

# NWQEP NOTES

The NCSU Water Quality Group Newsletter

Number 92

November 1998

ISSN 1062-9149

NC STATE UNIVERSITY

## IN THIS ISSUE

Instruction Exchange .....	1
Editor's Note .....	2
Information .....	10
WWW Resources .....	10
Meetings .....	10
Meeting Report.....	11

*NWQEP NOTES* is issued bimonthly. Subscriptions are free. *NWQEP NOTES* is also available on the World Wide Web at <http://www.bae.ncsu.edu/bae/programs/extension/wqg/issues/index.html>. To request that your name be added to the mailing list, use the enclosed publication order form or send an email message to [wq\\_puborder@ncsu.edu](mailto:wq_puborder@ncsu.edu). A publications order form listing all publications on nonpoint source pollution distributed by the NCSU Water Quality Group is included in each hardcopy issue of the newsletter and is also available at [http://www.bae.ncsu.edu/bae/programs/extension/wqg/issues/pub\\_order.html](http://www.bae.ncsu.edu/bae/programs/extension/wqg/issues/pub_order.html)

## Instruction Exchange

### Detecting Water Quality Changes Before and After BMP Implementation: Use of a Spreadsheet for Statistical Analysis

*Garry L. Grabow, Jean Spooner, Laura A. Lombardo, and Daniel E. Line*

NCSU Water Quality Group  
Biological and Agricultural Engineering Department  
NC State University

Detection of differences or trends in water quality data requires organization and analysis of data collected from the field. Such analysis may be done on a computer using statistical packages such as SAS, or by using spreadsheets. Recent versions of spreadsheet software include several statistical functions, and with minimal data setup, may be used to analyze and interpret data<sup>1</sup>.

This article illustrates the use of a spreadsheet to organize and analyze data for detecting a water quality change due to changes in land management. The spreadsheet used here is Excel version 7.0<sup>2</sup>. The parametric<sup>3</sup> statistical test, regression analysis (analysis of covariance), will be presented for use with paired watershed, upstream/downstream, and before/after water quality monitoring designs.

See Grabow et al. (1998) for a presentation of these analyses using SAS (also to be demonstrated in the next issue of *NWQEP NOTES*).

<sup>1</sup> It is assumed that the reader has had some experience with a spreadsheet, and an introductory course in statistics, or some statistical experience.

<sup>2</sup> While directions refer specifically to Excel version 7.0 commands, other spreadsheets have similar capabilities.

<sup>3</sup> Excel does not have non-parametric data analysis tools.



## Monitoring Designs

Three monitoring designs common to water quality studies are paired watershed, upstream/downstream, and before/after. A *paired watershed* design (Clausen and Spooner, 1993) comprises two watersheds (control and treatment) of similar location and land use and two time periods of study (calibration and treatment). Typically, one sampling station is positioned at the outlet of each watershed. During the calibration period (typically at least two years), land use at both control and treatment sites should remain the same. The goal is to establish a relationship between the watersheds. At the end of the calibration period, best management practices (BMPs) are implemented at the treatment site. The project then proceeds into the treatment period (usually at least 2 years).

Again, the goal is to establish a relationship between control and treatment watersheds. At the end of, or at some point during, the treatment period, the relationships are compared to see if a change has occurred, or is occurring, due to BMP implementation.

An *upstream/downstream-(before/after)* design (Spooner et al., 1985) also requires calibration and treatment periods (before and after BMP implementation); however, unlike the paired watershed design, only one watershed is monitored, with sampling stations positioned upstream and downstream of the treatment area.

With a *before/after* monitoring design (Spooner et al., 1985), water quality data from one downstream station is collected for a period of time before and after BMP implementation.

### Example

As an example, data obtained from the Morro Bay, California EPA 319 National Monitoring Program Project will be used to demonstrate how to detect whether a change in average storm turbidity has occurred due to BMP installation. Regression analysis will be performed using a spreadsheet. The sampling design is a *paired watershed* on two sub-basins within the Morro Bay watershed. Chumash Creek is the treatment watershed, while Walters Creek is the control.

### Raw Data Organization and Exploratory Data Analysis

With most spreadsheet operations, analysis is easiest when data are placed in columns with no blank rows, and labels only at the top. Different worksheets (pages) within a workbook may be used to logically separate data, such as from different sampling stations. Benefits of this format include easy generation of intermediate variables which may be required for the analysis (e.g., log transformed values) and greater ease of specifying data ranges for the statistical functions found in the spreadsheet.

With all three monitoring designs, the water quality data are paired by date and should be placed on the same spreadsheet row (see Figure 1). For a *paired watershed* design, the paired data are the control and treatment watershed data collected from the same time period or sampling event. In an *upstream/downstream* design, the pairing is between data collected from upstream and downstream stations for the same time period or sampling event. A lag may be required if the travel time between the two stations is significant relative to the sampling frequency.

## EDITOR'S NOTE

For nonpoint source pollution control projects, it is never too early to start thinking about how all of the data that is going to be collected, or has been collected, will be analyzed. Many of the 319 National Monitoring Program projects are in the post-BMP implementation phase, and as long as one year's worth of post-BMP monitoring data has been collected, analysis for detecting improvements in water quality can begin.

In this issue of *NWQEP NOTES*, our feature article describes how to use the tools available in a spreadsheet for determining statistical significance of water quality changes due to land treatment. In our next issue of *NWQEP NOTES*, we will illustrate similar analyses using the statistical package SAS.

As always, please feel free to contact us regarding your ideas, suggestions, and possible contributions to this newsletter.

