

NWQEP NOTES

The NCSU Water Quality Group Newsletter

Number 93

January 1999

ISSN 1062-9149

NC STATE UNIVERSITY

Instruction Exchange

IN THIS ISSUE

Instruction Exchange	1
Editor's Note	2
Information	11
WWW Resources	12
Meetings	12

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. 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

Detecting Water Quality Changes Before and After BMP Implementation: Use of SAS 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.

This article illustrates the use of SAS¹ to organize and analyze data for detecting a water quality change due to changes in land management². The parametric statistical test, regression analysis, with covariates or explanatory variables (analysis of covariance), will be presented for use with paired watershed, upstream/downstream, and before/after water quality monitoring designs. Analyses for a discrete step-change that compares pre- and post-BMP values is presented. In future articles, we will discuss analyses for other types of changes, such as trends over time.

See our last issue of *NWQEP NOTES* or Grabow et al. (1998) for a presentation of these analyses using a spreadsheet.

¹ SAS (SAS, 1985ab) version 6.12 for PCs is used for this publication; however, other versions and platforms are supported and identical in language. Other statistical packages are available that can perform similar analyses.

² It is assumed that the reader has had an introductory course in statistics or some statistical experience.



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 SAS. 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

Data can be read by SAS from a database (e.g., SAS or Access), a spreadsheet, or an ASCII file. In our example, SAS reads the data from an ASCII file.

With any of the three monitoring designs stated earlier, the data are paired by date, and may be either placed on the same record (row) as shown in Figure 1 or entered by station. 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. For a *before/after* design, the explanatory variable (see below) is paired with the water quality parameter of interest.

For this example, the required data are organized by storm and each record contains the beginning date of the storm, turbidity of the control watershed, and turbidity of the treatment watershed (see Figure 1). Any intermediate variables which may be required for the analysis (i.e., log-transformed values), as well as categorization and sorting of the variables can be done within SAS.

EDITOR'S NOTE

Analyzing water quality data from nonpoint source pollution control projects to detect change can be done on a computer using either spreadsheets or statistical packages.

In our last issue of *NWQEP NOTES*, our feature article described how to use a spreadsheet for determining statistical significance of water quality changes due to land treatment. While spreadsheet applications may be more readily available to our readers, statistical packages potentially provide a more powerful tool for data analysis and therefore are recommended, particularly for dealing with complicating issues such as autocorrelation. In this issue of *NWQEP NOTES*, our feature article illustrates similar analyses using the statistical package SAS.

It should be noted that there is more than one way to define a change in water quality. The analyses presented in this issue and the last issue of *NWQEP NOTES* evaluate discrete step-changes, comparing pre-BMP to post-BMP water quality. However, depending on how long it takes for BMP's to be established (instantaneous vs. gradual), analyzing for a trend over time may be more appropriate than a discrete change. Future issues of *NWQEP NOTES* will discuss and present analyses for gradual types of changes, such as trends over time.

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

```

data file "stormturb.sds" contains avg
storm turbidity data (NTUs). Avgs. are
calculated on paired 30 minute data.

```

Begin Date	Walters (Control) avg turb	Chumash (Treatment) avg turb
01/04/95	603.7	3169.3
01/07/95	113.3	1584.9
01/09/95	613.3	1004.8
01/12/95	110.5	116.7
01/14/95	108.7	430.0
01/20/95	92.8	578.0
01/22/95	1132.0	1206.9
03/05/95	163.0	368.0
01/31/96	626.0	780.0
02/03/96	394.0	875.2
02/04/96	872.3	668.1
02/19/96	614.8	532.4
12/09/96	1121.2	3947.8
.	.	.
.	.	.

Figure 1: ASCII data file for SAS input

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). Technically, the regression analysis requires this of the *residuals* (discussed later in this article). Note that non-parametric tests also assume independence.

Covariates (Explanatory Variables)

With any watershed design, it is important to “factor out” variables other than land treatment that may have changed over time and caused a change in the water quality variable of interest. Accounting for these variables allows for better documentation of the water quality change due to land treatment. These variables are called covariates.

For a paired watershed design, the covariate is the water quality variable of interest in the control watershed. In an upstream/downstream design, the covariate is the upstream water quality parameter of concern. In a before/after single station design, these covariates are often called explanatory variables, and may be rainfall, streamflow or other hydrologic factors.

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.

A more formal way to check for autocorrelation is to use PROC ARIMA or PROC AUTOREG in SAS. Figure 3 shows the SAS code for examining autocorrelation of the raw turbidity data from the Chumash and Walters sites using PROC ARIMA. In this figure, the “nlag=12” option in the “identify” statement tells SAS to check for autocorrelation of all data points at spacings from 1 to 12 time periods (lags) apart. The number of lags is arbitrary and is selected by the user. It is recommended that autocorrelation be explored out to at least 3 or 4 lags.

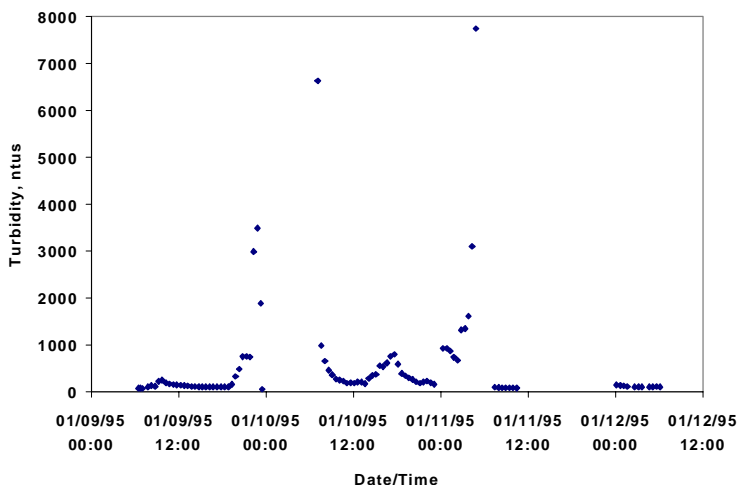


Figure 2: Autocorrelation exhibited in Walters Creek Turbidity data

```

data turbid;
/*
  open raw data file and read turbidity data
  the first record (line) of data is 5
*/
infile 'g:\garry\california\data\wq\tssturb3.sds' firstobs=5;
input mo 4-5 yr 10-11 @18 wal chu;
proc arima;
  identify var=chu nlag=12;
run;
proc arima;
  identify var=wal nlag=12;
run;
    
```

