significance level of α = 0.10.

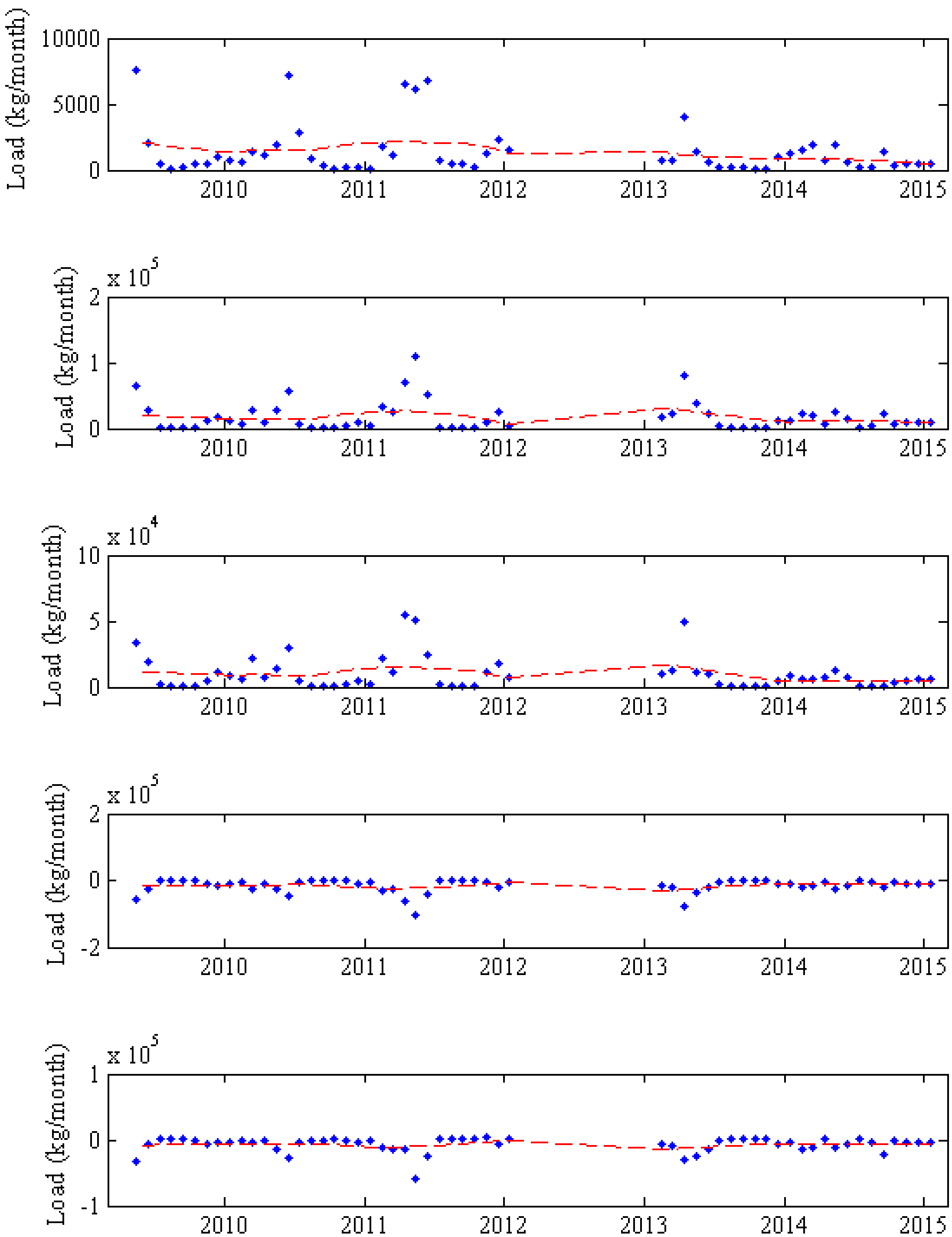


Figure 36. Trend analysis for monthly nitrate load in: a) Elliot Ditch, b) Little Pine Creek, c) Little Wea Creek, d) Elliot Ditch – Little Pine Creek and e) Little Wea – Little Pine Creek.

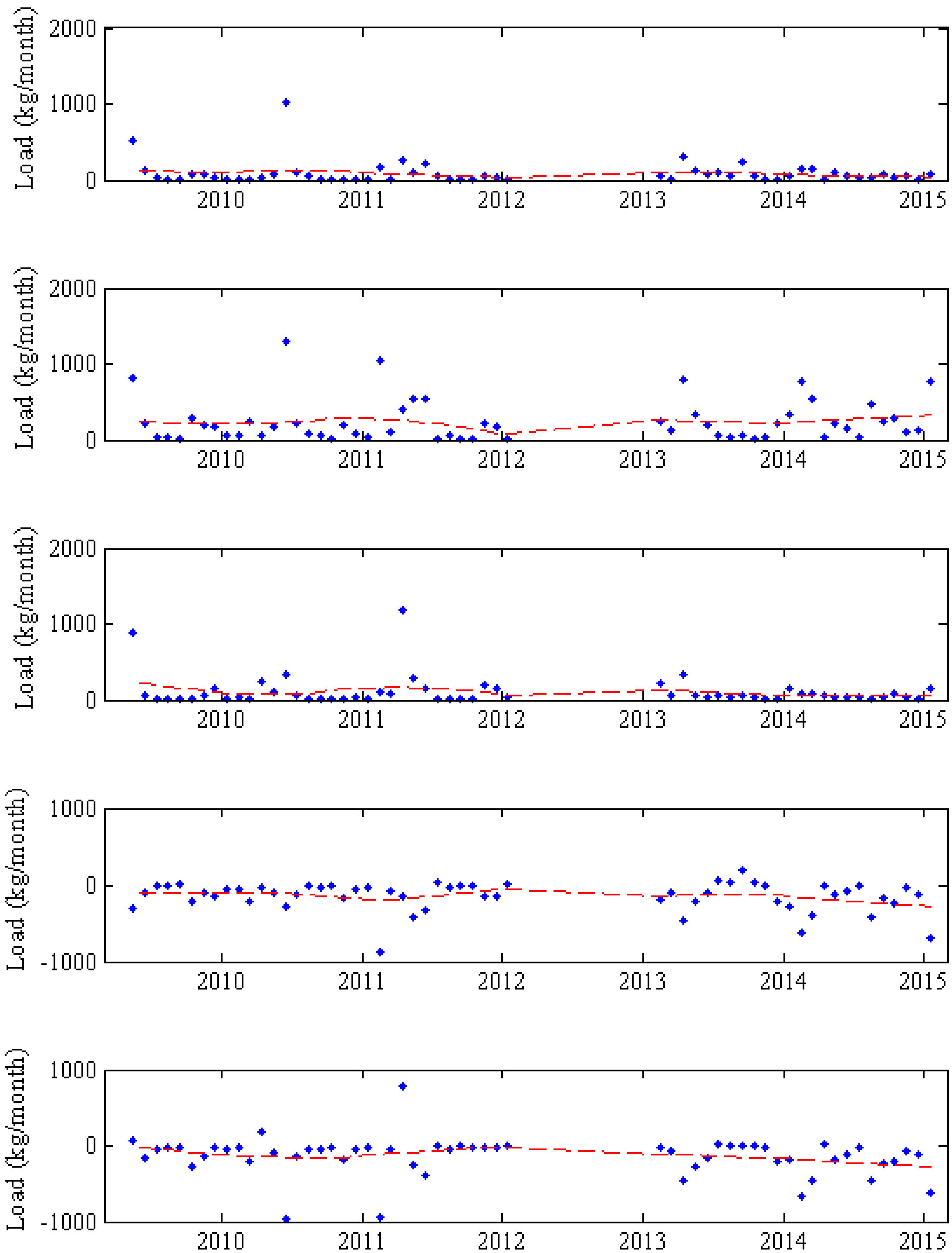


Figure 37. Trend analysis for monthly Total Phosphorus load in: a) Elliot Ditch, b) Little Pine Creek, c) Little Wea Creek, d) Elliot Ditch – Little Pine Creek and e) Little Wea – Little Pine Creek.

