OCC Paired Watershed Project: Beaty Creek

In August 1999, the Oklahoma Conservation Commission Water Quality Division began a project to demonstrate the effectiveness of best management practices in water quality improvements in Beaty Creek. The primary objective of the project was to reduce nonpoint source loading of principal pollutants, most importantly phosphorus, contributing to excessive algae growth and associated taste and odor problems in downstream Lake Eucha, a major water supply for the City of Tulsa. Best management practices (BMPs) including riparian management/improvement, pasture planting and nutrient management, offsite watering, and construction of heavy use areas for animal feeding were implemented throughout the project duration with the majority of implementation occurring by the fourth full year (Table 1).

The Beaty Creek Project was undertaken per EPA guidelines utilizing a paired watershed approach (EPA 841-F-93-009, 1993). The basic method requires a minimum of two watersheds, a control and a treatment, and two definable periods of study, calibration and treatment. The control watershed is chosen to account for environmental variability over the periods of study throughout the short duration that may otherwise mask the overall effect of BMPs on NPS pollutant loads in the treatment watershed. Therefore, it is necessary to choose a control watershed that will experience the same weather and seasonally induced changes as the treatment watershed. For this study, the Little Saline Creek watershed was chosen as the control as it is proximal to Beaty Creek and would be expected to experience the same climatic and other environmental impacts (Figure 1). Because its purpose is to account for natural variability, the control watershed must not incur any major landuse changes during the course of the study period.

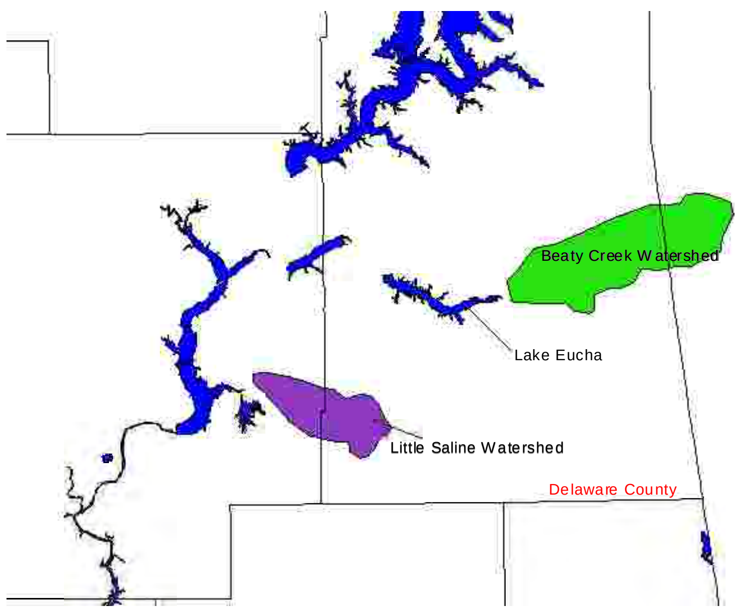

Figure 1. Beaty and Little Saline watersheds, Delaware County, OK.

Table 1. Best management practices and associated costs implemented over project duration.