Figure 3: SAS Program to Check for Autocorrelation in Raw Data From Chumash and Walters Watersheds

Figure 4 is the SAS output for the autocorrelation test. Visually, the shape of the plot generated illustrates a strong autocorrelation trend, showing highest correlations at lower lags, or spacing between time periods. Also, since the probability value for autocorrelation to lag 6 (shown under “Autocorrelation Check for White Noise”) is 0.000 (or actually less than 0.0005 but SAS output will only report up to three decimal places), we can reject the null hypothesis of no autocorrelation at a confidence level of almost 100 percent.

Correction for Autocorrelation

When addressing autocorrelation, it is first necessary to recognize the time scale of the experiment or project (e.g., 1 year or 10 years) and determine the most appropriate time step for data collection and analysis (e.g., hourly, daily, weekly, or storm event). For detecting change over the long term, it is usually appropriate to aggregate data within each storm. It may also be appropriate to aggregate data from hourly to daily or daily to weekly. This averaging of data into larger time steps is often termed “data reduction.” Normally, this aggregation reduces or eliminates the autocorrelation; however, data reduc-

tion is not appropriate in all cases, particularly for short-term projects.

In the example, since the 30-minute turbidity data collected from Chumash and Walters Creeks were autocorrelated, averages were calculated from paired data for each storm event lasting one to three days, which eliminated the autocorrelation.

If the data used in the analysis are autocorrelated (either the reduced data is still autocorrelated, or it is inappropriate to reduce the data), then PROC AUTOREG should be used to perform the regression analysis. An example SAS code for PROC AUTOREG is given in Figure 9 (beneath the code for PROC GLM, with output not shown). Typically, an autoregressive model with lag=1 is appropriate for weekly or more frequent water quality data.

If correction for autocorrelation has been made, organization of the reduced data should follow the same layout presented earlier in Figure 1.

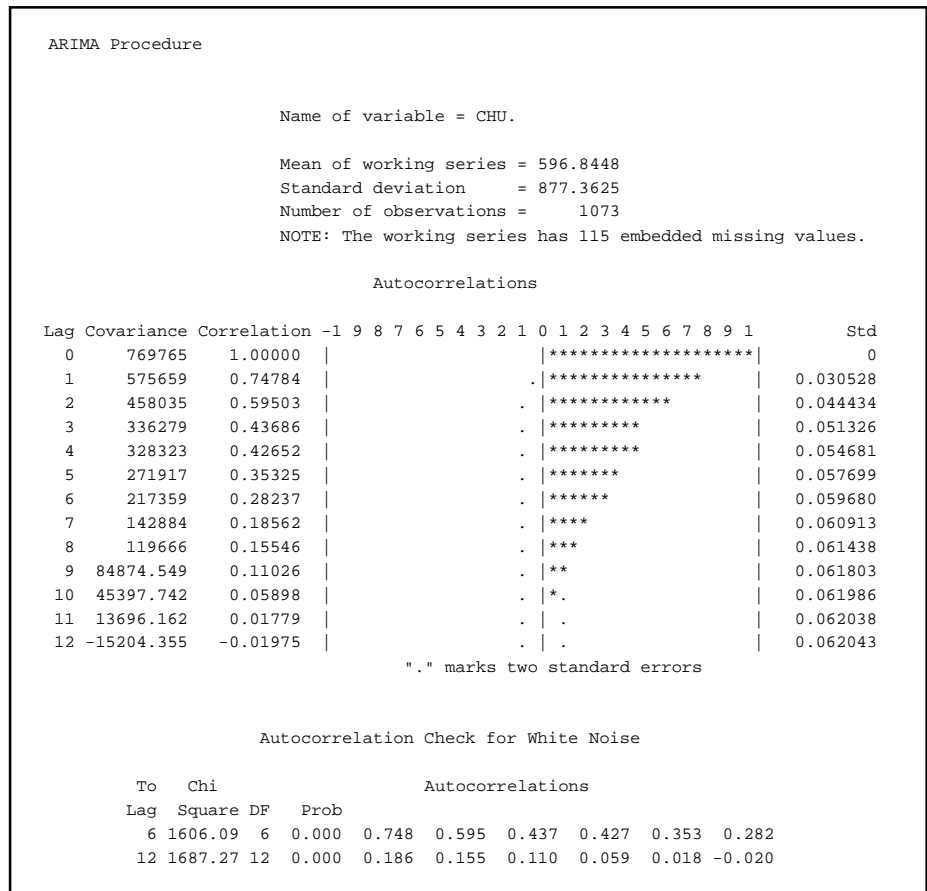


Figure 4: SAS Output Showing Autocorrelation Diagnostics for Raw Turbidity Data from Chumash Watershed

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 5). In most storm event-related water quality data, the skew will be positive reflecting many relatively low values and a few very high values.