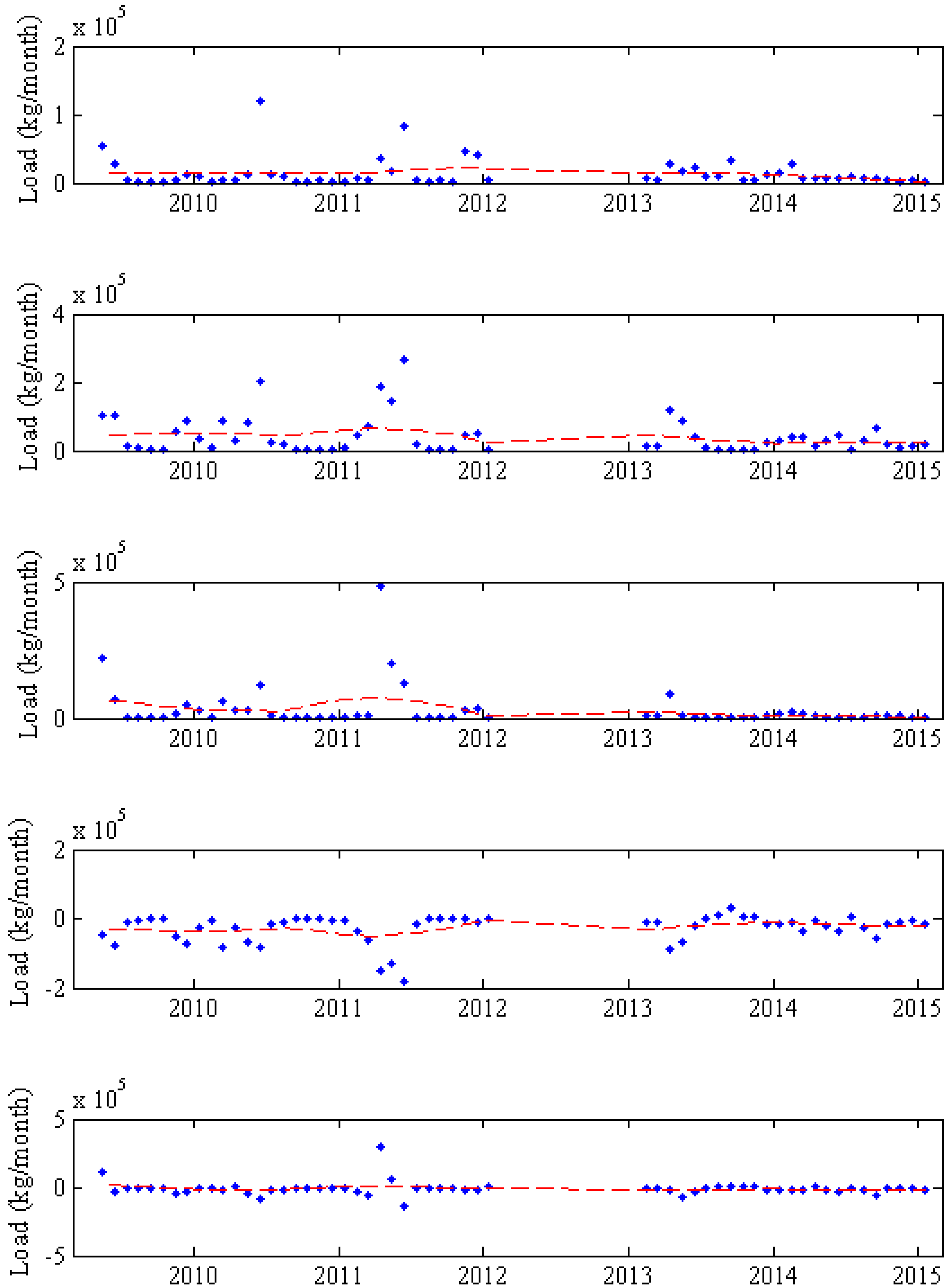


Figure 38. Trend analysis for monthly Total Suspended Solid load in: a) Elliot Ditch, b) Little Pine Creek, c) Little Wea Creek, d) Elliot Ditch – Little Pine Creek and e) Little Wea – Little Pine Creek.

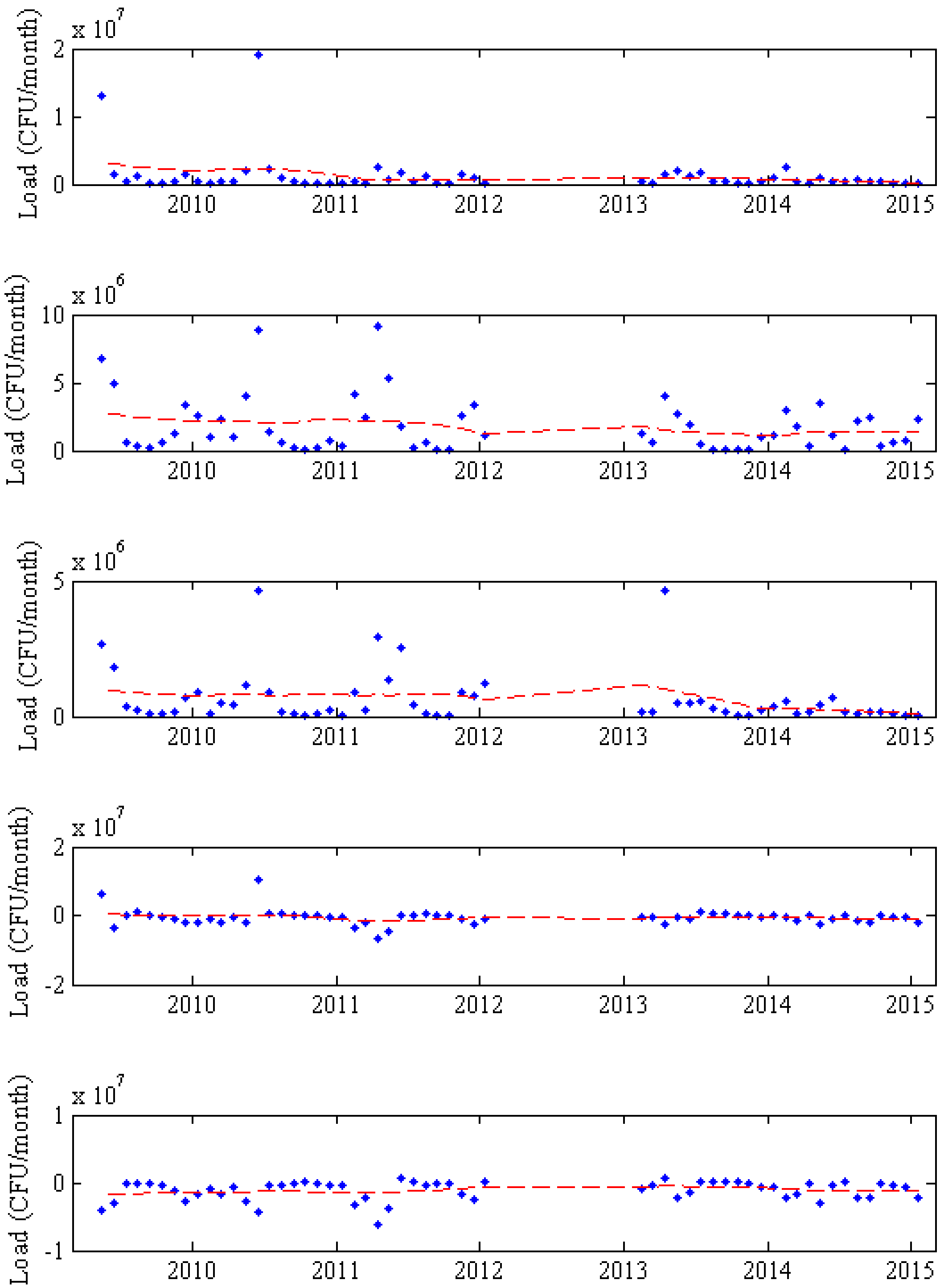


Figure 39. Trend analysis for monthly E. coli load in: a) Elliot Ditch, b) Little Pine Creek, c) Little Wea Creek, d) Elliot Ditch – Little Pine Creek and e) Little Wea – Little Pine Creek.

Table 5. Seasonal Mann-Kendall test trend results for monthly load.

Watershed & Pollutant	p-value	Trend Slope
Nitrate-Nitrogen		
Elliot Ditch	0.41	-50 kg/yr
Little Pine	0.41	-237 kg/yr
Little Wea	0.01^{**}	-183 kg/yr
Elliot Ditch - Little Pine	0.11	247 kg/yr
Little Wea - Little Pine	0.71	-26 kg/yr
Total Phosphorus		
Elliot Ditch	0.60	1.3 kg/yr
Little Pine	0.71	1.0 kg/yr
Little Wea	0.94	0.3 kg/yr
Elliot Ditch - Little Pine	0.94	1.0 kg/yr
Little Wea - Little Pine	0.82	-5.5 kg/yr
Total Suspended Solids		
Elliot Ditch	0.50	346 kg/yr
Little Pine	0.08^{*}	-1829 kg/yr
Little Wea	0.20	-540 kg/yr
Elliot Ditch - Little Pine	0.20	1745 kg/yr
Little Wea - Little Pine	0.94	181 kg/yr
E. coli		
Elliot Ditch	0.60	-25055 CFU/yr
Little Pine	0.06^{*}	-63523 CFU/yr
Little Wea	0.08^{*}	-6068 CFU/yr
Elliot Ditch - Little Pine	0.41	40670 CFU/yr
Little Wea - Little Pine	0.26	40510 CFU/yr

**Statistically significant at α = 0.05

*Statistically significant at α = 0.05

2.3.3 Discharge Data

Two-sample t-test: The first step in examining the results, a time series was generated for both time periods (pre/post), and the mean of each time period (pre/post) was also calculated to examine overall decreases or increases in discharge metrics. For both storm peaks and storm volume, all three watersheds demonstrated a decrease in mean. Indicating that the post periods for LW, ED and LP saw reductions in the total discharge volume per storm, and the peak flow rate per storm. These results are in Figure 40 to Figure 39. The results of the F-test for equal variance did not show any significant results. There were no statistically significant differences in the mean of either the peak discharge rates or the monthly storm volumes, as shown in Table 11.

Table 6.T-test results for discharge analysis.

Watershed	p-value	T-value	Difference in Mean[1] untransformed
Monthly Storm Volume			
LW	0.54	0.61	-232,121,360 l/month
LP.TP	0.20	1.31	-292,753,305 l/month
ED.TP	0.49	0.70	-216,648,467 l/month
Peak Discharge			
LW	0.41	0.84	-1776.43 liters/s
LP	0.69	-0.40	2504.67 liters/s
ED	0.90	-0.13	-3561.52 liters/s