Laura Lombardo  
 Editor, *NWQEP NOTES*  
 Water Quality Extension Associate  
 NCSU Water Quality Group  
 Campus Box 7637, NCSU  
 Raleigh, NC 27695-7637  
 Tel: 919-515-3723, Fax: 919-515-7448  
 Email: notes\_editor@ncsu.edu

Design	Date/Time	Treatment	Control	Period
		Watershed	Watershed	Before/After
Paired Watershed	t <sub>1</sub>	y <sub>1</sub>	x <sub>1</sub>	Before
	t <sub>2</sub>	y <sub>2</sub>	x <sub>2</sub>	Before
	t <sub>3</sub>	y <sub>3</sub>	x <sub>3</sub>	Before
	.	.	.	After
Upstream/Downstream	Date/Time	Downstream	Upstream	Period
	t <sub>1</sub>	y <sub>1</sub>	x <sub>1</sub>	Before
	t <sub>2</sub>	y <sub>2</sub>	x <sub>2</sub>	Before
	t <sub>3</sub>	y <sub>3</sub>	x <sub>3</sub>	Before
.	.	.	After	
Before/After	Date/Time	Wat Quality Parameter	Explanatory Variable	Period
	t <sub>1</sub>	y <sub>1</sub>	x <sub>1</sub>	Before
	t <sub>2</sub>	y <sub>2</sub>	x <sub>2</sub>	Before
	t <sub>3</sub>	y <sub>3</sub>	x <sub>3</sub>	Before
.	.	.	After	

Figure 1: Data Setup in Spreadsheet for 3 Monitoring Designs

For a *before/after* design, the explanatory variable (see below) is paired with the water quality parameter of interest.

Before performing any statistical analyses to detect differences or change, some “exploratory data analysis” should be done to see if the data are in the proper form for analysis. The statistical test discussed here, regression analysis, is parametric and requires that the data be approximately normally distributed and independent (not autocorrelated). Regression analysis also requires this of the *residuals* (discussed later in this article). Note that non-parametric tests also assume independence.

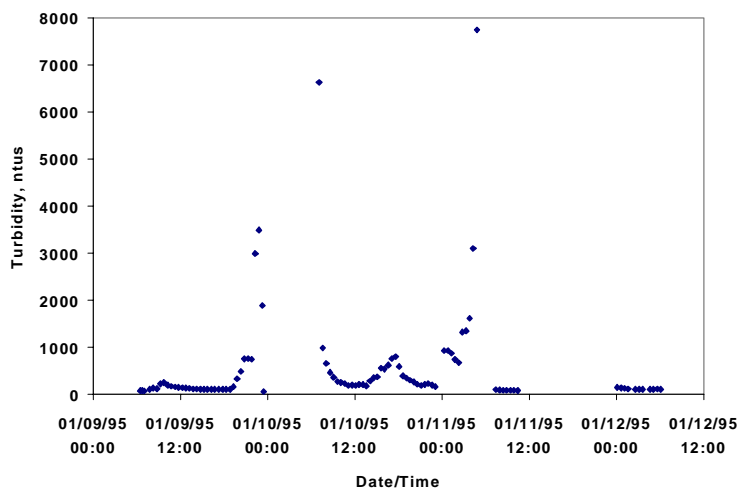


Figure 2: Autocorrelation exhibited in Walters Creek Turbidity data

Check for Autocorrelation

If the raw data are collected at short time intervals (e.g., sub-hourly, hourly, daily, or even weekly), there is a good chance it will be autocorrelated. With regression analysis, the real issue is whether the *residuals* are autocorrelated. By checking and adjusting the data for autocorrelation, this most likely will insure that the residuals are independent.

Autocorrelation occurs quite often in water quality and hydrologic data, as evidenced by hydrographs, sedigraphs, and chemographs, which exhibit an autocorrelated trend in the data. Autocorrelation, sometimes referred to as serial correlation, is reflected by data that is related or “correlated” to previous and subsequent samples. Autocorrelated data contain redundant information and less information than would the same number of independent observations. Therefore, statistical tests performed on non-independent or autocorrelated data may not be as conclusive as the tests imply. As a result, adjustment for autocorrelation must be made to ensure proper interpretation.

Perhaps the easiest way to check for autocorrelation is to plot the data over time. Repeated trends or patterns are an indication of autocorrelation. Autocorrelation can be seen in Figure 2 for the raw 30-minute paired turbidity data collected on Walters Creek.

More formal tests for autocorrelation can be performed by statistical packages such as SAS (PROC AUTOREG)<sup>4</sup>.

Correction for Autocorrelation

The easiest way to eliminate autocorrelation, and thus simplify the subsequent analysis, is to aggregate or average the data into larger time steps; for instance, from hourly to daily, or from daily to storm period. This is often termed “data reduction.” In the example, since the 30-minute turbidity data collected from Chumash and Walters Creek were autocorrelated, averages were calculated from paired data for each storm event lasting one to three days. If it is not possible to take an average – perhaps this results in too few data points – there are other ways of accounting for autocorrelation<sup>5</sup>. Once correction for autocorrelation has been made, organization of reduced data in the spreadsheet should follow the same layout presented earlier in Figure 1.

<sup>4</sup> See Grabow et al., 1998.

<sup>5</sup> Regression using “unreduced” data and accounting for autocorrelation of the residuals may be done in SAS.

**Explanatory Variables**

With a *before/after* single downstream station design, it is necessary to “factor out” or “block” variables other than land treatment that may have changed over time and caused a change in the water quality parameter of interest. Accounting for these variables allows for better documentation of the water quality change due to land treatment. These variables are called “explanatory variables” since they explain much of the variation in data, and may include rainfall, streamflow, or other hydrologic factors.

Explanatory variables are useful in all monitoring designs to document the relationship between water quality and hydrologic and/or meteorologic changes. With a *before/after* design, explanatory variables must be integrated into the analysis. For example, suspended sediment yield is highly dependent upon flow. Therefore, flow could be selected as an explanatory variable, and flow values corresponding to specific sampling events could then be added to the same spreadsheet line with the sediment values, the water quality parameter of interest (see Figure 1 for case of before/after design). In a *paired watershed* or *upstream/downstream* design, including explanatory variables in the statistical analysis is less important since concurrent or “paired” data are collected under similar hydrologic conditions. Explanatory variables must also be checked for autocorrelation.