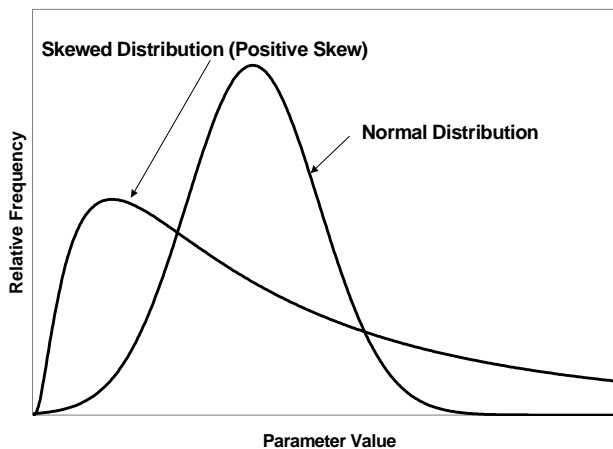


Figure 5: Normal and Skewed Distributions

To check for skewness with SAS, use the “UNIVARIATE” procedure with the “normal” and “plot” options. Figure 6 shows the SAS code for performing this analysis. The program inputs data from an external ASCII file, creates log-transformed variables of the data (which will be compared to untransformed values for correction of skewness), categorizes data into calibration (pre) and treatment (post) periods (which will be used for regression analysis), and tests the data (both untransformed and log-transformed) for normality. July 1, 1996 is the date dividing calibration and treatment periods.

The “normal” option under PROC UNIVARIATE produces descriptive statistics including the skewness and the Shapiro-Wilk W statistic. If the data are perfectly “normal,” the 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. The W statistic ranges between 0 and 1, with low values leading to a rejection of the hypothesis of normality.

```
data turbid;
*open data file-first data on line
number 7;
infile
'g:\california\data\wq\stormturb.sds'
firstobs=7;
*read data;
input mo 8-9 yr 14-15 @16 wal chu;
wallog=log10(wal);
chulog=log10(chu);
*assign pre-post bmp;
if yr>=97 then per='post';
else if yr=96 and mo>=7 then
per='post';
else per='pre';
*increment counter;
N=_N_;
*test for skewness and normality;
proc univariate plot normal;
var chu wal;
run;
*test for skewness of log transformed
variables;
proc univariate plot normal;
var chulog wallog;
run;
```

Figure 6: SAS Program File

The “plot” option under PROC UNIVARIATE produces a visual way of checking for normality. Three graphs will be created for a visual assessment of the distribution of the water quality variable; a stem and leaf plot or histogram, a box plot, and a normal probability chart. Normally distributed data will generate stem and leaf plots and histograms that are symmetrical, box plots that are symmetrical with the mean “+” and median “*-----*” indicators close to each other, and normal probability plots that graph as a straight line “*****” that is superimposed upon the ideal line³ “+++++”.

The SAS UNIVARIATE output for the untransformed turbidity data from Chumash watershed is shown in Figure 7. This output indicates that the data depart markedly from a normal distribution. This can be seen from the high degree of skewness (3.02), relatively low value for W of 0.64, and the probability of the W statistic (0.0001), leading to a rejection of the null hypothesis of normality at $\alpha=0.05$. Also, the highly skewed stem and leaf plot and box plot, and the curvilinear normal probability plot which deviates from the ideal line provide visual evidence of data which is not normally distributed.

³ For further diagnostic details, see SAS Institute Inc. (1985a).