[1]All differences are presented as post-treatment – pre-treatment

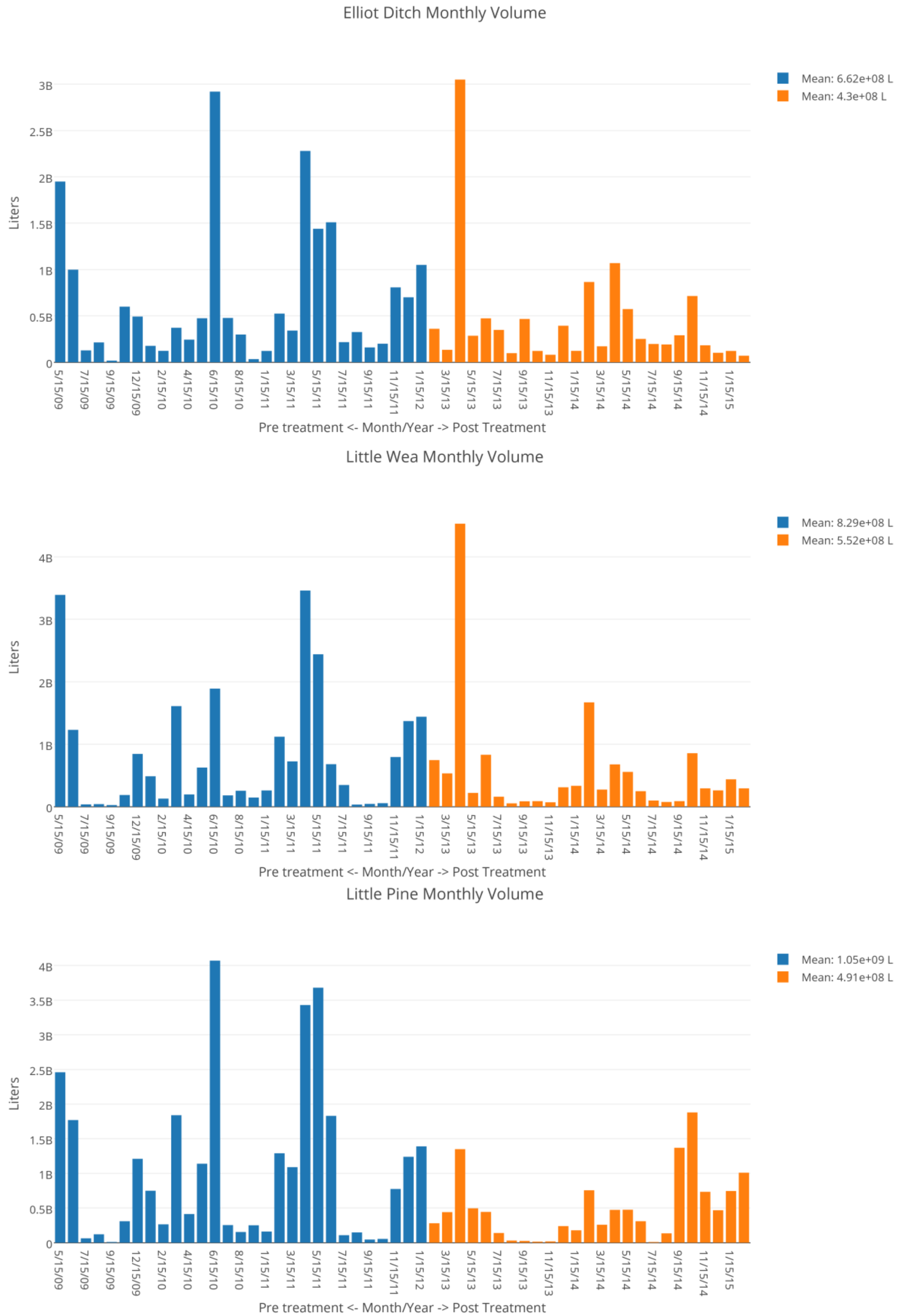


Figure 40. Pre-treatment and post-treatment time series for monthly storm volume Elliot Ditch, Little Wea Creek and Little Pine Creek.

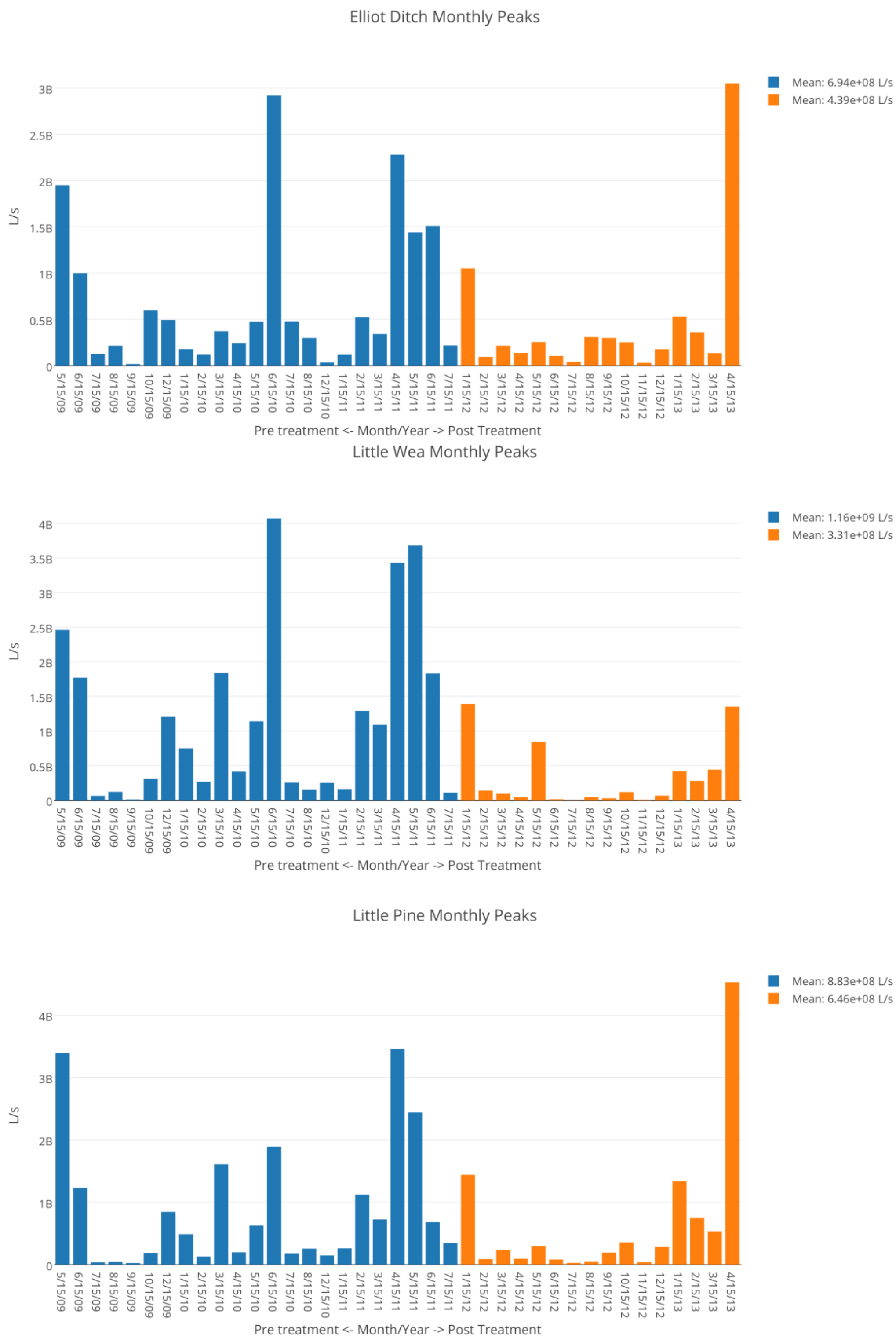


Figure 41. Pre-treatment and post-treatment time series for monthly storm peaks Elliot Ditch, Little Wea Creek and Little Pine Creek.

ANCOVA: ANOVA results for pre/post treatments were significant for both pairs, confirming the presence of a relationship, as shown in Figure 44 to Figure 45. There was no detectable change for peak flows in either watershed pair based on the ANCOVA results in Table 7, but there were statistically significant differences in stormflow volume in both watersheds.

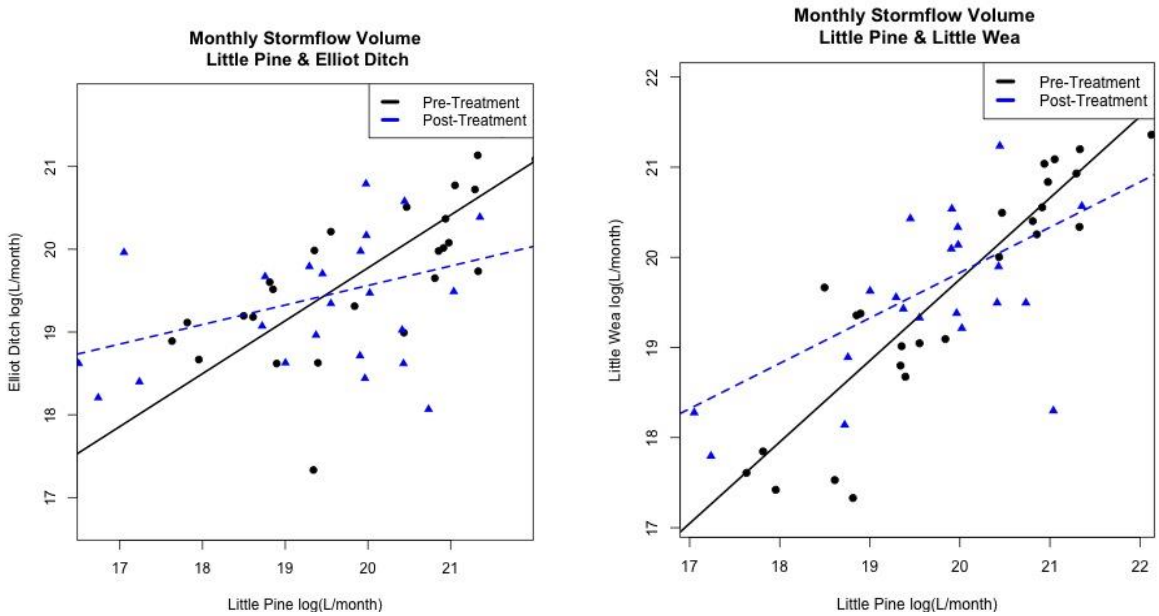


Figure 42. Pre-treatment and post-treatment regression relationships for monthly storm volume between a) Elliot Ditch and Little Pine Creek and b) Little Wea Creek and Little Pine Creek.