	A	B	C	D	E	F	G	H
1	**BMP**	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**Totals**
2	Riparian area managemer	$3,695.00	$5,295.00	$9,640.00	$11,446.00	$14,104.08		$44,180.08
3	Well Drilling	$4,199.40	$8,674.90	$3,144.15	$4,599.74	$25,500.37		$46,118.56
4	Freeze-proof tanks		$3,335.45	$2,160.00	$1,316.92	$13,470.31	$640.00	$20,922.68
5	Litter Clean-Out structures		$12,000.00	$4,935.00	$7,609.22	$9,000.00		$33,544.22
6	Pasture planting	$30,173.98	$13,003.86	$13,285.90	$34,373.72	$27,420.48	$2,215.38	$120,473.32
7	Pasture Nutrient Managem	$22,101.80	$23,362.70	$28,078.00	$31,206.50	$29,060.50		$133,809.50
8	Fence	$11,508.65	$14,595.58	$26,104.23	$18,893.48	$82,481.19	$48,746.32	$202,329.45
9	Ponds	$8,026.90	$1,811.20	$6,495.74	$15,328.95	$3,706.50		$35,369.29
10	Freeze-proof tanks/pipelin	$8,031.07	$15,863.15	$9,618.94	$16,127.55	$67,415.39	$6,276.29	$123,332.39
11	Cow shade		$4,200.00					$4,200.00
12	Poultry Waste Utilization	$9,677.84	$13,212.77	$9,496.87	$7,844.32	$8,381.09		$48,612.89
13	Heavy use areas	$12,096.22	$18,150.47	$13,819.64	$24,741.38	$32,663.28	$5,335.27	$106,806.26
14	Waste Storage/Feeding Facility			$6,300.00	$6,300.00	$100,800.00	$22,680.00	$136,080.00
15	Rural Waste Mgt Systems		$5,912.00	$7,578.40	$5,390.40	$13,639.85	$11,260.00	$43,780.65
16								
17	**Grand Totals**	$109,510.86	$139,417.08	$140,656.87	$185,178.18	$427,643.04	$97,153.26	$1,099,559.29

The foundation of the paired watershed approach is that there is a quantifiable relationship between the watersheds for a parameter of interest, and that this relationship remains valid until major changes (i.e., BMPs) occur in one or the other (EPA 841-F-93-009, 1993). Following these changes, a new relationship will exist and have to be determined. It is the comparison of these relationships to one another that allows the determination of land management effects, slight though they may be due to short project duration, on NPS pollutant loads. It is necessary to note that the difference in quality of runoff between the control and treatment watersheds is not the issue, but rather that the relationship between paired observations between the two remain the same through time, except for the effects of the BMPs (EPA 841-F-93-009, 1993). Thus, if litter is spread in one watershed contributing to higher levels of phosphorus than in the other, it has no bearing on the paired watershed approach. Differences in water quality between the two watersheds are expected, but it is the predictable response of the two watersheds together that is the foundation of the paired watershed method.

To monitor pollutant loads through the systems, automated samplers were placed in both watersheds to allow continuous, flow weighted sampling. Specifically, samplers were programmed to pull samples based upon rate of water passage (e.g., sample pulled for every 10,000 cubic feet of water). Thus, during periods of runoff, sampling frequency was heavier than during seasonal base flows. The integration of such sampling over a period of time results in more accurate estimates of pollutant loads (i.e., simply how much of a certain item of interest is being transported via the system) than single weekly or monthly grab samples.

Upon completion of the project period, data were compiled, collated into calibration and treatment periods, and analyzed per EPA Paired Watershed Study protocol (EPA 841-F-93-009, 1993). Because there was at least some implementation throughout the study period, the first year and last years of the project were chosen for calibration and treatment periods, respectively. Weekly total phosphorus (T-P) loads were determined by multiplying T-P concentrations from weekly, integrated samples by the total flow for the sampling period. The first step in the analysis was to determine the relationship, if any, between the watersheds for both the calibration and treatment phases. To meet assumptions necessary to implement certain statistical methods, weekly T-P loads were converted to log base ten values before analysis. These log T-P load values were paired between the watersheds by date of collection and analyzed by linear regression to determine relationship. Figure 2 indicates strong, statistically significant ($P<0.001$) linear relationships between the two watersheds for both the calibration and treatment periods. Dotted bands are the upper and lower bounds of the interval within which the regression would be expected to occur 95 times if the project were repeated 100 times. Since both regression lines fall outside the 95% CI bounds (i.e., the dotted bands) of each other, the relationships are said to be significantly different.

After determining that the relationship between the watersheds is significant for both periods, it is necessary to determine what level of change in T-P load between the calibration and treatment periods the sampling effort is sufficient to detect. Discussion of the exact procedure is too involved to include in this summary, but the method involves computing the ratio of the residual variance for the treatment regression to the percent difference expected (e.g., 20% reduction in T-P load between calibration and treatment periods). The results of this analysis show that sampling effort during the calibration period is sufficient enough to allow detection of at least a 12.75 percent change in weekly T-P load between the periods. Thus, to detect a lesser change like 10%, it would have been necessary to increase sampling effort to 88 samples for both creeks.