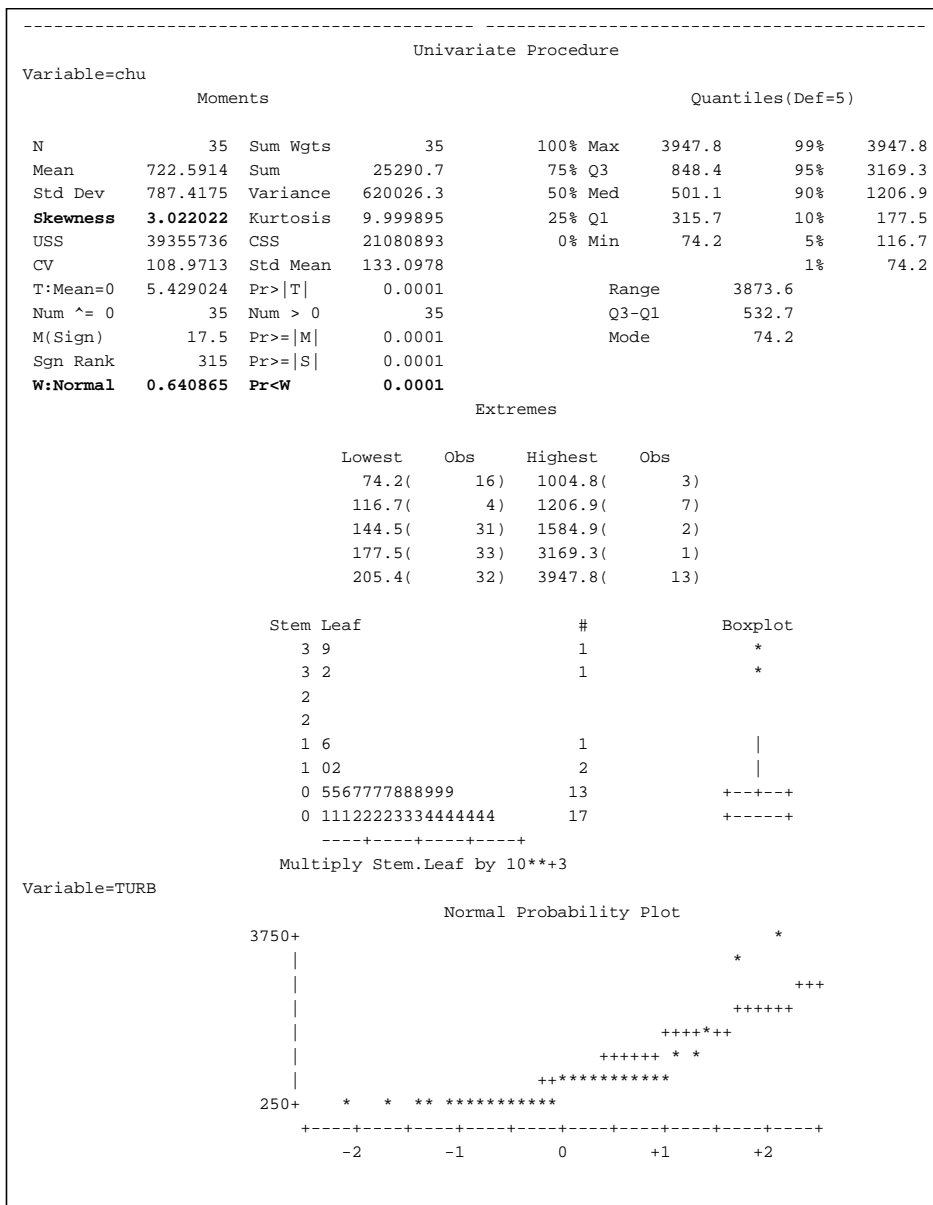


Figure 7: SAS Univariate Output for Chumash (Treatment Watershed) Turbidity Data

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 SAS by using the function “LOG10(x)” (see Figure 6). This will return the base 10 logarithm of the value x. After doing this transformation, repeat the diagnostic steps given previously to check for normality.

If the data remain highly skewed after various transformation attempts, nonparametric data analysis may be required. Nonparametric data analysis is not addressed in this article.

The SAS output for the log-transformed turbidity data from Chumash watershed (Figure 8) indicates that the log-transformed data are normally distributed. This is evidenced by the reduced skewness, relatively high value for W (at $\alpha=0.05$, we do not reject the null hypothesis of normality since $p > \alpha$), greater symmetry of the stem and leaf and box charts, and straighter normal probability plots that fall closer to the ideal line. Therefore the log-transformed variables will be used in the subsequent analyses.

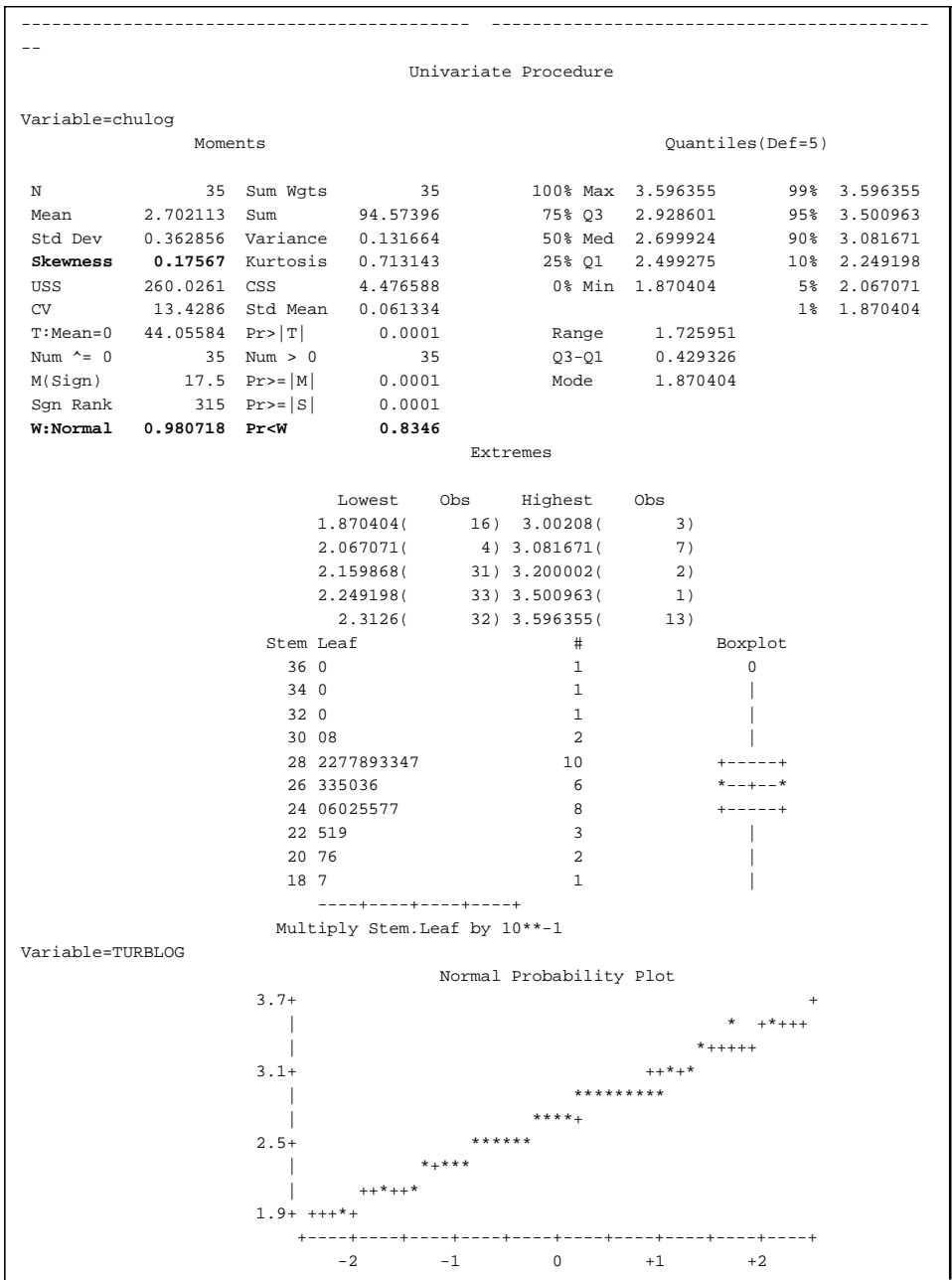
Regression Analysis to Detect Change

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.