Once the data (including explanatory) are corrected for autocorrelation and organized in the spreadsheet, further evaluation is required to see if the data are normally distributed.

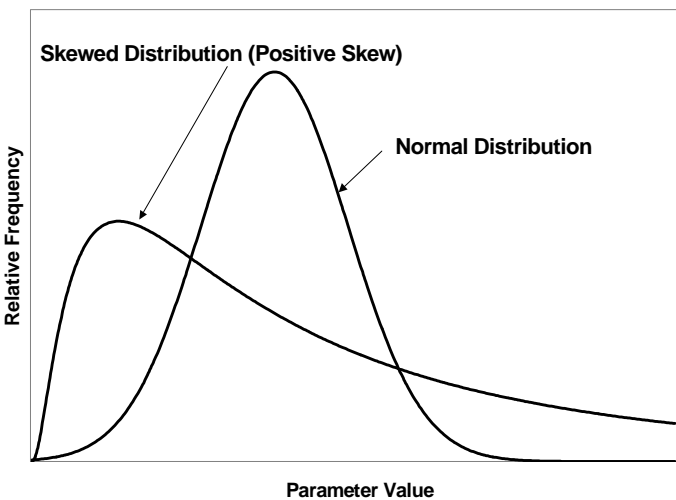


Figure 3: Normal and Skewed Distributions

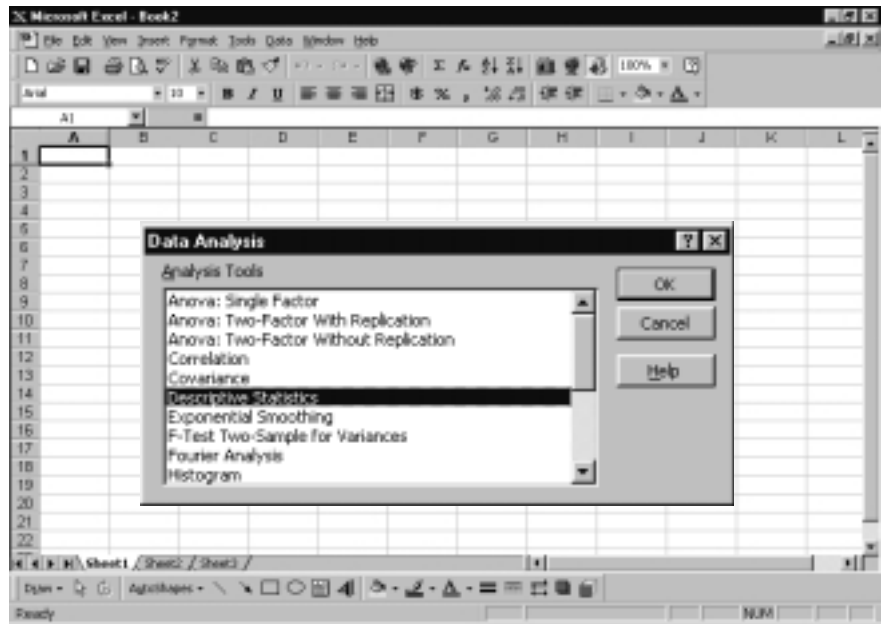


Figure 4: Excel Analysis Tools

**Check for Normality**

Parametric statistical tests, or those tests based upon estimates of statistical parameters such as mean and standard deviation, require that the data be normally (or nearly so) distributed. A normal distribution is indicated by a bell-shaped, symmetrical curve. Many times water quality and hydrologic data are not normally distributed, but are skewed. The skewness indicates the departure of the distribution from the normal curve. Data are positively skewed if the curve has a longer “right tail” and negatively skewed if it has a longer “left tail” (see Figure 3). In most storm event-related water quality data, the skew will be positive reflecting many relatively low values and a few very high values.

There are two simple ways to check for skewness in data using a spreadsheet. One way is to use the descriptive statistics<sup>6</sup> function available in Excel. The *Data Analysis*<sup>7</sup> command under *Tools* allows access to this feature (see Figure 4). Once in the descriptive statistics dialog box, select the input range (your data) and a place in the spreadsheet to put the results (output range), and check the “Summary Statistics” box. Among the results will be “Skewness” (see Figure 5). If the data are perfectly “normal” the coefficient of skewness will be zero. As a rule of thumb, an absolute value of skewness above 1.0 indicates a degree of skewness in the data that should be addressed.

<sup>6</sup> This and other statistical functions may be found in the “Analysis ToolPak” add-in. This will need to be installed if not available under the “Tools” command.

<sup>7</sup> Italics indicate a menu command in Excel.

Figure 5 shows a high degree of skewness (3.02) for Chumash storm average turbidity data.

Column1	
Mean	722.58
Standard Error	133.10
Median	501.07
Mode	#N/A
Standard Deviation	787.42
Sample Variance	620034.65
Kurtosis	10.00
<b>Skewness</b>	<b>3.02</b>
Range	3873.63
Minimum	74.17
Maximum	3947.80
Sum	25290.45
Count	35

Figure 5: Summary Statistics for Chumash Turbidity Data

A more visual way of checking for normality is to develop a histogram of the data. This is also accessed through the Analysis ToolPak via the *Tools/Data Analysis* menu commands. In the histogram dialog box, enter the data range and check the box for “chart output”. A graph will be generated and the degree of normality, or conversely skewness, can be seen by the symmetry of the histogram. A bell shaped symmetry indicates normality. If you check the “cumulative probability” box, a symmetrical sigmoidal or “S” shaped curve would indicate normally distributed data.

Figure 6 shows that average storm turbidity data from Chumash Creek are skewed.

**Correction for Skewness**

If the data are skewed, transformations of the variable(s) should be done. The most common transformation in water quality and hydrology is a logarithmic transformation. This is done in Excel by using the LOG function “=LOG(x)” to return the base 10 logarithm of the value x in a new column of the spreadsheet. After doing this transformation, repeat the diagnostic steps given previously to check for normality.