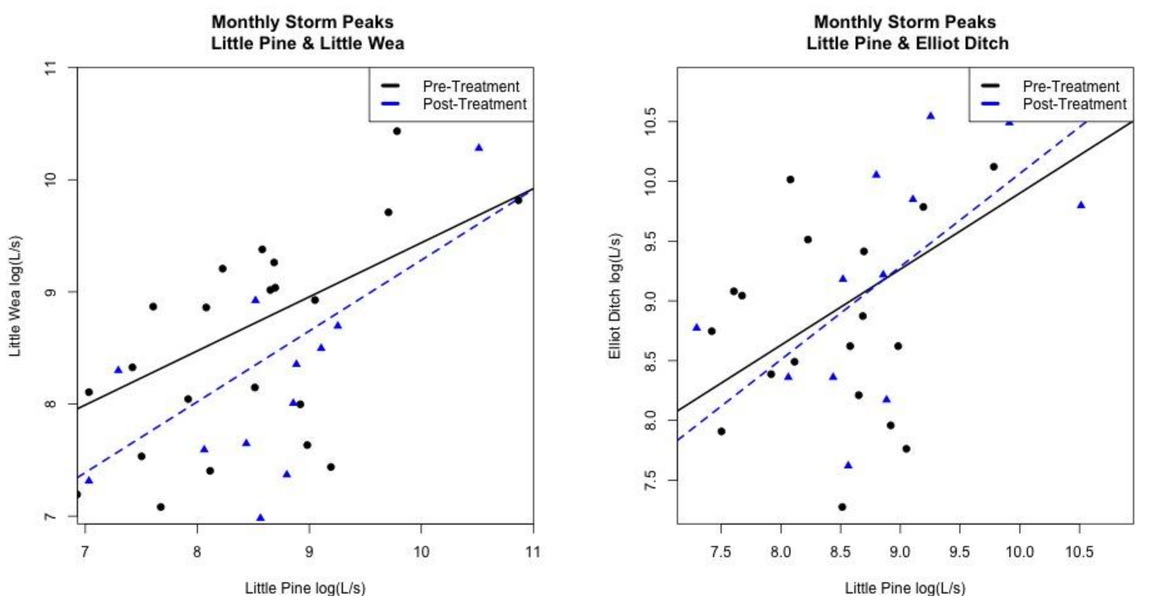


Figure 43. Pre-treatment and post-treatment regression relationships for peak discharge rate between a) Elliot Ditch and Little Pine Creek and b) Little Wea Creek and Little Pine Creek.

Table 7. ANOVA and ANCOVA test results for discharge analysis.

Watershed	ANOVA p-value	ANCOVA p-value	Is there a detectable relationship?	Percent Magnitude of Change[1]
Monthly Stormflow				
Elliot Ditch & Little Pine	Pre: 0.00** Post: 0.05**	Slope: 0.00** Intercept: 0.00**	Significant Relationship	-20.8
Little Wea & Little Pine	Pre: 0.00** Post: 0.00**	Slope: 0.00** Intercept: 0.00**	Significant Relationship	-28.5
Peak Flow rates				
Elliot Ditch & Little Pine	Pre: 0.00** Post: 0.00**	Slope: 0.59 Intercept: 0.60	No Significant Relationship	0.001
Little Wea & Little Pine	Pre: 0.02** Post: 0.02**	Slope: 0.55 Intercept: 0.64	No Significant Relationship	0.049

[1]Calculated by predicting the treatment watershed response for the post-treatment period using the pre-treatment regression equation for the untransformed data and post-treatment observation from the controlled watershed. Difference is 100% *(observed – predicted)/observed, so a positive value indicates an increase in the treatment watershed response.
**Statistically significant at □ = 0.05

<u>Mann-Kendall Test:</u> Time series graphs of the concentration data sets with a smoothed line are shown in Figure 44 and Figure 45. As summarized in Table 8, there were no statistically significant trends in discharge over time.

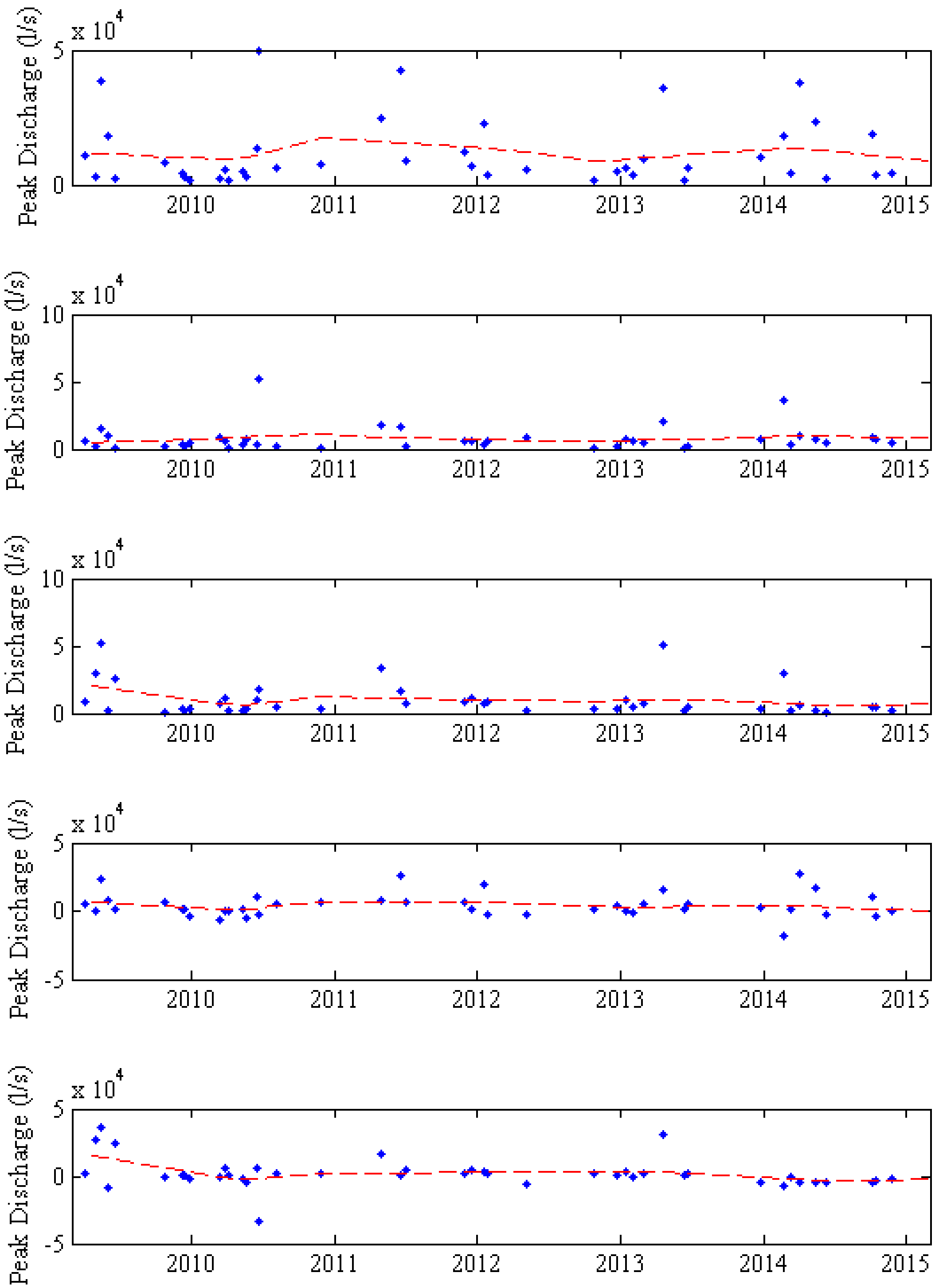


Figure 44. Trend analysis for peak discharge rates in: a) Elliot Ditch, b) Little Pine Creek, c) Little Wea Creek, d) Elliot Ditch – Little Pine Creek and e) Little Wea – Little Pine Creek.

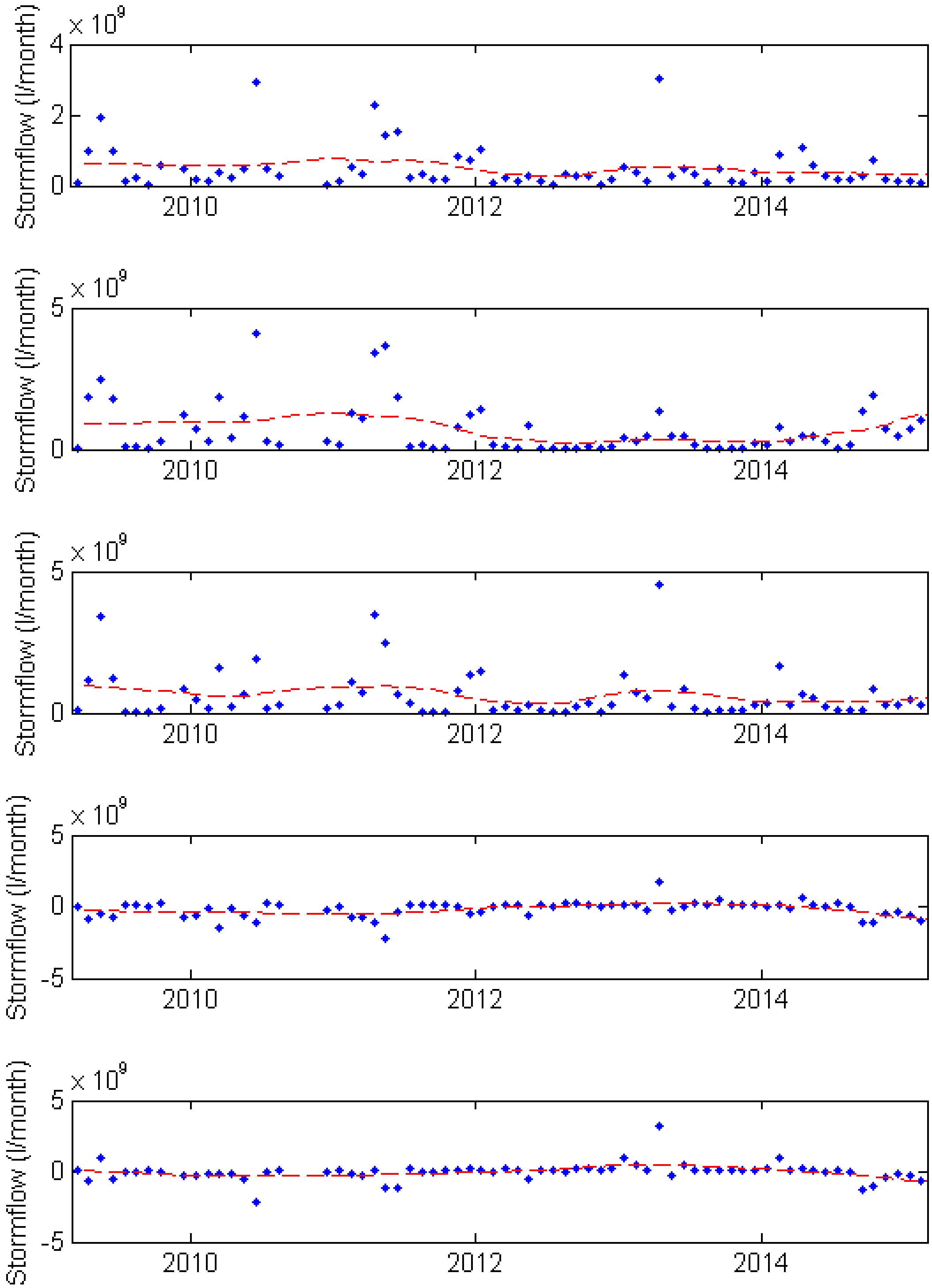


Figure 45. Trend analysis for stormflow volume in: a) Elliot Ditch, b) Little Pine Creek, c) Little Wea Creek, d) Elliot Ditch – Little Pine Creek and e) Little Wea – Little Pine Creek.