To determine the effect of the BMPs on weekly T-P load in Beaty Ck, it was necessary to employ a statistical tool called analysis of covariance (ANCOVA). This powerful tool allows

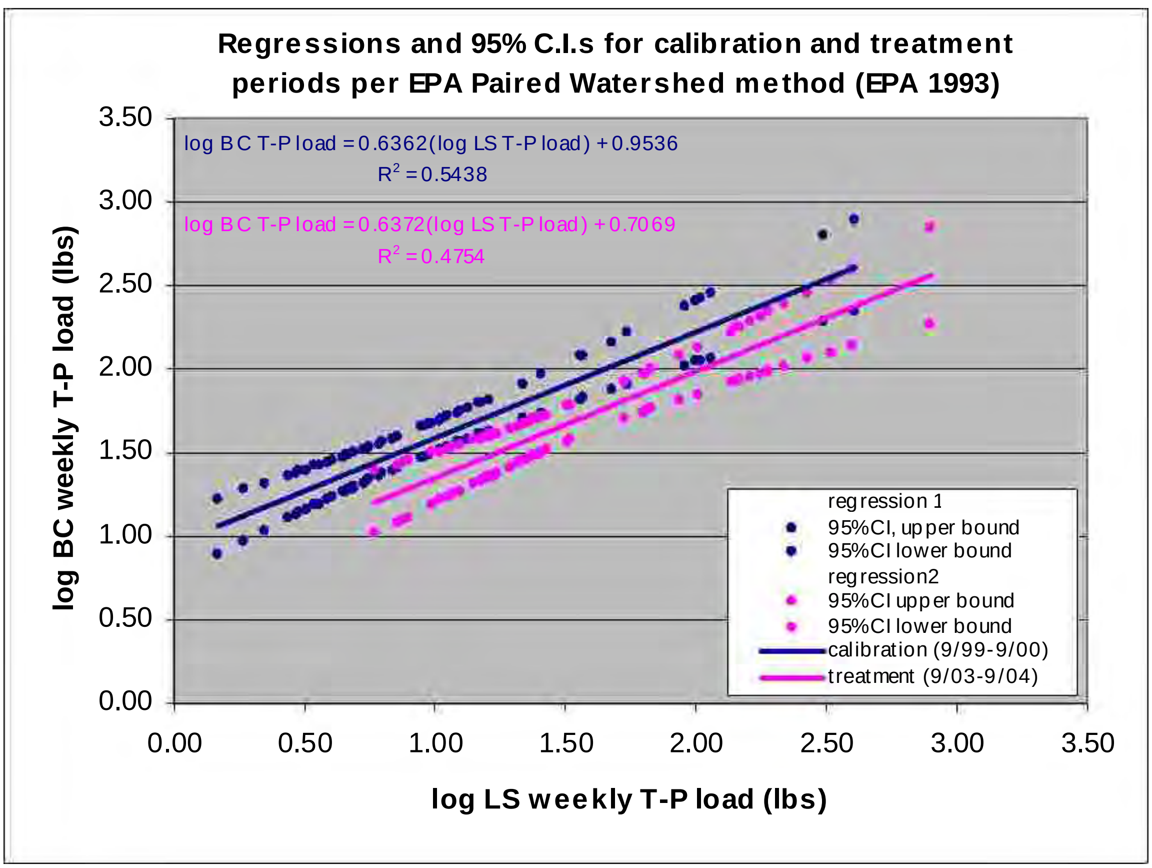

Figure 2. Relationships of log transformed Beaty T-P load to Little Saline T-P load for calibration and treatment periods. Both regressions are significant at the P<0.001 level.

the determination of difference between the calibration and treatment periods despite whatever difference might have occurred because of environmental variability (e.g., wet year vs. dry year) or other factors as accounted for by the Little Saline data. The statistical software package Minitab, V. 14 was employed to conduct the analysis. The results of the ANCOVA analysis are included in Table 2 (below).