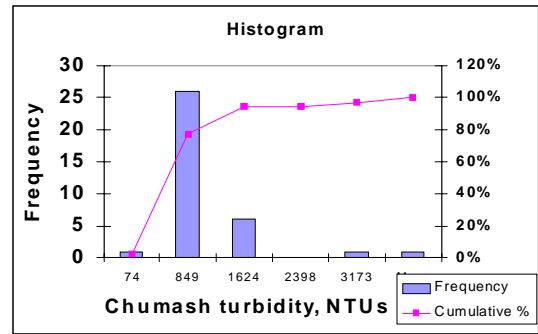


Figure 6: Histogram of Chumash Turbidity Data

In the example, log transformation of the Chumash average storm turbidity data reduced the skewness from 3.02 to 0.17. The storm-averaged example data set, including log transformation, is shown in Figure 7. Even if the data are moderately skewed, in many cases log transformation will improve the distribution and therefore improve the subsequent analysis. If the data remain highly skewed after transformation, non-parametric data analysis may be required<sup>8</sup>.

Observation (storm)	Control Walters avg turb ntu	Treatment Chumash avg turb ntu	(1) log Chumash	(2) log Walters
1	603.7	3169.3	3.50	2.78
2	113.3	1584.9	3.20	2.05
3	613.3	1004.8	3.00	2.79
4	110.5	116.7	2.07	2.04
5	108.7	430.0	2.63	2.04
6	92.8	578.0	2.76	1.97
7	1132.0	1206.9	3.08	3.05
8	163.0	368.0	2.57	2.21
.	.	.	.	.
.	.	.	.	.

Figure 7: Example Data Set (Partial) after Storm Averaging and Log Transformation

**Regression Analysis to Detect Change**

This discussion is limited to detecting discrete changes, such as *before* and *after* BMP implementation, as opposed to gradual changes. A good way of detecting a discrete water quality change due to land treatment changes is **regression analysis**. This test requires that at least one year of post-BMP data have been collected. The test may be repeated as more data are collected.

<sup>8</sup> Non-parametric analysis is not treated in this fact sheet since it is not directly available in Excel.

Regression analysis as presented here is actually an analysis of covariance (ANCOVA), where two variables (“covariates”) are being compared for a linear relationship. This analysis is stronger than an ANOVA or *t-test*, particularly for covariates with a strong linear relationship. In our paired watershed example, the covariates are log-transformed turbidity data from the control (Walters) and treatment (Chumash) watersheds.

The regression analysis will establish a linear relationship (expressed as an equation for a line) between the covariates, where the dependent variable will be defined in terms of the independent variable. A relationship will be established for both the calibration and treatment periods. This relationship expressed as a line is referred to as a “best fit line.” The regression analysis will detect whether or not these lines of best fit from the calibration and treatment periods differ. A difference in the regression lines indicates a change due to land treatment. The analysis will test for the magnitude of this difference, by examining the slope and intercept of the lines, and will define the surety or “significance” of the difference.

In a *paired watershed* design, the independent variable (X) represents the water quality parameter of interest collected from the control watershed, while the dependent variable (Y) is the same parameter from the treatment watershed.

In an *upstream/downstream* monitoring design, the independent variable (X) represents the water quality parameter of interest collected from the upstream station while the dependent variable (Y) represents data from the downstream station.

For a *before/after* design, the independent variable (X) represents the explanatory variable, such as water discharge, while the water quality parameter of interest is the dependent (Y) variable.

To prepare data for regression analysis in Excel, it is first necessary to generate two variables in addition to the two paired water quality variables. The first is an “indicator” variable, which assumes a value of 0 or 1 to differentiate the data between calibration and treatment periods (X2 in Figure 8). The indicator variable allows the regression lines to have different intercepts. The second variable is created by multiplying the first indicator variable (0 or 1) by the independent variable (X1\*X2 in Figure 8). The X1\*X2 term is called an “interaction” term and allows the two regression lines to have different slopes.

Addition of these two independent variables is required to develop and compare the two separate regression lines. In the example data set, the control watershed variable (log Walters TSS) and two created variables are the independent or X variables (X1, X2, and X1\*X2 in Figure 8), while the treatment watershed variable (log Chumash TSS) is the dependent or Y variable. The independent variables need to be in contiguous columns when using regression in Excel.

Y Log Chumash	X1 log Walters	X2 Pre(0) Post(1)	X1*X2 log Walters
3.501	2.781	0	0.000
3.200	2.054	0	0.000
3.002	2.788	0	0.000
2.067	2.043	0	0.000
2.633	2.036	0	0.000
2.762	1.967	0	0.000
3.082	3.054	0	0.000
2.566	2.212	0	0.000
2.892	2.797	0	0.000
2.942	2.595	0	0.000
2.825	2.941	0	0.000
2.726	2.789	0	0.000
3.596	3.050	1	3.050
2.650	2.433	1	2.433
.	.	.	.
.	.	.	.

Figure 8: Variables Required for Calibration/Treatment Period Regression Analysis

The regression analysis in Excel is accessed through the “Analysis ToolPak” under *Tools/ Data Analysis/ Regression*. First, an input range for the independent variable(s) and dependent variable must be selected. For the example, the independent variables include the data in columns X1, X2, and X1\*X2, selected together as one block. The dependent variable is the data in column Y (Figure 8). Second, a location on the spreadsheet must be selected for the output, which is a summary of the analysis results.

As an important side note, in regression analysis, the residuals should be normally distributed and uncorrelated, and have a constant variance. Residuals are the differences between the actual Y values of the data points and the predicted values on the regression lines. To check for this, under *Regression*, mark the boxes for “standardized

residuals” and “residual plots.” This will generate a table called “Residual Output,” listing for each observation the predicted Y value and the corresponding residual. Also generated are plots of each of the independent X variables versus the residuals. In the example, since there are three independent variables in the analysis, three residual plots are generated.