Table 8. Seasonal Mann-Kendall trend test results for monthly load.

Watershed & Pollutant	p-value	Trend Slope
Peaks over Threshold		
Elliot Ditch	0.91	-26 l/s/yr
Little Pine	0.34	238 l/s/yr
Little Wea	0.25	-306 l/s/yr
Elliot Ditch - Little Pine	0.24	-435 l/s/yr
Little Wea - Little Pine	0.08	-711 l/s/yr
Storm Runoff Volume		
Elliot Ditch	0.17	-15.5 $x10^6$ l/yr
Little Pine	0.57	-7.7 $x10^6$ l/yr
Little Wea	0.78	-3.4 $x10^6$ l/yr
Elliot Ditch - Little Pine	0.64	-15.5 $x10^6$ l/yr
Little Wea - Little Pine	0.78	+8.9 $x10^6$ l/yr

3.0 <u>Minimal Detectable Difference (MDD)</u>

Figure 46 and Figure 47display the calculated minimum detectable differences under a paired catchment design and a single-site monitoring design for the weekly concentration data, for a range of sampling lengths. Each graph illustrates the relationship between MDD and sample size for the single site analysis (based on the power of the t-test) and the paired site analysis (based on the power of the ANCOVA). A fixed power, or probability of rejecting the null hypothesis of no change when it is false, of 0.8 is used in these calculations. The MDD range on the y-axis is based on the mean of the pre-treatment data in the treatment watershed and varies from 10 to 90% of the mean. All results have a fixed pre-treatment sample size (based on this study's collection data), and are based on a once per week sampling (52 samples per year). The sample size shown on the x-axis is for the number of sampling occurrences post-treatment.

The results for nitrate-nitrogen concentration in Figure 46 and for TSS in Figure 47 show that in both cases the same effect size can be detected with fewer sampling events using the single site approach, although in Little Wea Creek the difference is much smaller. For example, in Elliot Ditch a decrease in concentration of 0.3 mg/L could be detected with approximately 100 sampling events, while it would take almost 300 using the paired catchment approach. This is a surprising result, since when the paired watersheds respond similarly to weather events, the paired approach is better able to detect change because the relationship between the control and treatment watersheds helps to isolate the impact of weather variation from the treatment.

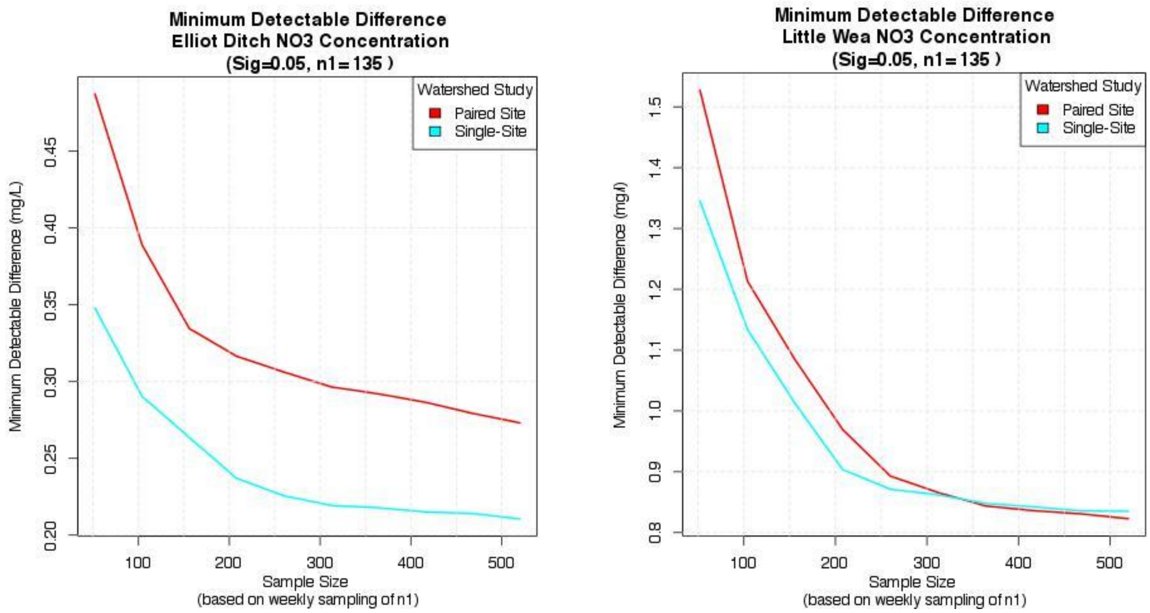


Figure 46. Minimum detectable difference in mean nitrate-nitrogen concentration versus sample size for a) Elliot Ditch and b) Little Wea Creek.

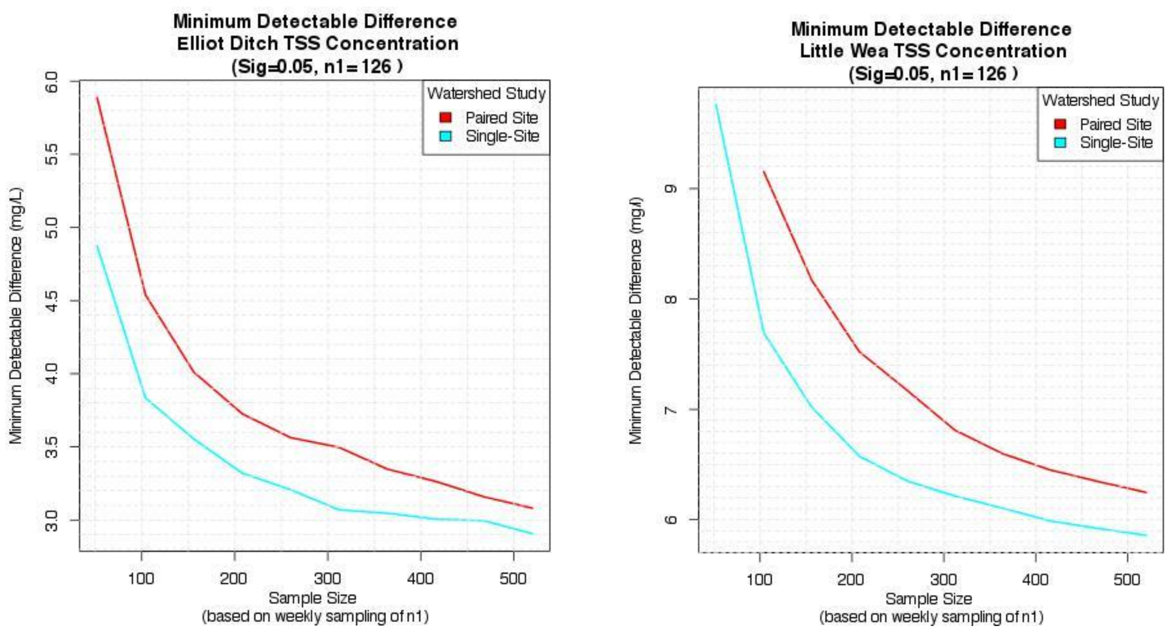
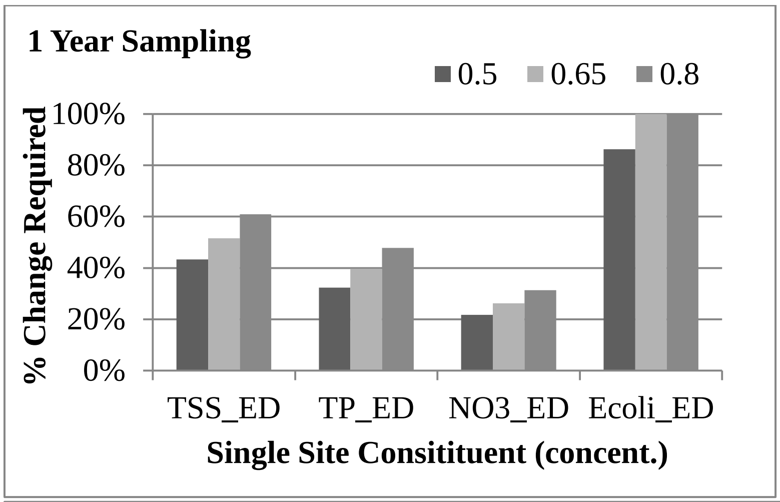


Figure 47. Minimum detectable difference in mean Total Suspended Solids versus sample size for a) Elliot Ditch and b) Little Wea Creek.

To illustrate the results of the power analyses better, results were extracted from the overall curves based on fixed weekly sampling of one and six years duration, for three different power levels of 0.5, 0.65, and 0.8. Once again, a lower power indicates that the statistical test is less likely to detect change when a change has occurred. The analysis shows the percent change that would be statistically significant based on the defined power levels. Figure 48a) illustrates that a higher likelihood of correctly detecting a change that has occurred, means that the change is going to have to be larger. This pattern is true regardless of length of sampling. However, when the sampling is increased to six years in Elliot Ditch (Figure 48b), it does reduce the percent change needed. In both cases, E.coli would require the largest change amongst constituents, with a 1-year sampling scheme requiring over 100% change in concentration for it to be reliaby detected (power > 0.5).

