Statistical Consulting Services
3405 Chukar Lane, Clarkston, Washington 99403 (509) 758-8330
stat@cableone.net

• Design • Analyze • Report

Status: Preliminary Report
Date: June11, 2009
Project Title: Water Quality: 2008 Data, BPA-51

2008 Water Quality

I) Data

Print Out #1 presents a listing of the initial data. The variables included were: *SITE*, *REP*, *NH4*, *NO2_3*, *SRP*, *TDP, TN*, *TP*, and *JULIAN* , representing site code, replication number, ammonia nitrogen, nitrate and nitrite nitrogen, soluble reactive phosphorus, total dissolved phosphorus, total nitrogen, total phosphorus, and Julian date, respectively. All nutrient values are recorded as g/L.

Thirteen sites were included: KR1, KR2, KR3, KR4, KR6, KR7, KR9, KR9.1, KR10, KR11, KR12, KR13, and KR14. Each site was represented by up to six replications over the dates of April 26, 2008 to September 30, 2008. The data contained 383 observations.

II) Summary Statistics and Trends for Water Quality Variables

Mean, variance and the number of observations for the response variables are presented in Print Out #2. All computations were carried out separately for each site. In what follows, each response variable will be discussed in succession. Further detailed summary statistics are provided in Print Out #3.

Ammonia Nitrogen (NH4)

The number of observations over time for ammonia varied from 21 to 36 observations. Mean values ranged from 3.56 to 5.88 g/L. The larger value was evident in the site KR14 with lower value found in KR9.1. Sample variances were higher at KR2 (23.99) and minimal at KR9.1 (3.49). The higher value may be due to more <u>extreme</u> values at KR2 (see Print Out #3).

Nitrate and Nitrite (NO2_3)

Observation numbers for NO2_3 varied from 21 to 36. Mean values were higher than ammonia and ranged from 82.2 to 153.9 g/L for sites KR14 and KR3, respectively. Variances were <u>substantially higher</u> in some cases, ranging from 148 to 15573, at sites KR7 and KR3, although most sites had variances which were smaller than 400.

Soluble Reactive Phosphorus (SRP)

As seen in the other variables, SRP was available in 21 to 36 values across time. Mean values were lowest (1.60   g/L) at sites KR11 and KR13 and largest at site KR7 (1.29   g/L). Variability ranged from 0.00 (KR11, KR13) to 5.79 (KR7).

Total Dissolved Phosphorus (TDP)

The number of observations ranged from 21 to 36. Mean values were lowest (1.60   g/L) at site KR11 and largest at site KR1 (2.74   g/L). Variance estimates ranged from 0.56 at KR9.1 to 5.69 at KR7.

Total Nitrogen (TN)

Sample sizes ranged from 21 to 36 observations for this response variable. Mean TN values were 179.5 at KR14 to 286.3   g/L at KR3. Levels of variability showed a very wide range across sites from 637 at KR4 to over 54424 at KR3.

Total Phosphorus (TP)

Like the other responses, the number of observations for the TP variable ranged from 21 to 36. Means showed a fair amount of variability across sites with a low value of 5.5   g/L at KR13 to a high value of 61.8   g/L at KR3. These differences appeared again in the variance measures where TP variances ranged from 11.0 at KR10 to 25930.6 at KR3. As with other nutrient variables, this is probably due to extreme values at some sites (see Print Out #3).

Trends Over Time and RKM

Plots of the nutrient variable trends for each site across the sampling periods are given in Print Out #4. These plots provide both the mean trend (solid lines) and a vertical bar representing two standard deviations for each time period or RKM. In addition, a red reference line is provided to indicate the Julian date (166) or RKM (263.8) at which the nutrient addition treatments were begun in 2008. The presence of trends varies from site to site as well as among variables. For most responses, the variability was high, making definitive trends very difficult to detect. Some trends for TP, TN, and NO2_3 may be evident over time and RKM, however, most trends are obscure due to variability.

III) Determination of Sample Sizes

The formulation for calculating sample size is:

$$n = (z*s/d)^2$$

where s, d and z are related to the variability, desired precision, and confidence levels, respectively. As a means of increasing the number of observations available to estimate the variability, s, it was necessary to pool the data across sampling periods. A value of z=1.96 was assumed to provide a confidence level of 95%. The desired **precision** levels, d, were set to values of 10% of the overall mean where feasible.

The resulting sample sizes are given in Print Out #5. This shows a range of sample sizes across responses and sites. Overall, all the responses show reasonable results for the target sample size of 6 replications. However, some sample size estimates for sites KR3 and KR7 are higher. This may be due to a higher degree of variability and the respective sample sizes reflect this (see individual response discussions above and Print Out #3). Other combinations also result in relatively higher sample size estimates ranging from n = 9 to 12. In such cases, there is an indication of high variability associated with the responses and, therefore, it is recommended that sample sizes be kept at higher levels for those sites. For the remaining sites, while some sample sizes are estimated to be small, i.e. 1 or 2, any actual sample sizes used should be kept at the nominal levels of 6 replicates to ensure that adequate information is available for more advanced statistical procedures which may be carried out in the future.

Print Out #6 presents a comparison of sample sizes from 2002 to 2008 water quality data. For the sites in common across these years, **the results are relatively similar in the later years** indicating consistent variability in the respective responses for 2008. With the exceptions noted above, it would appear that the observations taken within each channel **were effective** in reducing the average variability allowing the precision level to be maintained at 10% of the response means.

It should be emphasized that for all the calculations shown, the resulting sample size values are approximate. Therefore, policy regarding future sampling protocols should not rely on this information alone. That said, SCS would recommend that sample sizes be maintained at a level between **5 and 10 units** (minimum of 5) for each sampling site and time. This should provide sufficient levels of information for the continued long term monitoring and analysis of most water quality responses.

IV) 2002 - 2008 Water Quality Comparisons

Print Out #7 presents overlaying plots of each nutrient versus Julian date or RKM from 2002 through 2008. As with the plots presented earlier, the red reference line indicates the RKM at which the 2008 nutrient addition treatment was initiated. While nutrient trends do differ across years, some similarity among years is evident. In addition, certain nutrients (NO2_3) show a trend towards higher levels across years, although this pattern is present at all RKM. This could be due to annual differences in ecological factors, such as nutrient decomposition/depletion rate, and physical factors such as temperature and water flow. Any future analyses utilizing water quality data should consider these factors.

V) Additional Remarks for 2008 Water Quality Data

1) The 2008 water quality data received by SCS required considerable manipulation and data management prior to analysis. If it is anticipated that water quality data received by SCS in the future will be of the same format, the time to carry out the necessary reformatting of the data should be taken into consideration.

2) The levels of SRP from water quality data of previous years were often below detection limits.

The data from 2008 showed elevated levels for this and other responses. This pattern was seemingly unrelated to nutrient addition treatments, however, as they appeared consistently across the study area.

3) Considerable variability was observed for some responses in site KR3 and KR7 in 2008. The large values causing this variability should be examined for accuracy or potential contamination before more detailed analyses are carried out on this data.

4) The river fertilization program was begun in 2005. Because the procedures for detection of nutrients and metals are quite sensitive, SCS recommends that any future water quality samples taken on, or close to, the dates of fertilizer application be carried out with the utmost care to avoid contamination issues. Doing so will ensure consistency and reliability in the resulting data.