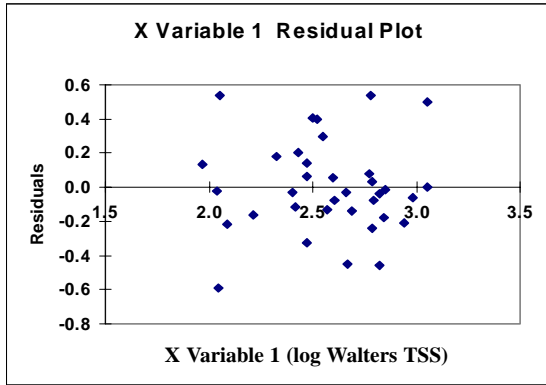


Figure 9: Residual Plot

Typically, if water quality data have been log transformed, the residuals from a regression analysis are normally distributed and have a constant variance. A non-constant variance will show up in the residual plot as a “funnel” or “mounding” rather than an even distribution. Figure 9 shows that the residuals from a plot with the X1 variable have a constant variance. Also, since these data were aggregated to a storm basis, autocorrelation is not a concern. If autocorrelation was a concern, it would be necessary to plot the residuals over time. Once again, repeated trends or patterns are an indication of autocorrelation. In order to do this, go to the X1 Variable residual plot and click on the data points, then select *source data*. From there, change the x values of the data series to the column labeled “Observation” in the Residual Output table. This will generate a plot of the residuals over time, assuming that the data from which the residuals are generated are entered in chronological order. Evidence of autocorrelation in the residuals plotted over time may indicate that the trend is non-linear. Finally, normality of the residuals can be checked as discussed in the section *Check for Normality*.

The results from the regression analysis for the example, generated by Excel, are illustrated in Figure 10, Summary Output. The regression equation resulting from this analysis is

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 \quad (1)$$

where Y is the estimate of the dependent variable, X<sub>1</sub> is the independent variable, X<sub>2</sub> is the indicator variable, and  $\beta_{i, i=0,3}$  are the regression coefficients. Recall that X<sub>2</sub>=0 for the calibration period, and X<sub>2</sub>=1 for the treatment period. This equation reduces to:

$$Y = \beta_0 + \beta_1 X_1 \quad (2)$$

for the calibration period and

$$Y = \beta_0 + \beta_2 + (\beta_1 + \beta_3) X_1 \quad (3)$$

for the treatment period,

where

$\beta_0$  = y-intercept of the calibration regression line

$\beta_1$  = slope of the calibration regression line (X Variable 1)

$\beta_2$  = difference in y-intercept between calibration and treatment lines (X Variable 2)

$\beta_3$  = difference in slope between calibration and treatment lines (X Variable 3)

SUMMARY OUTPUT					
<i>Regression Statistics</i>					
Multiple R		0.65774571			
R Square		0.43262941			
Adjusted R Square		0.37772258			
Standard Error		0.28624898			
Observations		35			
ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	3	1.936862683	0.645621	7.879337	0.000475621
Residual	31	2.540092922	0.081938		
Total	34	4.476955605			
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>
Intercept, $\beta_0$	1.810872	0.536154243	3.377521	0.001986	0.717377598
X Variable 1, $\beta_1$	0.41480983	0.21149849	1.96129	0.058881	-0.01654443
<b>X Variable 2, <math>\beta_2</math></b>	<b>-1.9125649</b>	0.888344182	-2.15295	<b>0.039216</b>	-3.72435582
<b>X Variable 3, <math>\beta_3</math></b>	<b>0.63063967</b>	0.34345195	1.83618	<b>0.075938</b>	-0.06983559

Figure 10: Excel Regression Analysis Output for Equation with Interaction Term



**Is There a Difference?**

Once again, the purpose of regression analysis is to test for a difference in the regression lines between calibration and treatment periods. Both the y-intercepts and slopes of the lines are examined. The *statistical significance* of a difference is revealed in the *P-values*, while the best guess of the *magnitude* of the difference is revealed in the *Coefficient* values. So when interpreting the Summary Output, it is important to first check the *P-values* to see if there is a difference, and if so, then look at the *Coefficient* values of the  $\beta$ s for an estimate of the difference.

The Summary Output (Figure 10) indicates that for the calibration period, the y-intercept ( $\beta_0$ ) of the regression line is 1.81 and the slope of this line ( $\beta_1$ ) is 0.41.

Looking at  $\beta_2$ , the *P-value* of 0.039 indicates that a statistically significant difference in y-intercepts between calibration and treatment regression lines exists at a 96% confidence level. The *Coefficient* value of -1.91 represents an estimate of the magnitude of this difference. Furthermore, the negative sign indicates that the y-intercept of the treatment regression line is lower than the y-intercept of the calibration line, indicating a reduction in turbidity at the treatment watershed relative to the control watershed.

Looking now at  $\beta_3$ , the *P-value* of 0.076 indicates that a statistically significant difference in slopes of the regression lines exists at a 92% confidence level. The *Coefficient* value of 0.63 represents an estimate of the magnitude of this difference. Since this is a positive number, the slope of the line for the treatment period is greater or steeper ( $0.41 + 0.63 = 1.04$ ) than the line for the cali-