As shown in Figure 49, concentration results for Little Wea Creek follow a similar pattern as Elliot Ditch. For a three-year sampling duration, with a 50% probability of correctly detecting change the single site analysis requires fewer samples. Unlike Elliot Ditch however, E.coli and NO_3 require a smaller percent change than TP and TSS.



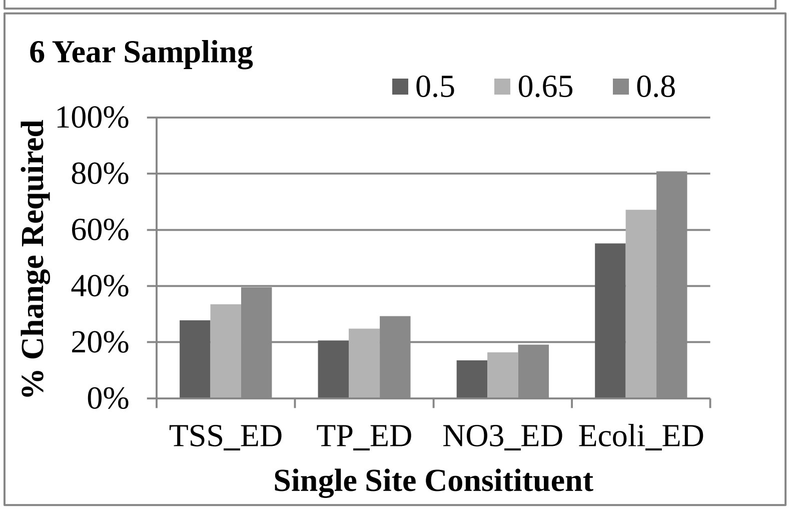


Figure 48. The percent change in concentration for statistical significance in Elliot Ditch given (top) one year of weekly sampling and (bottom) six years of weekly sampling for different levels of power. Increasing power represents the increased probability of identifying change that has occurred. Detecting NO_3, and TP would require the least amount of change under a one-year sampling schedule.

Figure 49. Single-site versus paired-site percent required changes under a three-year sampling design at a power of 0.5 for a) Elliot Ditch and b) Little Wea concentrations.

Figure 50 and Figure 51 show similar MDD curves for nitrate-nitrogen load and TSS load, respectively. The MDD curves for our two discharge metrics, POT and storm volume, are shown in Figure 52 and Figure 53. In contrast to the MDD results for concentration, where the single-site approach detected smaller changes, monthly load, volume and peaks over threshold results all detected smaller changes using a paired design. The cases where the single site line is not visible on the graphs indicates that a change could not be detected within the range of sample sizes tested. In the case of TSS load for Little Wea, more than a 100% change would have to occur to be detected with the single-site method. Another way of interpreting this particular result is to conclude that a single-site method for TSS load would never detect a reduction in load. In Elliot Ditch, the MDD is much smaller overall due to the lower mean and variance of nitrate in Elliot Ditch. Similarly, the paired approach is better able to detect differences in total phosphorus and total suspended solids concentrations in both Elliot Ditch and Little Wea Creek.

The MDD curves highlight the importance of having good catchment pairs for the paired catchment analysis. These watersheds are close enough together that they experience similar weather events, and therefore streamflow is highly correlated, even in watersheds with differing land use. Concentration tends to be less correlated, particularly between Little Pine Creek and Elliot Ditch that have contrasting land use. By comparing monitoring design scenarios given a constant three years of sampling for both watersheds, we found that larger percent changes are necessary under a paired-site design versus a single-site design when the objective was measuring change in concentration (Figure 49). For example, if one were interested in measuring TP concentration change for a 3-year sampling period under a single site design in Elliot Ditch, an 25% change would need to occur, but a paired-site design would need approximately a 35% change in concentration. The opposite is true when measuring discharge, so that the paired-site design will show detectable change sooner than a single site design; Figure 54 shows these data for Elliot Ditch watershed. Percent change requirements for peaks-over-threshold were all greater than 40% and therefore not shown.

So can discharge be used to measure change? Well, that depends on the constituent and land use amongst other factors. Must urban stormwater BMPs have little influence on concentration, so load reductions are expected to come primarily from discharge reduction. Looking at Elliot Ditch with a 3-year sampling period and a power of 0.5, we see that the MDD for the storm volume (30-35%) is lower than that of E.coli (~65%) and TSS (35-40%) using either single or paired designs (Figure 54). In the case of NO_3 however, a less important constituent in the urbanizing watershed, a single site design would detect concentration changes of 18% (Figure 50), so a measurable percent change will be visible in concentration sooner than discharge.

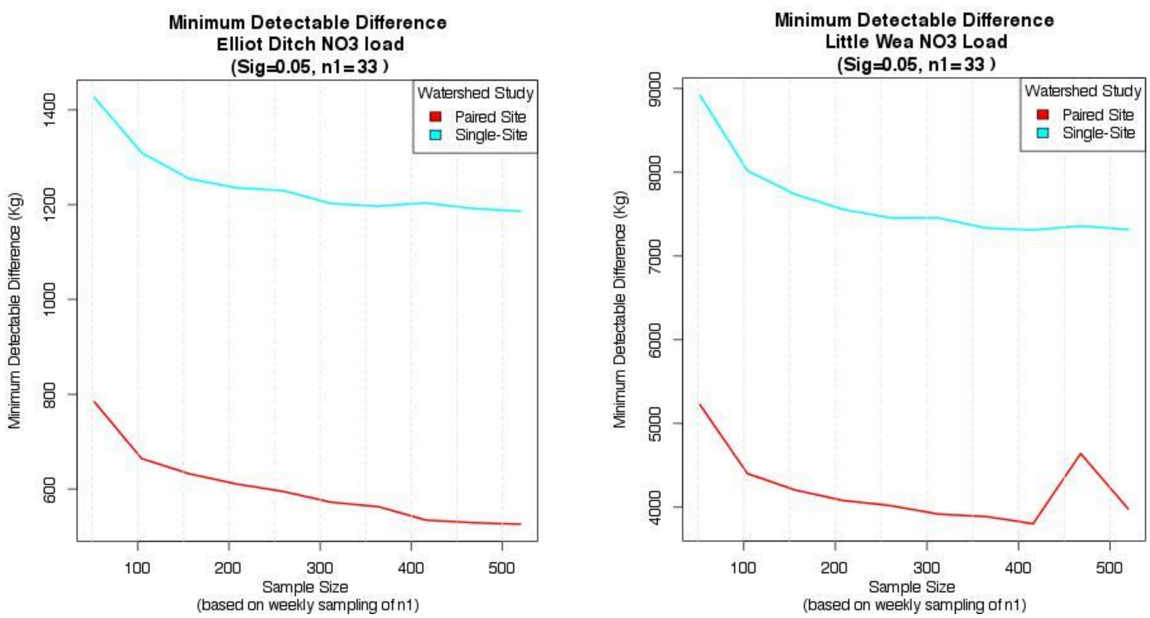


Figure 50. Minimum detectable difference in mean Nitrate Load versus sample size for a) Elliot Ditch and b) Little Wea Creek.

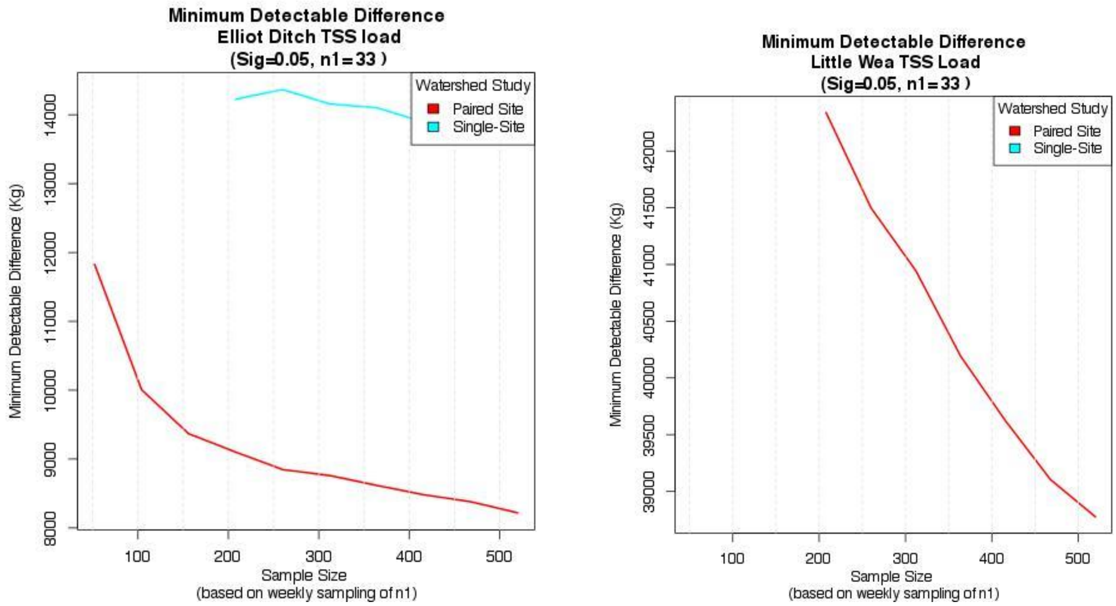


Figure 51. Minimum detectable difference in mean TSS Load versus sample size for a) Elliot Ditch and b) Little Wea Creek.