The ANCOVA results show that both log Little Saline T-P load and calibration/treatment period are strongly related to Beaty T-P. The items in the table of most significance are the P-values, 0.000 and 0.002, which convey the statistical significance of the relationships to log Beaty T-P load of Little Saline T-P and study period, respectively. Specifically, the P-value of 0.002 indicates that there is strong evidence of a difference between the calibration and treatment periods, even after adjusting for difference due to other things as accounted for in the Little Saline data. The P-value associated with "log LS T-P load" shows that the Little Saline T-P data is related to the Beaty T-P data quite significantly for both periods combined. A test of difference in slopes and intercepts per EPA method (EPA 841-F-93-009, 1993) show no difference in slopes between the calibration and treatment regressions but a highly significant difference in intercepts ($P<0.005$), corroborating the overall parallel shift in regression as seen in Figure 2.

To aid in visualizing any change in log weekly T-P load between the calibration and treatment periods, a plot of the difference between weekly Beaty T-P loads observed during the treatment period and those predicted by the calibration equation was constructed (Figure 3). Since the calibration period regression represents the relationship between the two watersheds before any significant BMP implementation, input of treatment period log weekly Little Saline T-P loads into the equation will result in log weekly T-P loads for Beaty Creek that would be expected under the same circumstances. Thus, subtraction of these "predicted" loads from the actual loads seen during the treatment period would result in determination of a change, if any, from what would be expected given no BMPs. Theoretically, if the relationship is adequate, most of the differences should be slight to none if BMPs have had no effect on T-P loads (i.e., T-P loads during the treatment shouldn't be that different from those during the calibration). In this case, it is obvious by the shift in the regression line (Figure 2) that log weekly T-P loads on average are lower during the treatment period when adjusted for differences due to environmental variability and other factors accounted for in the Little Saline data. Thus, most of the observed weekly

Table 2. Minitab results of the ANCOVA for the combined calibration and treatment T-P data (log transformed) for Beaty and Little Saline Creeks.

Factor	Type	Levels	Values
Period	fixed	2	calibration, treatment

Analysis of Variance for log BC T-P load, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
log LS T-P load	1	11.2189	12.1397	12.1397	103.83	0.000
Period	1	1.2258	1.2258	1.2258	10.48	0.002
Error	100	11.6916	11.6916	0.1169		
Total	102	24.1364				