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.



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.

In order to perform regression analysis in SAS, two additional independent (X) variables are required to develop and compare the two separate regression lines. The first is a class variable PER, which separates the observations into calibration “pre” and treatment “post” periods and allows the regression lines to have different intercepts. In our example, this variable is generated in Figure 6. The second independent variable is the interaction term PER*WALLOG, and is the class variable multiplied by the water quality parameter from the control site, which in our example is log-transformed turbidity from Walters Creek watershed. This variable allows the two regression lines to have different slopes. WALLOG is also generated in Figure 6.

The ANCOVA can be done with the General Linear Models Procedure (PROC GLM)⁴ in SAS. Figure 9 shows the SAS statements required. In the “model” statement, the control watershed variable (WALLOG) and the two created variables (PER and PER*WALLOG) are the independent or X variables, while the treatment watershed variable (CHULOG) is the dependent or Y variable. The “solution” option in the “model” statement generates the parameter estimates (β’s) for the regression lines.

Figure 8: SAS Univariate Output for Log-Transformed Chumash (Treatment Watershed) Turbidity Data

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.

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 Y values on the regression lines. Typically, if water quality data have been log transformed, the residuals from a regression analysis are normally distributed and have a constant variance.

⁴ For a full description of the GLM procedures, options, and uses, see SAS Institute Inc. (1985b).

```

proc glm;
class per;
model chulog=wallog per per*wallog/solution;
lsmeans per/pdiff;
output rstudent=res2 p=pred2;
run;
proc gplot;
title1 'Residual Plot-Regression Model';
plot res2*pred2/vref=0;
run;
plot res2*n/vref=0;
run;
/*
For the case of regression analysis with autocorrelated
residuals, the following code should be used in place of that
above. The interaction term in the model statement called
“cross” was generated previously by multiplying “wallog”
by “per” since PROC AUTOREG does not allow creation of
interaction terms in the model statement.
*/
proc autoreg;
model chulog=wallog per cross/nlag=1;
output r=res;
run;
*plot residuals to see if autocorrelation removed;
proc gplot;
title1 'AR(1) Model Standardized Residuals';
plot res*n/vref=0;
run;

```

Figure 9: SAS Program for Calibration and Treatment Regression Lines and Residual Check

Constant variance and autocorrelation of the residuals can be checked in SAS using the GLM and GPLOT procedures. In Figure 9, an output statement is included under GLM to generate standardized residuals “rstudent” and predicted Y values “p” for plotting of the residuals under PROC GPLOT⁵ (“plot res2*pred2/vref=0”). A non-constant variance will show up in the residual plot as a “funnel” or “mounding” rather than an even distribution. Figure 10 shows that the residuals, plotted against the predicted Y variables, have a constant variance. Also, since these data were aggregated to a storm basis, autocorrelation is not a concern. If autocorrelation were a concern, it would be necessary to plot the residuals over time (“plot res2*n/vref=0” statement in Figure 9) and visually check for autocorrelation. Once again, repeated trends or patterns are an indication of autocorrelation. Finally, normality of the residuals can be checked in SAS using PROC UNIVARIATE as discussed in the section *Check for Normality*.

The SAS output from the regression analysis for the example is given in Figure 11, General Linear Models

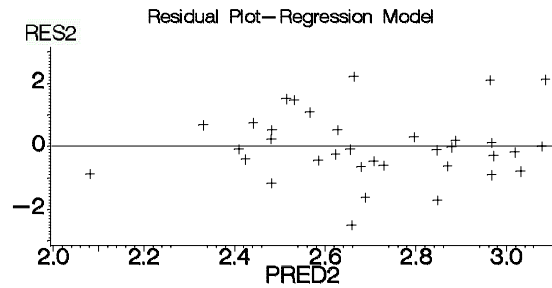


Figure 10: Residual Plot of Regression Model

Procedure. The regression equation resulting from this analysis is

$$Y = \beta_0 + \beta_1 X_1 + \beta_{2j} + \beta_{3j} X_1 \tag{1}$$

or, rearranging the terms to ease interpretation

$$Y = (\beta_0 + \beta_{2j}) + (\beta_1 + \beta_{3j}) X_1 \tag{1a}$$

where Y is the estimate of the dependent variable, X_1 is the independent variable, and $\beta_{i,j}, i=0,1,2,3; j=pre,post$ are the regression coefficients generated from the regression analysis, with

β_0 = y-intercept of the calibration regression line

β_1 = slope of the calibration regression line

β_2 = difference in y-intercept between calibration and treatment lines

β_3 = difference in slope between calibration and treatment lines

The values of β_{2j} and β_{3j} depend on the period j, which in this case is either “pre” or “post” (see Figure 11). SAS internally orders the elements of the class variable PER (“pre” and “post”) either numerically or alphabetically, and then assigns 1 to the first variable (“post” since alphabetically it comes before “pre”) and 0 to the next variable (“pre”). This is why zeros are generated for the Estimate values in Figure 11 for β_2 and β_3 for the “pre” period.

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

⁵ PROC GPLOT produces a graphical plot, while PROC PLOT produces the same plot, but in a text format.

