Retrospective Analysis of Big Thompson Water Quality, 2000 – 2006

Volume 2: Appendices

(THIS PAGE INTENTIONALLY LEFT BLANK)

Appendix A: Statistical Methods: Box-and Whisker Plots

and Seasonal Regression on Ranks

Box-and-whisker plots

In this type of plot, also known as "boxplots" each box and set of whiskers corresponds to one group, or in this case, one month. The box corresponds to the middle 50% of the data in that group. The bottom of the box represents the 25th percentile; the horizontal line within the box indicates the median or 50th percentile; and the top of the box represents the 75th percentile. The whiskers extend to representative extreme values of the group, which can be the minimum and maximum, the 5th and 95th percentiles, or other values.

Minitab extends the upper whisker to the highest value within an upper limit defined as

Upper limit = Q3 + 1.5 (Q3 − Q1)

Where Q3 is the third quartile or 75th percentile and Q1 is the first quartile or 25th percentile.

The lower whisker extends to the lowest value within a lower limit defined as

Lower limit = Q1− 1.5 (Q3 − Q1)

Note that the distance (Q3-Q1) is equal to the length of the box. Individual values beyond the whiskers can be individually plotted or not as the user chooses.

Seasonal regression on ranks

The regression model used to test for trend in the BTWF data is

Rank (concentration) = B0 + B1*I1 + B2*I2 + B3*(Sample Date)

Where

Rank(concentration) = the rank of the observed concentration or other water quality measurement, with 1 corresponding to the smallest observation, 2 the next smallest etc.

> The rank is used rather than the actual measurement in order to obtain a test for trend that is not greatly affected by outliers and does not depend on the data being normally distributed. This is very important for water quality data, which tend to contain outliers and to be non-normally distributed. In some cases, a log transformation might be appropriate to achieve normality, but since a rank transformation is always appropriate, it is used throughout the present study.

For consideration in future studies, it would be possible to add additional predictive variables, such as streamflow or precipitation, into the regression equation in an attempt to better explain the causes of apparent trends.

Seasonal indicator variables

I1 and I2 are indicator variables that take on a value of 0 or 1 to indicate the season for each observation. Three seasons are used for the BTWF data. The breakdown of seasons and corresponding values of the indicator variables is shown in the following table.

Months	Season #	I1	I2
May - July	1	0	0
August - October	2	1	0
Nov – April	3	0	1

B0, B1, and B3 are regression coefficients estimated by Minitab.

The test for statistical significance

The only regression coefficient that is of particular interest in this study is the time trend slope, B3 in the above model. The regression analysis by Minitab or other software computes a **p-value** for B3 (as well as for the other coefficients), and this p-value may be roughly interpreted as the probability of obtaining a slope as large or larger than the computed value if in fact there were no trend in the data. One concludes that there is a statistically significant trend present if the computed p-value is smaller than some critical value.

For this study a critical value of 0.10 is used. So if the computed p-value for B3 for a particular variable and location is less than 0.10, then one concludes that a significant time trend exists. Otherwise, one concludes that no significant trend exists. In statistical terminology, the test for trend is performed at a 90% confidence level since a critical value of $p = 10\% = 0.10$ is chosen. It is also common to use a critical value of $p = 0.05$, which corresponds to a 95% confidence level. The 90% test is better able to detect smaller trends than is a 95% test, but the 90% test has a slightly higher probability of falsely detecting trends that do not really exist.

Estimating the trend magnitude

Once a trend is deemed statistically significant, the question remains of how large the trend is. Unfortunately the rank regression model does not provide this information, since it uses ranks rather than actual observations. There are many ways to estimate the trend magnitude, but the simplest is to perform the regression again using the actual observations rather than the ranks. When the actual observations are used, the magnitude of B3 will be the time trend magnitude or slope which will generally have units of (mg/L) per day. This approach provides a valid least-squares estimate of the trend, no matter what the distribution of the data, but it is significantly influenced by outliers, meaning that one or two extreme values can markedly affect the slope estimate. However, this problem is easy to spot by plotting the data against time with the regression trend line and can be handled by repeating the estimate with the outliers removed if necessary.

Limitations of trend analysis

This regression approach has the advantage of providing a nice yes/no answer to the question of whether or not an apparent trend is really statistically significant as opposed to being the result of a chance arrangement of observations. However, this and other statistical approaches to trend analysis definitely have limitations, including the fact that they are more likely to detect significant trends as the number of observations in the historical record increases. Thus if the number of observations is very small, a practically significant trend will not be detected as statistically significant. Conversely, as the number of observations becomes very large, eventually even very small trends will be determined to be statistically significant. Therefore, the regression test should always be combined with a visual inspection of the data or annual boxplots to see if there appears to be a trend that is large enough to be of interest or importance. The rule of thumb is “if you can’t see a trend in the data, it is probably not there.” Additionally the computation of the trend magnitude should help determine whether or not a trend is of practical importance.

Another important consideration in trend analysis of the type performed here is that a detected trend applies to the period of record only. Nothing can be inferred about the probability that an observed trend in the past will continue into the future.