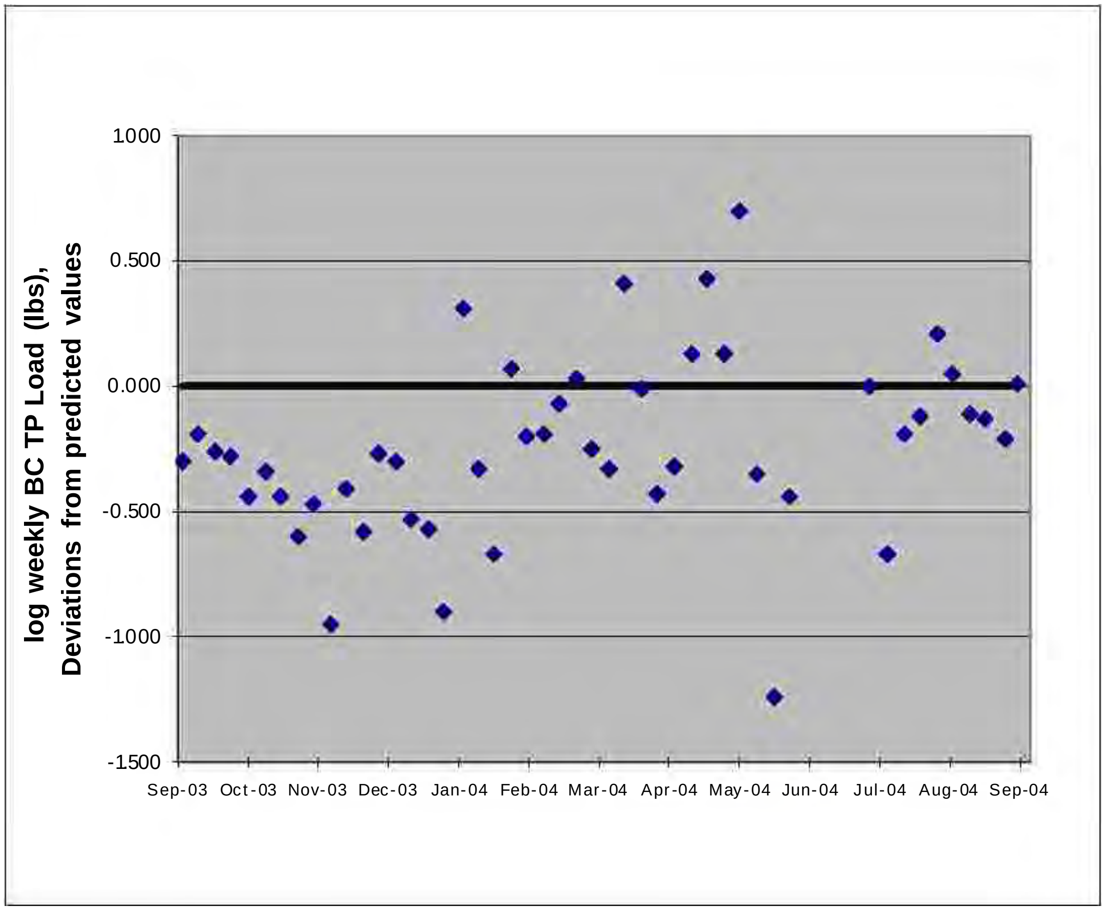

Figure 3. Plot of deviations of predicted weekly log BC T-P from observed values during the treatment period.

loads for Beaty Ck. during the treatment period are less than those predicted by the calibration relationship and contribute to the many negative differences seen in Figure 3.

Again, to arrive at an estimate of BMP influence on the weekly T-P load in Beaty Ck., it is necessary to consider the difference between the observed and expected loads for the treatment period as opposed to the difference in calibration and treatment period means, which are not corrected for environmental variability (Table 3). Although the treatment mean weekly T-P load for Beaty Ck is a little higher than the calibration period mean, it is actually 14% lower than what would be predicted before BMPs were implemented, and this is the value of interest in paired watershed studies. Even if the difference in observed and predicted loads had not been statistically significant, a presumed benefit of BMPs is still realized in the difference in observed loads between the periods when compared to Little Saline Creek. Despite no adjustment in watershed management practices, Little Saline exhibited a 176 percent increase in observed weekly T-P load between the periods as opposed to an 18 percent increase in Beaty Creek, which is assumed to have experienced a similar magnitude of difference in load driving runoff between calibration and treatment periods.

Table 3. Mean observed weekly T-P loads (lbs) for study periods and predicted load for Beaty Ck. during treatment period. Change in Beaty T-P load is calculated using the weekly loads observed and predicted during the treatment period.

	Mean weekly T-P load (lbs)	
Calibration		
Little Saline	30.29	
Beaty	90.66	
Treatment		
Little Saline	83.56	
Beaty (observed)	106.81	
Beaty (predicted)	123.60	
Change in T-P load*	*(observed-predicted/observed)*	-14%

*adjusted for environmental variability as accounted for in Little Saline

A more concise and perhaps clearer method to relay the difference between study periods in the Beaty Creek T-P data is to present the adjusted means of the log transformed data from the ANCOVA analysis previously discussed. Using all the data combined, an overall project mean of log weekly Little Saline T-P load (1.268) was computed and used in both regression equations in Figure 2 to generate corrected log weekly Beaty Creek T-P load means for the calibration and treatment periods. The Minitab results outlined in Table 4 (next page) exhibit clearly the drop in adjusted means of log weekly Beaty Creek T-P load between the periods (0.2458), along with a 95 % confidence interval around the difference and the associated statistical confidence (P=.0016). Again, the percent difference from calibration is approximately 14 percent.

To observe the change in climate adjusted log weekly T-P loads in Beaty Creek, ANCOVAs were performed for each succeeding pair of years for the project using the same procedure outlined in the preceding discussion. By subtracting the first year from the second for each pair, a difference in climate adjusted means was determined and plotted along with the 95% confidence interval to track the effect of BMPs. The results of this analysis show not only a decrease in climate adjusted means from year to year but also a significant increase in the difference between them over the project duration (Figure 4). The statistical confidence in the difference estimates between the first and last years of the project is evident in the non-overlap of the 95% C.I.s.