General Linear Models Procedure					
Dependent Variable: CHULOG					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	1.93669600	0.64556533	7.88	0.0005
Error	31	2.53989206	0.08193200		
Corrected Total	34	4.47658805			
	R-Square	C.V.	Root MSE	CHULOG Mean	
	0.432628	10.59310	0.28623767	2.70211313	
Source	DF	Type I SS	Mean Square	F Value	Pr > F
WALLOG	1	1.00493313	1.00493313	12.27	0.0014
PER	1	0.65570489	0.65570489	8.00	0.0081
WALLOG*PER	1	0.27605797	0.27605797	3.37	0.0760
Source	DF	Type III SS	Mean Square	F Value	Pr > F
WALLOG	1	1.48117822	1.48117822	18.08	0.0002
PER	1	0.37954313	0.37954313	4.63	0.0393
WALLOG*PER	1	0.27605797	0.27605797	3.37	0.0760
Parameter	Estimate	T for H0: Parameter=0	Pr > T	Std Error of Estimate	
INTERCEPT β_0	1.810505768 B	3.38	0.0020	0.53620060	
WALLOG β_1	0.414947950 B	1.96	0.0588	0.21151378	
PER β_2 post	-1.912058740 B	-2.15	0.0393	0.88837728	
pre	0.000000000 B	.	.	.	
WALLOG*PER β_3 post	0.630454023 B	1.84	0.0760	0.34346308	
pre	0.000000000 B	.	.	.	
Least Squares Means					
	PER	CHULOG LSMEAN	Pr > T H0: LSMEAN1=LSMEAN2		
	post	2.58797266	0.0086		
	pre	2.87805029			

Figure 11: SAS Output from Regression (ANCOVA) Model

slopes of the lines are examined. The *statistical significance* of a difference is revealed in the $Pr > T$ values, while the estimate of the *magnitude* of the difference is revealed in the *Estimate* values. So when interpreting the GLM output, it is important to first check the $Pr > T$ values to see if there is a difference, and if so, then look at the *Estimate* values of the β s for an estimate of the difference.

The SAS Output (Figure 11) indicates that for the calibration period, the y-intercept (β_0) of the regression line is 1.81 and the slope of this line (β_1) is 0.41.

Looking at β_2 (post), the $Pr > T$ 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 *Estimate* 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 β_3 (post), the $Pr > T$ value of 0.076 indicates that a statistically significant difference in slopes of the regression lines exists at a 92% confidence level. The *Estimate* 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 calibration 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 12, 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 β_3 *Estimate* 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* (PER*WALLOG) term should be dropped from the regression model. In this case, Equation 1a would now be

$$Y = (\beta_0 + \beta_{2j}) + \beta_1 X_1$$

which would result in regression lines with the same slope (parallel) but with y-intercepts differing by a magnitude of β_2 (see Figure 13).

In this scenario (Figure 13) with the interaction term left out, the $Pr > T$ value (0.01) of β_2 (the y-intercept term) indicates that the difference in intercepts is statistically significant at a 99% confidence level. The Coefficient value for β_2 is -0.29 (1.212 - 0.920, Figure 13) which once again indicates a reduction in turbidity of the treatment watershed relative to the control watershed (SAS output not shown).