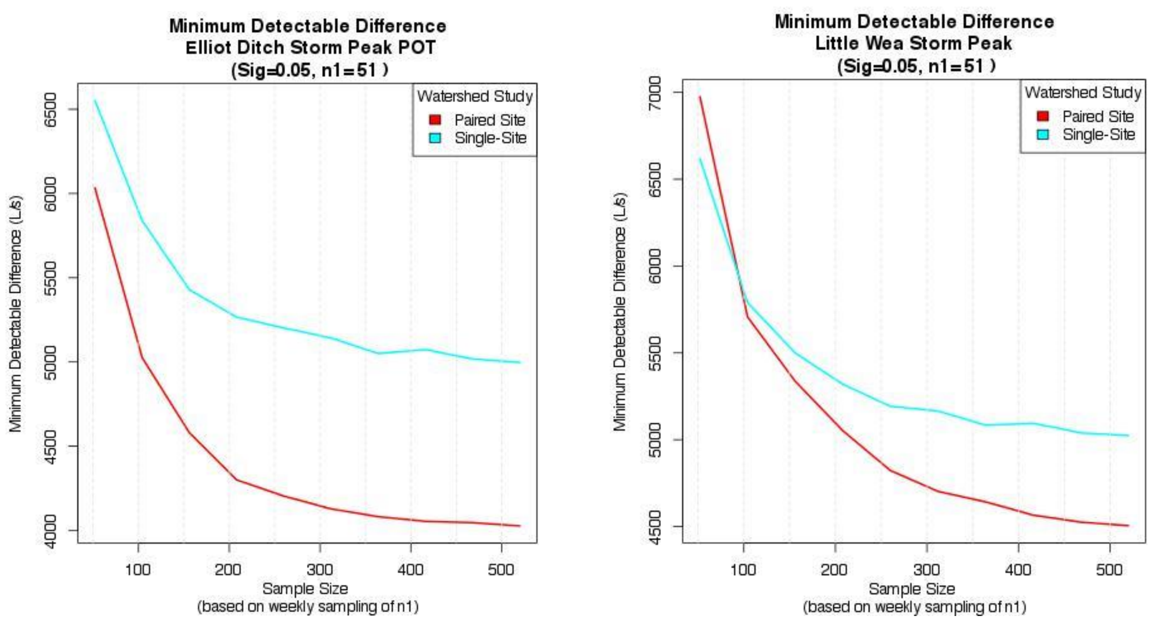


Figure 52. Minimum detectable difference in mean storm peaks versus sample size for a) Elliot Ditch and b) Little Wea Creek.

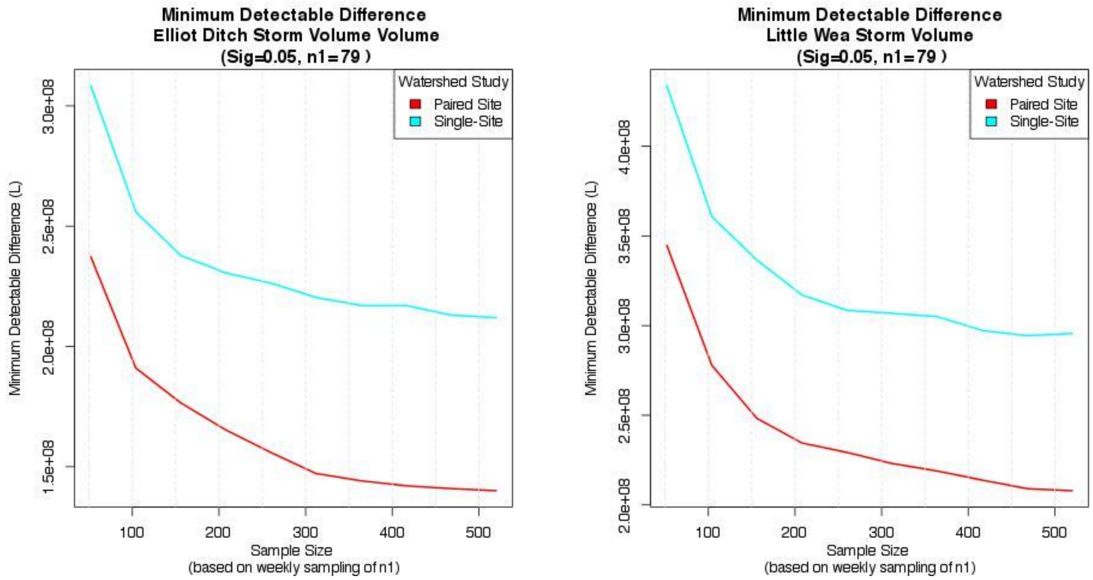


Figure 53. Minimum detectable difference in mean storm volume versus sample size for a) Elliot Ditch and b) Little Wea Creek.

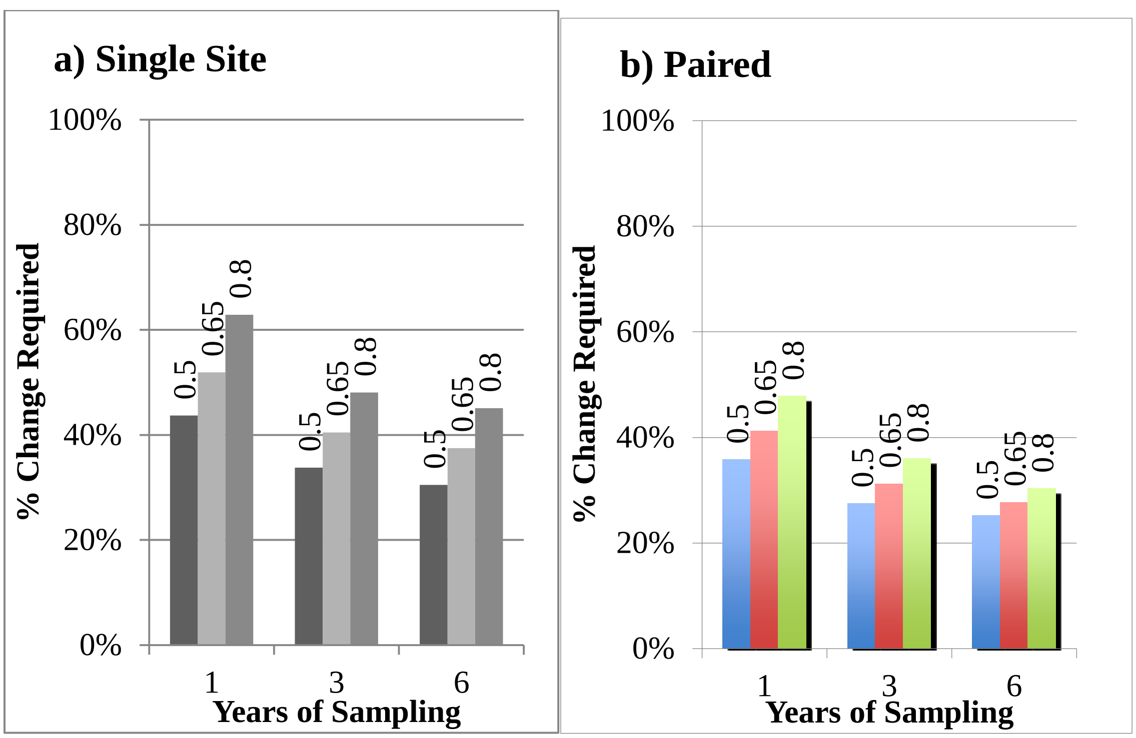


Figure 54. Percent required change in discharge for powers of 0.5, 0.65, and 0.8 in Elliot Ditch under a) single site analysis and b) paired-site design.

4.0 <u>Summary</u>

Yes! We have seen changes. The single site analysis has detected statistically significant changes in nitrate, E. coli , total suspended soils and total phosphorus concentrations for different watersheds. We have also detected statistically significant changes in nitrate load. Our paired watershed analysis allows us to conclude whether these changes represent true differences between the treatment and control watersheds, or changes that may be happening in all watersheds due to weather and other factors. These results are summarized below.

4.1 Discharge

Looking at each watershed in isolation (single site analysis) both the t-test and the Mann-Kendall test indicate that storm volume decreased in all three watersheds. The peaks-over-threshold values decreased in Elliot Ditch and Little Wea but not Little Pine. The paired-catchment analysis helps to isolate if these results are due to weather alone. The ANCOVA analysis showed a small increase in POT, while the Mann-Kendall difference series shows that POT decreased in both Elliot Ditch and Little Wea Creek relative to Little Pine Creek. ANCOVA shows that storm volumes decreased significantly in Elliot Ditch and Little Wea when we control for the response in Little Pine, while the Mann-Kendall test shows an increase in Little Wea relative to Little Pine and a slight decrease in Elliot Ditch. Based on this we cannot conclude anything regarding changes in storm peaks in Elliot Ditch and Little Wea Creek. There is evidence to suggest that the observed decreases in storm runoff volume are greater in Elliot Ditch and Little Wea Creek than those observed in Little Pine Creek and this result is statistically significant.

4.2 Nitrate-Nitrogen

Little Pine Creek: Both the t-test and the Mann-Kendall trend analysis indicate that nitrate concentration has increased in Little Pine Creek, while the load has decreased. The decrease in load is statistically significant according to the t-test. This difference between direction of change for concentration and load is consistent with the decrease in storm runoff volume observed in Little Pine Creek, suggesting the decrease in discharge compensates for the increase in concentration.

Elliot Ditch: Both the t-test and the Mann-Kendall trend analysis indicate that nitrate concentration and load has decreased in Elliot Ditch. The decrease in load is statistically significant.
ANCOVA shows an increase, while the Mann-Kendall difference shows nitrate concentration in Elliot Ditch decreased relative to expectations. However the nitrate load decreased relative to expectations according to the ANCOVA and increased according to the Mann-Kendall test. The ANCOVA was statistically significant. This suggests that there is a small but statistically significant decrease in nitrate load in the treatment watershed that is greater than what is expected given the response in the control catchment. This is likely due to the combined effect of decreasing concentrations and decreasing storm volumes.

Little Wea Creek: Both the t-test and the Mann-Kendall trend analysis indicate that nitrate concentration and load has decreased in Little Wea Creek. The decrease in concentration is statistically significant for both the t-test and the Mann-Kendall trend test and the decrease in load is significant according to the trend test.

ANCOVA shows nitrate concentration increased in Little Wea Creek relative to Little Pine Creek (statistically significant), while the Mann-Kendall difference tests indicates that it decreased relative to Little Pine Creek (not significant). In contrast nitrate load in Little Wea Creek decreased relative to Little Pine according to both the ANCOVA and the Mann-Kendall difference test. We conclude that nitrate concentrations and load have decreased in Little Wea Creek. We cannot yet say definitively if this is due to treatment in the watershed, relative to overall regional changes.