Table 4. Minitab output detailing adjusted means analysis.

```
Means for Covariates

Covariate         Mean   StDev
log LS T-P load  1.268  0.6104

Least Squares Means for log BC T-P load

Period        Mean  SE Mean
calibration  1.761  0.04939
treatment    1.515  0.05215

Tukey 95.0% Simultaneous Confidence Intervals
Response Variable log BC T-P load
All Pairwise Comparisons among Levels of Period
Period = calibration  subtracted from:

Period       Lower   Center     Upper  ---+---------+---------+---------+---
treatment  -0.3964  -0.2458  -0.09520  (------------*-----------)
                                       ---+---------+---------+---------+---
                                       -0.36     -0.24     -0.12      0.00

Tukey Simultaneous Tests
Response Variable log BC T-P load
All Pairwise Comparisons among Levels of Period
Period = calibration  subtracted from:

           Difference       SE of            Adjusted
Period       of Means  Difference  T-Value    P-Value
treatment     -0.2458     0.07592   -3.238     0.0016
```

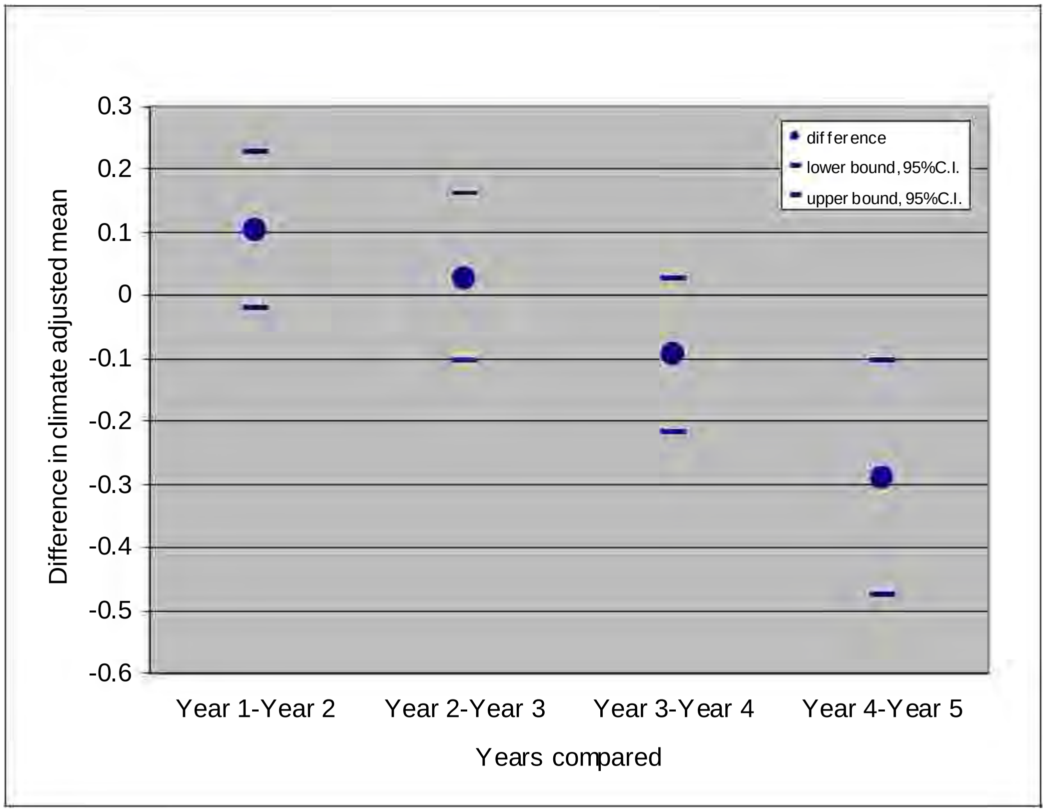

Figure 4. Difference in year-to-year climate adjusted means of log weekly T-P load for Beaty Creek. Hash marks depict the 95% confidence interval around the estimate.