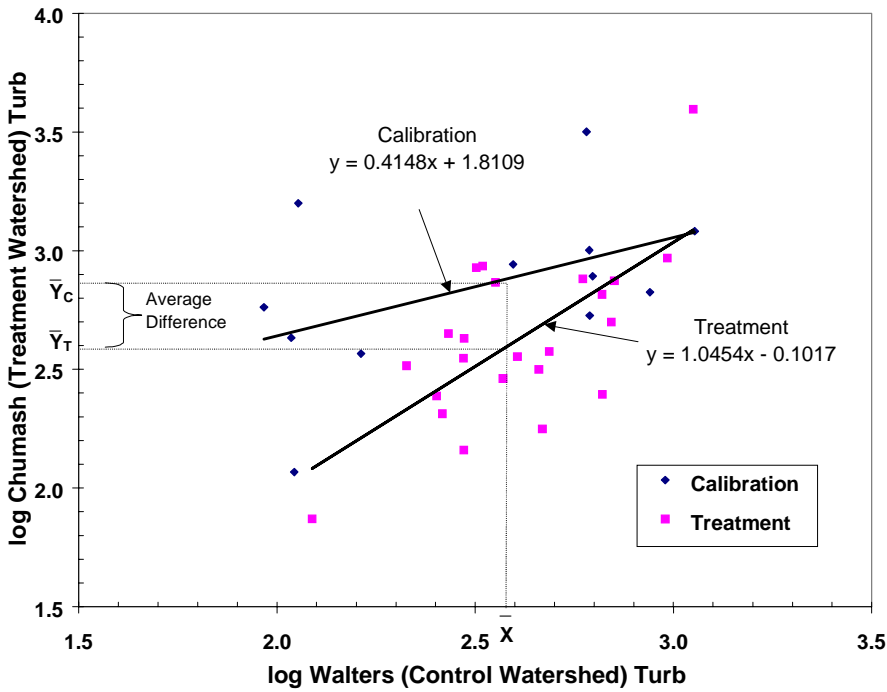


Figure 12: Calibration and Treatment Period Regression Lines for Full Model

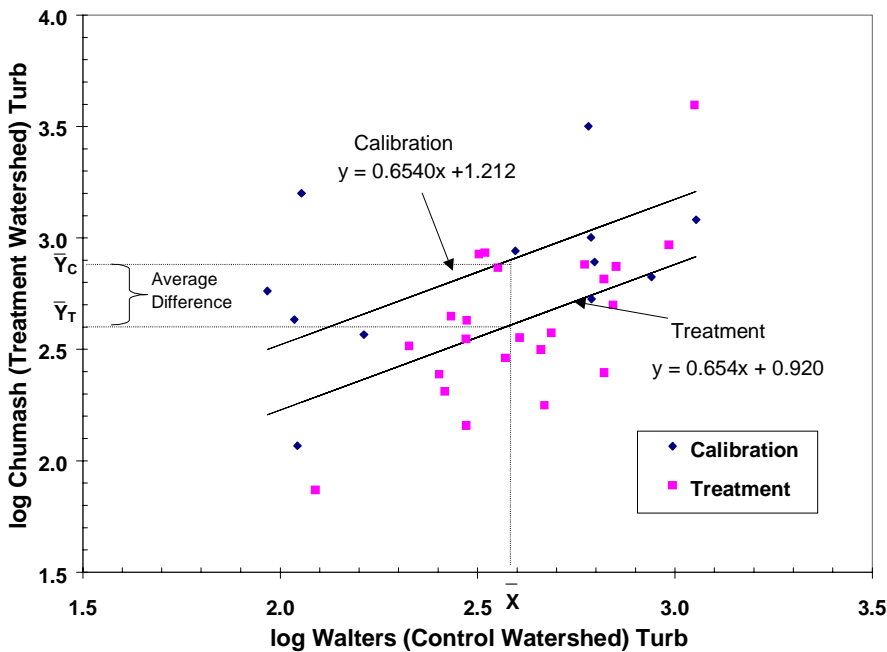


Figure 13: Calibration and Treatment Period Regression Lines for Reduced Model

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 may be obtained by using the least square means “lsmeans” command in PROC GLM (for the “full model” with interaction term, see Figure 9 for code and Figure 11 for output). This produces predicted values for the treatment watershed (Chumash) for the calibration and treatment periods at the average log-transformed value of turbidity for the control watershed (Walters) for the entire period (both calibration and treatment). By using an X value averaged over the entire study period (\bar{X} in Figure 12) to calculate the average difference in Y value, changes in turbidity at the control watershed occurring over the study period will be accounted for.

As can be seen in Figure 11 under the “Least Square Means” output, for the example using the full model with interaction term, the predicted log-transformed turbidity values for Chumash, the treatment watershed, are 2.878 for the calibration period and 2.588 for the treatment period (\bar{Y}_C and \bar{Y}_T respectively, in Figure 12). The following equation is then used to determine the percent decrease in the original, untransformed scale:

$$\left[1 - \left(\frac{10^{\bar{Y}_T}}{10^{\bar{Y}_C}}\right)\right]100$$

or

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

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

$$\left[1 - \left(\frac{10^{2.602}}{10^{2.894}}\right)\right]100$$

or 49%.

For More Information

Garry Grabow, Ph.D., P.E.
Water Quality Extension Specialist
NCSU Water Quality Group
Campus Box 7637
Raleigh, NC 27695-7637
(919) 515-3723
garry_grabow@ncsu.edu

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 Spreadsheet for Statistical Analysis of Paired Watershed, Upstream/Downstream and Before/After Monitoring Designs*. U.S. EPA-NCSU-CES Grant No. X825012.

SAS Institute Inc. 1985a. *SAS® Procedures Guide for Personal Computers*, Version 6 Edition. Cary, NC: SAS Institute Inc., 373 pp.

SAS Institute Inc. 1985b. *SAS/STAT™ Guide for Personal Computers*, Version 6 Edition. Cary, NC: SAS Institute Inc., 378 pp.

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.

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 appreciation to Dr. Francis Giesbrecht, Professor of Statistics, North Carolina State University, for his review and contributions.

■

INFORMATION

STORET Update Released

An update to the STORET water quality database system is now available. STORET, the EPA's main repository for marine, freshwater, and biological monitoring data, is free to users and easy to use. It is available from EPA on CD ROM by calling 1-800-424-9067 or sending an email to STORET@epa.gov. Data stored in STORET will be accessible to the public on the Internet early in 1999.

Toward Sustainable Growth for the 21st Century

The Chesapeake Bay Program has created a video detailing six techniques to curtail sprawl development patterns. The video discusses urban boundaries, infill development, transit-oriented development, transfer of development rights, rural clustering, and traditional neighborhood development. To order a copy or to request more information, contact the Chesapeake Bay Local Advisory Committee, 416 Goldsborough Street, Easton, MD 21601, Tel: 410-822-9630, Fax: 410-820-5039.

Low-Impact Development Design Manual Available

Prince George's County, Maryland, has released a manual on low-impact development, which uses micromanagement-level planning techniques to incorporate stormwater BMPs into landscaping plans for each developed parcel. This approach maximizes environmental protection from development impacts through reduction of clearing, use of existing grading, and use of forest- and habitat-enhancing techniques to protect ground water, streams, floodplains, and wetland areas. It combines an environmentally sensitive and functional site design with active public outreach and education, water conservation and reuse, and public/private partnerships. To order the \$35 manual, contact Prince George's County Government, DER Programs and Planning, Attn: Larry S. Coffman, 9400 Peppercorn Place, Suite 600, Largo, MD 20744.

■

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>

Help for Grant Writing

The new EPA Grant-Writing Tutorial, available at <http://www.epa.gov/grtlakes/seahome/grants.html>, can make the grant-writing process easier: the tool walks users through the process of grant-writing and helps them learn how to write more competitive grants.

Source Water Protection Ordinances

Model ordinances for source water protection can be viewed at <http://www.epa.gov/r10earth/offices/water/swp.htm>. Currently the list includes ordinances from:

- New Hampshire Office of State Planning
- Massachusetts Department of Environmental Protection
- Oregon, Lancaster County (PA)
- Oklahoma Department of Environmental Quality
- Skagit County (WA)
- Hernando County (FL)
- Brevard County (FL)

“Antidotes to Sprawl” Website

EPA Region 5 has created a website listing Federal agencies and their role in assisting local communities with sustainable urban development alternatives. Assistance is provided in the form of grant funds, data and information, and technical assistance. The Federal agency contacts are organized by environmental and metropolitan issues so municipal officials can refer quickly to their own specific concerns. The site’s address is <http://www.epa.gov/region5/sprawl/>.