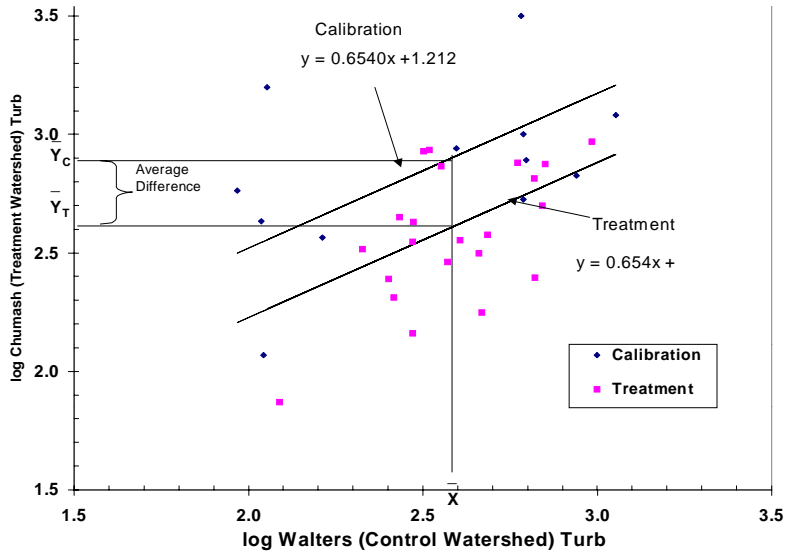


Figure 12: Calibration and Treatment Period Regression Lines for Reduced Model

bration period, indicating a greater reduction at lower levels of turbidity and a lesser reduction at higher levels. This can be seen in the calibration and treatment period regression lines shown in Figure 11, as the greatest distance between the two lines is at the lower levels of turbidity, with the difference lessening at higher levels. Note that if the  $\beta_3$  *Coefficient* value was negative, indicating a decrease in slope, this would mean greater reductions at higher levels of turbidity.

If a 92% confidence level in difference in slope is not acceptable (say, for example, a significance level was set at 95% ( $\alpha=0.05$ )), then the *interaction* ( $X_1 * X_2$ ) term should be dropped from the regression model. In this case, the calibration period equation would be the same as Equation 2 and the treatment period equation would now be

$$Y = \beta_0 + \beta_2 + \beta_1 X_1$$

This would result in regression lines with the same slope (parallel) but with y-intercepts differing by a magnitude of  $\beta_2$  (see Figure 12). Figure 13 shows the Summary Output generated in Excel for the revised regression model.

In this scenario with the *interaction* term left out, the *P-value* (0.01) of  $\beta_2$  (the y-intercept term) is lower than in the previous model (0.039), indicating a greater confidence level (99%) that the difference in intercepts is statistically significant. The *Coefficient* value of -0.29 once again indicates a reduction in turbidity of the treatment watershed relative to the control watershed.

Figures 11 and 12 can be generated in Excel through *Chart Wizard*, selecting XY (Scatter) plot.

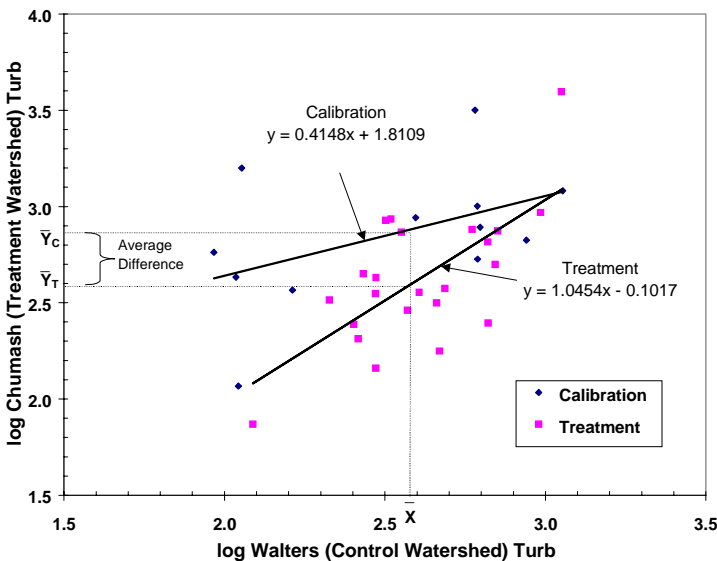


Figure 11: Calibration and Treatment Period Regression Lines for Full Model



**How Much is the Difference?**

Once a difference in the water quality parameter of interest between calibration and treatment periods has been statistically proven, the next step is to estimate the amount of the difference. The average difference for the “full model” with interaction term may be obtained by setting  $X_1$  in equations (2) and (3) equal to the average of the values in the  $X_1$  (log Walters) column in Figure 8 over the *entire* period (both calibration and treatment). By using an  $X_1$  value averaged over the entire study period to calculate the average difference in  $Y$  value, changes in turbidity at the control watershed occurring over the study period will be accounted for. For the example data set, this average  $X_1$  value is calculated to be 2.573 log turbidity units ( $X$  in Figure 11). Now, substitute 2.573 in for  $X_1$  in equations (2) and (3), and using the  $\beta$  Coefficient values from the Summary Output (Figure 10), solve for the  $Y$ s. The predicted  $Y$  values for Chumash, the treatment watershed, are 2.878 log units using the calibration period regression equation, and 2.588 log units using the treatment period regression equation ( $Y_c$  and  $Y_T$ , respectively, in Figure 11). The following equation is then used to determine the percent decrease in the origi-

$$\left[1 - \left(\frac{10^{2.588}}{10^{2.878}}\right)\right]100$$

nal, untransformed scale:

which equals a 49% average reduction in turbidity. If the “reduced model” without the interaction term is used, and with the same method to calculate the amount of the difference, the average reduction in turbidity would be 49%.

**Acknowledgements**

The authors would like to express their thanks to Dave Paradies and Karen Worcester, Morro Bay 319 National Monitoring Program Project, for their collaboration. The authors also express their thanks to Dr. Francis Giesbrecht, Professor of Statistics, North Carolina State University, for his review and contributions.

Regression Statistics	
Multiple R	0.60903385
R Square	0.37092223
Adjusted R Square	0.33160487
Standard Error	0.29666654
Observations	35

ANOVA					
	df	SS	MS	F	Significance F
Regression	2	1.660602366	0.830301	9.434057	0.000601547
Residual	32	2.816353239	0.088011		
Total	34	4.476955605			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%
Intercept, $\beta_0$	1.21187302	0.440975063	2.748167	0.009765	0.313636969
X Variable 1, $\beta_1$	0.65395623	0.172704397	3.786564	0.000635	0.302169186
X Variable 2, $\beta_2$	-0.2924882	0.107144126	-2.72986	<b>0.010214</b>	-0.51073349

Figure 13: Excel Regression Analysis Output for Equation without Interaction Term

or 49%.