4.3 Total Phosphorus

Little Pine Creek: Both the t-test and the Mann-Kendall trend analysis indicate that total phosphorus concentration has increased in Little Pine Creek (not significantly). The load has also increased.
Elliot Ditch: Both the t-test and the Mann-Kendall trend analysis indicate that total phosphorus concentration increased in Elliot Ditch. The t-test indicates that load has decreased, while the Mann-Kendall test indicates that load has increased. None are statistically significant, suggesting that the true change in load is near zero.

ANCOVA and the Mann-Kendall difference shows Total Phosphorus concentration in Elliot Ditch decreased slightly (ANCOVA) or increased slightly (Mann-Kendall) relative to expectations. The ANCOVA and Mann-Kendall indicates that TP load decreased relative to expectations. Based on these mixed results, with no statistical significance, we conclude that there is no detectable change in total phosphorus concentration due to treatment in Elliot Ditch, but there are indications of a decrease in Total Phosphorus load consistent with a decrease in stormflow volume

Little Wea Creek: Both the t-test and the Mann-Kendall trend analysis indicate that total phosphorus concentration increased in Little Wea (t-test was statistically significant). The t-test suggests that load has decreased, while Mann-Kendall shows an increasing trend.

The ANCOVA analysis shows that TP concentration decreased in Little Wea relative to Little Pine, as did the load (statistically significant). The Mann-Kendall difference test for trend shows decreasing trends in both TP concentration and load in Little Wea relative the values in Little Pine. We can conclude that although TP has increased in Little Wea Creek, the increase is less than what we would expect based on what was observed in the control catchment, meaning there is a statistically significant reduction in TP concentration and load due to treatment effect.

4.4 Total Suspended Solids

Little Pine Creek: The t-test indicates a decrease in TSS concentration and load in Little Pine Creek while the Mann-Kendall test indicates an increase in concentration but not load. None are statistically significant so we concluded that there is no evidence of change in TSS in Little Pine Creek.

Elliot Ditch: Both the T-test and Mann-Kendall indicate that TSS concentration increased (significantly for the t-test) in Elliot Ditch. In contrast the t-test shows a decrease in load while the Mann-Kendall test indicates an increasing trend in load. Both ANCOVA and Mann-Kendall difference show that TSS load increased in Elliot Ditch relative to Little Pine Creek (the ANCOVA is marginally significant), but the ANCOVA reported a decrease in concentration (concentration is statistically significant). Therefore, there seems to be some evidence that TSS has increased in Elliot Ditch, but results are currently conflicted.

Little Wea Creek: Both the T-test and Mann-Kendall indicate that TSS concentration and load decreased in Little Wea Creek. ANCOVA shows that TSS concentration increased in Little Wea relative to Little Pine, while the TSS load decreased. In contrast, the Mann-Kendall difference shows that TSS concentration decreased in Little Wea relative to Little Pine Creek, while the TSS load increased. We conclude that although TSS is decreasing in Little Wea Creek the amount of decline is consistent with the changes seen in the control watershed.

4.5 <u>E. coli</u>

<u>Little Pine Creek:</u> Both the t-test and Mann-Kendall indicates a decrease in e. coli concentration and load in Little Pine Creek. None are statistically significant so we concluded that there is no evidence of change in E. coli in Little Pine Creek.

<u>Elliot Ditch:</u> The t-test indicates a decrease in e. coli concentration in Elliot Ditch, but the Mann-Kendall test indicates an increasing trend. Both the t-test and Mann-Kendall indicate a decrease in E. coli load. Both the ANCOVA analysis and Mann-Kendall difference indicate that E. coli concentration increased relative to Little Pine Creek, but only the Mann-Kendall showed an increase in load as well. We conclude that although e. coli is decreasing in Elliot Ditch, the decrease is consistent with what would be expected given the decrease observed in Little Pine Creek.

<u>Little Wea Creek:</u> Both the t-test and Mann-Kendall indicate a decrease in e. coli concentration and load in Little Wea Creek. ANCOVA shows that E. coli concentration and load decreased in Little Wea Creek relative to Little Pine Creek (statistically significant). The Mann-Kendall difference test indicates that both concentration and load increased relative to Little Pine Creek. We conclude that E. coli is decreasing in Little Wea Creek, and the decrease is more than what would be expected given the decrease observed in Little Pine Creek.

5.0 <u>What We Learned/Transference to Other Watershed Groups</u>

We have analyzed the extensive data collected in this project in multiple ways in order to analyze not only if change has occurred in our treatment sub-watersheds, but also which methods of analysis may prove to be most useful. Overall, we have found the following:

- Our sampling design has proven successful. With this level of water quality sampling intensity and the inclusion of USGS stream gauges, we were able to detect statistically significant changes in the treatment watersheds that are greater than the changes expected in the control watershed. Statistically significant changes were detected in nitrate-nitrogen, TSS and *E.coli* concentration, nitrate, TP and E. coli load and stormflow volumes. So the sampling design is successful in detecting change and should be given consideration for wider use.
- If the primary concern is concentration, then using a single-site design will likely detect changes sooner in the watershed than under a paired-site design. However, the change detected may not be due to the treatment using a single-site design, but using a paired-design requires a strong correlated pair.
- When determining a 'good' pair, similar land use type is an important factor to pairing for concentration as is evidenced by Figure 46. In this case, the difference in MDD between monitoring designs in Little Wea are smaller in comparison to Elliot Ditch. This relationship can also be seen in the ANCOVA results where a changes in nitrate were inconclusive for Elliot Ditch but significant for Little Wea. Results were inconclusive in Elliot Ditch due to a weaker correlation to the control. A good pair requires not only good correlation between observed values (concentration, discharge and load) between the two watersheds, but also similar distribution of the data (that is the data variance is similar between watersheds).
- If primary concern is load, discharge, or peaks-over-threshold, a paired-site design will likely detect changes in the watershed sooner than a single-site design. Streamflow response tends to be more correlated between regional watersheds, so neighboring watersheds are more likely to form good pairs.
- Our analysis of minimum detectable differences indicates that with the current sampling frequency and design, that if the implemented practices resulted in our targeted level of change (e.g. reduction of nitrate-nitrogen concentrations in Little Wea Creek to below 2 mg/L

we would have been able to detect that level of change. Using a more typical monthly sampling scheme during the pre-treatment period would result in the need for a significantly longer post-treatment sampling period to detect change (not shown).

- Not only is the total number of practices implemented – and their spatial extent – an issue in enacting change, but the location of the practices will also have an impact on observed changes in water quality at the watershed outlet.
- Overall, the paired watershed approach has the potential to 1) increase the ability to detect statistically significant change and 2) reduce the effect of weather variation on the ability to observed differences. The MDD analysis shows that in many cases even though the paired watershed approaches requires double the number of analyses per sampling event, the overall number of analysis needed to detect change may still be less, and the time period of sampling needed is definitely reduced. This result does depend on whether or not the identified treatment and control watersheds make a good pair.
- If a confirmation of change without attribution is the primary goal then monitoring before and after a targeted implementation period is all that is required.
- Trend analysis is appropriate for analyzing the results of a protracted, multi-year, implementation period. In this case monitoring should extend continuously from before the treatment and throughout implementation.
- Seasonal trend analysis is good approach for detecting water quality changes given the fact that most implementation is not done in a single year, but rather continues to build over a period of years. This results in a gradual change in water quality over time rather than an abrupt shift between arbitrarily defined pre-treatment and post-treatment periods. However, a monitoring period of at least ten years is generally recommended for trend analysis (extending from initial baseline monitoring through implementation). The sampling frequency could potentially be reduced, if adequate samples are taken to characterize the mean concentration or load during each season, where seasons can be defined as monthly, 3, 4, or 6 month intervals depending on the characteristics of the watersheds.

5.1 <u>Comparison on Minimal Detectable Change to the Number of Collected Samples</u>

In a few cases, the analysis presented above has resulted in statistically significant differences using the two sample t-test, while the MDD analysis shows that they should not be detectable. For example, there were 89 post-treatment samples in our cleaned data and statistically significant differences in the mean concentration of 0.28 mg/L and 0.67 mg/L were detected in Elliot Ditch and Little Wea Creek, respectively. There are a couple of explanations. First, data that does not fit the assumptions of the parametric statistical tests such as non-normality and dependence are likely to result in false positives (rejection of the null hypothesis when it is true). Although we made many efforts to prepare the data appropriately, there could still be some effect. Since the data used in the MDD was randomly generated, it perfectly fits the assumptions of the statistical tests. Secondly, the MDD analysis assumed equal variance between the pre and post-treatment samples, while the t-test used the observed difference in variance when the F-test was statistically significant.

5.2 <u>Is Concentration or Load the Answer?</u>

Both concentration and loading analysis provide benefits to understanding water quality within the Region of the Great Bend of the Wabash River Watershed. Statistical analyses that occurred as part of this report indicate that concentration measurements may be more important in agricultural watersheds, such as Little Wea Creek, while discharge and loading measurements may be more important in urban watersheds, such as Elliot Ditch. Agricultural best management practices focus on stabilizing in-field conditions which often target specific nutrient and sediment reductions – holding soils in place, filtering field runoff and tile drainage and more. Urban best management practices focus on infiltrating stormwater – sending larger volumes of water and the pollutants the water carries, to groundwater sources rather than it entering adjacent waterbodies as surface stormwater. This suggests that the collection of grab samples is especially important to characterize your watershed and that continued grab sample collection will provide additional, useful data if it is collected over a sufficient period of time to measure a change in water quality especially in agricultural watersheds. The addition of discharge may be especially important within urban watersheds where infiltration or peak discharge may be a way to identify a measurable change in water quality. Additional statistical analyses are required to assess whether discharge is sufficient on its own to identify these changes in water quality or if grab sampling is an equally or more important component of standard water quality monitoring programs.

It is also important to note that the preferred monitoring method is also dependent on the goals of the coordinator or monitoring team. The single-site method was best at detecting smaller changes in the watershed when concentration was the focus. However, it is also true that attributing a change in concentration due to the practice(s) and not changes in weather is impossible to do. On the other hand, load and discharge detected smaller changes in the watershed under a paired method, which also enables the removal of changes in weather as a contributing factor to changes in the watershed. As a final caveat, the results of the statistical analyses proved to be more conclusive and/or significant under the paired method opposed to the single-site method in this study.