MEETINGS

Notice of Training Event

Getting in Step: A Pathway to Effective Outreach in Your Watershed. Feb 7-10, 1999, Building Capacity in Environmental Community-Based Watershed Projects conference in Stevenson, WA.

This workshop promotes a step-by-step approach to planning and implementing outreach, education and public involvement programs for watershed protection to achieve significant results with limited resources. The building blocks for such a process include defining the program objective, identifying the audience, developing the message, selecting a format, identifying distribution venues, and evaluating the results. The use of innovative designs, graphics, photos, “hooks,” and textual materials will be explored for instilling a conservation ethic among various publics.

The latter part of the workshop focuses on partnering with the news media to enhance outreach, education and public involvement efforts. Approaches for increasing coverage of watershed issues, remediation projects and volunteer monitoring programs are examined in detail. A final section on environmental communication provides agency public information personnel and watershed group representatives with insights on public perception issues, how media interviews are conducted, and offers tools and techniques to employ when dealing with reporters or appearing in televised news programs.

Four additional workshops are tentatively scheduled for April (Seattle), May (Fayetteville, AR; Austin, TX), and September (Lafayette, IN). For more information on either attending or scheduling a workshop, contact Charlie MacPherson of TetraTech, Inc. or Barry Toning of the Council of State Governments. Telephone Charlie MacPherson at 703-385-6000, or email at macphch@tetrattech-ffx.com. Telephone Barry Toning at 606-244-8228 or email at bttoning@csg.org.

Call for Papers

Conference on Stormwater and Urban Systems Modeling: February 18-19, 1999, Toronto, Ontario. Abstracts (due January 30) are solicited on the use of state-of-the-art computer models for resolving real water pollution problems. Contact Lyn James, Computational Hydraulics, Int. 86 Stuart Street, Guelph, ON, Canada N1E 4S5. Tel: 519-767-0197, Fax: 519-767-2770, email: info@chi.on.ca, web site: <http://www.chi.on.ca>

4th International IAWQ Conference on Diffuse Pollution: January 16-20, 2000, Bangkok, Thailand. Submit abstract (500 words maximum) by April 15, 1999 to Ms. Nitayaporn Tonmanee, Department of Land Development, Phaholyothin Road, Chatuchak, Bangkok 10900, Thailand, Fax: 662-579-4430 email: ldd@mozart.inet.co.th, website: <http://www.ddd.go.th/iawq.htm>. In North America, contact Dr. Mike Burkart, USDA, at burkart@nsl.gov for information.

This conference will cover causes, impact, prevention and abatement of diffuse pollution in urban and rural areas and watershed management. In addition to topics widely recognized in developed countries, the conference will focus on worldwide problems of deforestation, land use conversion, and problems of drainage and diffuse pollution in megacities.

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>

FEBRUARY

The 1999 North Carolina Environmental Education Conference: February 10-12, Research Triangle Park, NC. Contact Judy Pope, Office of Environmental Education, P.O. Box 27687, Raleigh, NC 27611-7687. Tel: 919-733-0711.

Conference on Stormwater and Urban Systems Modeling: February 18-19, Toronto, Ontario. Contact Lyn James, Computational Hydraulics, Int. 86 Stuart Street, Guelph, ON, Canada N1E 4S5. Tel: 519-767-0197, Fax: 519-767-2770, email: info@chi.on.ca, web site: www.chi.on.ca

MARCH

Wetlands Engineering and River Restoration Conference: March 22-27, Denver, CO. Contact ASCE, Conferences and Expositions, P.O. Box 832, Somerset, NJ 08875-0832. Tel: 800-548-ASCE within the U.S., and 703-295-6050 outside the U.S. Fax: 703-295-6333.

EPA Region 7 Seventh Annual Nonpoint Source Conference and the Iowa State University Conference on Global Water Quality Issues: March 24-26, Ames, IA. Contact Richard Larson, Agribusiness Education Programs, Iowa State University, Tel: 515-294-6429,

email: rwlarson@iastate.edu, website: http://extension.agron.iastate.edu/aged/water_quality/wqconf.html

Fundamentals, Modeling, and Applications of Nitrification and Denitrification: March 28-31, Roanoke, VA. Contact Conference Registrar, 810 University City Blvd, Suite D, Mail Code 0272, Virginia Tech, Blacksburg, VA 24061, Tel: 540-231-5182, Fax: 540-231-3306, website: <http://www.conted.vt.edu/nitrification.htm>

MAY

International Conference on Diffuse Pollution: May 16-21, 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.environ.wa.gov.au>

JUNE

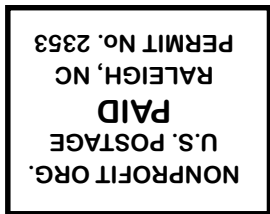
26th Annual Water Resources Planning and Management Conference: Preparing for the 21st Century: June 6-9, Tempe, Arizona. ASCE Conferences, 1801 Alexander Bell Drive, Reston, VA 20191-4400. Tel: 1-800-548-2723 (ASCE) or 703-295-6300, Fax: 703-295-6144, email: conf@asce.org, web site: <http://waterqg.asce.org>

3rd National Workshop on Constructed Wetlands/BMPs for Nutrient Reduction and Coastal Water Protection: June 9-12, New Orleans, Louisiana. Contact Dr. Frank Humenik, Box 7927, North Carolina State University, Raleigh, NC 27695-7927. Tel: 919-515-6767, Fax: 919-513-1023, email: frank_humenik@ncsu.edu, web site: http://www.cals.ncsu.edu/waste_mgmt/workshop.html

2nd Annual Mitigation Banking Conference: June 13-15, 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>

■

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>



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

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