**For More Information**

Garry Grabow, PhD., P.E.  
 Water Quality Extension Specialist  
 garry\_grabow@ncsu.edu

Laura Lombardo  
 Water Quality Extension Associate  
 laura\_lombardo@ncsu.edu

NCSU Water Quality Group  
 Campus Box 7637  
 Raleigh, NC 27695-7637  
 (919) 515-3723

**References**

Clausen, J.C. and J. Spooner. 1993. Paired Watershed Study Design. Biological and Agricultural Engineering Department, North Carolina State University, Raleigh, NC. EPA-841-F-93-009. 8p.

Grabow, G.L., J. Spooner, L.A. Lombardo, D.E. Line, and K.L. Tweedy. 1998. Has Water Quality Improved?: Use of a SAS for Statistical Analysis of Paired Watershed, Upstream/Downstream and Before/After Monitoring Designs. U.S. EPA-NCSU-CES Grant No. X825012.

Spooner, J., R.P. Maas, S.A. Dressing, M.D. Smolen, and F.J. Humenik. 1985. Appropriate Designs for Documenting Water Quality Improvements from Agricultural NPS Control Programs. *In: Perspectives on Nonpoint Source Pollution*. EPA 440/5-85-001. Pp. 30-34.

## INFORMATION

---

### Providing Safe Drinking Water in America

The U.S. EPA recently released its first annual national assessment of drinking water system compliance, entitled "Providing Safe Drinking Water in America: 1996 National Public Water System Annual Compliance Report and Update on Implementation of the 1996 Safe Drinking Water Act Amendments." Based on the most recently available 1996 state-provided information, the report shows that 86 percent of all Americans are served by water systems with no reported violations of health standards and that most violations of those standards occurred in the very smallest systems. The report also describes the many steps the Clinton Administration is taking to improve the safety of the nation's drinking water. The full report is available in PDF format for viewing or download. Summaries of each state's report are included, as is information for obtaining full state reports. The web site is <http://www.epa.gov/ogwdw/annual>

### Conservation Design for Stormwater Management

A manual (Brandywine Conservancy and the DE Dept. of Natural Resources and Environmental Control, 225 pp., 1998) for developers and municipal planning agencies on nonstructural conservation techniques for stormwater management that treat stormwater as a resource rather than a waste. Available for \$25 plus s/h from DE DNR&EC, 89 Kings Highway, Box 1401, Dover, DE 19903. Tel: 302-739-4411

### Planning for the Future: A Handbook on Community Visioning

This handbook (Center for Rural Pennsylvania, 1998, 52 pp.) provides a practical tool for individuals who want to plan for change in their communities. Available from CRP at 717-772-3587.

### Community Voices for Sustainability

Izaak Walton League of America (1998, 32 pp.) offers advice on organizing and conducting workshops to address community sustainability issues. 797 Conservation Lane, Gaithersburg, MD 20878-2983. Tel: 301-548-0149, email: [sustain@iwla.org](mailto:sustain@iwla.org), web site: <http://www.iwla.org>

■

## WWW RESOURCES

---

*A more complete list of World Wide Web sites that relate to nonpoint source pollution and water quality issues can be found at:*

*<http://www.bae.ncsu.edu/bae/programs/extension/wqg/issues/resource.html>*

### Stormwater Resources

Visit <http://www.stormwater-resources.com> for information on stormwater pollution, treatment methodologies, and related issues.

### Wetlands

The EPA Office of Wetlands, Oceans, and Watersheds has produced a series of fact sheets about wetlands, including information on wetland functions and values and economic benefits. They are available at <http://www.epa.gov/OWOW/wetlands/contents.html>

■

## MEETINGS

---

### Notice of Training Events

**The Institute for Wetland and Environmental Education & Research (IWEER)** is a private educational organization providing technical training to professionals in the environmental sciences. IWEER is presenting a series of field-oriented short courses, including wetland delineation and plant identification. Course schedules, outlines, instructors' bios, and a complete publication list are available at the following web page: <http://members.aol.com/iwer/index.htm>. For information about registering, call (413) 548-8866, send email to [iweer@aol.com](mailto:iweer@aol.com), or write to IWEER, P.O. Box 288, Leverett, MA 01054-0288.

**Working at a Watershed Level: Jan 11-15, 1999, California State University at Chico, California.** Dr. Donald Holtgrieve, Tel: 530-898-5780, Fax: 530-898-6781, email: [holtgrieve@faculty.csuchico.edu](mailto:holtgrieve@faculty.csuchico.edu). The course is designed as an introductory-level basic training program for watershed group coordinators and members, agency staff, consultants and others. This offering of the course is sponsored by the Anadromous Fish Restoration Program, and the focus will be on those watersheds within California's Central Valley that support natural production of chinook salmon, steelhead and other anadromous fish.

## Meeting Announcements — 1998

### NOVEMBER

**Agricultural Solutions for the Neuse River Basin: Nov 9-10, 1998, New Bern, NC.** Joni Tanner, Soil Science Department, NCSU, Box 7619, Raleigh, NC 27695-7619, Tel: 919-513-1678, email: joni\_tanner@ncsu.edu

**ENVIROSOFT '98 – Development & Application of Computer Techniques to Environmental Studies: Nov 10-12, Las Vegas, NV.** Sue Owen, Conference Secretariat, ENVIROSOFT '98, Wessex Institute of Technology, Ashurst Lodge, Ashurst, Southampton, S0407AA, UK, Tel: 44(0)170-329-3223, Fax: 44(0)170-329-2853, email: sue@wessex.ac.uk

**1998 Annual Conference on Water Resources & Symposia on Management of Human Impacts on the Coastal Environment and Applications of Water Use Information: Nov 15-19, Point Clear, AL.** AWRA, Attn: 1998 Annual Conference & Symposia, 950 Herndon Pkwy, Ste 300, Herndon, VA 20170-5531, Tel: 703-904-1225, Fax: 703-904-1228

**Society of Environmental Toxicology and Chemistry 19th Annual Meeting – The Natural Connection: Environmental Integrity and Human Health: Nov 15th-19th, Charlotte, NC.** SETAC Office, 1010 North 12th Ave., Pensacola, FL 32501-3367, Tel: 850-469-1500, Fax: 850-469-9778, email: setac@setac.org, web site: <http://www.setac.org>

### DECEMBER

**XVIII Interamerican Congress of Chemical Engineering: Dec 6-10 1998, San Juan, Puerto Rico.** Professor Federico Padron, Department of Chemical Engineering, PO Box 9046, Mayaguez, PR 00681-9046. Tel: 787-832-4040 ext 2593, Fax: 787-265-3818; email: f\_padron@rumac.upr.clu.edu, or iacche@cheme.cornell.edu

- \* **The North Carolina Lake Management Society's 5th Annual Fall Workshop:** Water Quality and Watershed Issues at Lake Rogers: Dec 11, 1998. Creedmoor, NC. Contact Bryn Tracy, NCLMS, PO Box 28348, Raleigh, NC 27611-8348. Tel: 919-733-6946, email bryn\_tracy@h20.enr.state.nc.us, web site: <http://www.nc-lakes@don-anderson.com/nc-lakes>

## Meeting Announcements — 1999

**Third Annual American Wetlands Month Conference: Communities Working for Wetlands.** February 18-20, New Orleans; March 18-20, San Francisco; April 8-10, Indianapolis; May 6-8, Andover, Massachusetts. Terrene Institute, 4 Herbert Street, Alexandria, VA 22305, Tel: 703-548-5473, Fax: 703-548-6299, email: terrinst@aol.com, web site: <http://www.terrene.org>

### MAY

**International Conference on Diffuse Pollution: May 16-21, 1999, Perth, Western Australia.** Dianne McLeod, Conference Secretariat, P.O. Box 257, South Perth, WA 6951. Tel: +61 8 9450 1662, Fax: +61 8 9450 2942, email: convlink@wantree.com.au, website: <http://www.viron.wa.gov.au>

### JUNE

**2nd Annual Mitigation Banking Conference: June 13-15, 1999, Atlanta, Georgia.** Contact Erin Foster, Terrene Institute, 4 Herbert Street, Alexandria, VA 22305, Tel: 703-548-5473, Fax: 703-548-6299, email: terrinst@aol.com, web site: <http://www.terrene.org>



### Meeting Report: Sixth National Nonpoint Source Monitoring Workshop

On September 21-24, 1998, the Sixth National Nonpoint Source Monitoring Workshop was held in Cedar Rapids, Iowa. The goal of this annual conference is to share information and progress made in monitoring nonpoint source pollution. The steering committee, lead by Lynette Seigley of Iowa Department of Natural Resources, Geological Survey Bureau, put together an outstanding conference, which was informative and fun.

This year's conference theme was interpreting water quality responses to land treatment. Session topics focused on Total Maximum Daily Loads (TMDLs), analysis of biological and water quality data, land treatment, sediment transport, and nutrient management. Two all-day field trips were offered. One trip visited central Iowa to see Walnut Creek National Wildlife Refuge, restored from rowcrop to native prairie, and the Bear Creek Riparian Management System Project. The other trip visited northeast Iowa to see the Big Spring Basin and Sny Magill Creek Watershed.

The conference brought in six nationally known speakers for the plenary sessions. Also, Don Meals of Vermont Dept. of Environmental Conservation gave a thought-provoking lunch-time presentation on why he believes BMPs will never solve the agricultural nonpoint source pollution problem.

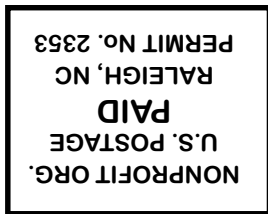
At the conclusion, conference attendees joined in song to bid farewell to Steve Dressing, USEPA, Chief of the Rural Sources Section, Nonpoint Source Control Branch, who directed the EPA 319 National Monitoring Program for the past 8 years. Tom Davenport, USEPA, will be the new program director.

Next year's conference will be held in San Luis Obispo, California on September 13-17, 1999, sponsored by the Morro Bay Watershed. Let's see if the East of the Mississippi softball team can defend their winning title over West of the Mississippi!

Production of *NWQEP NOTES* is funded through U.S. Environmental Protection Agency (EPA) Grant No. X825012. Project Officer: Tom Davenport, Office of Wetlands, Oceans, and Watersheds, EPA. 77 W. Jackson St. Chicago, IL. 60604  
Web Site: <http://www.epa.gov/OWOW/NPS>

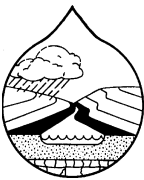
\*

Did not appear in printed newsletter



NCSU Water Quality Group  
Department of Biological and Agricultural Engineering  
North Carolina Cooperative Extension Service  
Campus Box 7637  
North Carolina State University  
Raleigh, NC 27695-7637

**NC STATE UNIVERSITY**



**NCSU Water Quality Group**  
**Campus Box 7637**  
**North Carolina State University**  
**Raleigh, NC 27695-7637**  
**Telephone: (919) 515-3723**  
**Fax: (919) 515-7448**  
**Web Site: <http://www.bae.ncsu.edu/bae/programs/extension/wqg>**

**Personnel**

<b>Jean Spooner</b>	<b>Daniel E. Line</b>
<b>Robert O. Evans</b>	<b>Laura A. Lombardo</b>
<b>Garry L. Grabow</b>	<b>Terry W. Pollard</b>
<b>William A. Harman</b>	<b>Rachel E. Smith</b>
<b>Frank J. Humenik</b>	<b>Kevin L. Tweedy</b>
<b>Gregory D. Jennings</b>	<b>Dani E. Wise</b>