To the Graduate Council:

I am submitting herewith a thesis written by Tim Pobst entitled "My Thesis." I have examined the final paper copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Environmental Engineering.

Dr.John Schwartz, Major Professor

We have read this thesis
and recommend its acceptance:

Dr. Bruce Robinson

Dr. Qiang He

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

To the Graduate Council:

I am submitting herewith a thesis written by Tim Pobst entitled "My Thesis." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Environmental Engineering.

Dr.John Schwartz, Major Professor

We have read this thesis
and recommend its acceptance:

Dr. Bruce Robinson

Dr. Qiang He

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

My Thesis

A Thesis Presented for
The Master of Science
Degree
The University of Tennessee, Knoxville

Tim Pobst

May 2014

dedication...

Acknowledgements

I would like to thank... Dr. Schwartz, Keil Neff, Matt, and Steve Moore.

I would like to thank all the faculty and staff in the civil and environmental engineering program. I was never lacking in offered help, from Larry and Ken to the Professors and the secretaries.

I would like to thank Dr.Schwartz for giving me the opportunity for a graduate degree. My responsibilities involved a lot of field work in the Great Smokey Mountains and I never bored of bragging about it.

I would like to thank Matt Kulp and Steve Moore of wildlife and fisheries in the GRSM for funding me and any work done in the park.

I would like to thank Mei Cai and Keil Neff who were post-docs for much of the time I was working on my degree. Sometimes it was like having multiple bosses saying different things at once but most of time it was multiple mentors that could help me if one was unavailable.

I would like to thank Chris Rollison and Keil Neff for teaching me everything I need to know to help on the GRSM projects.

I would like to thank Jordan Hayes and Matt Aplin who were the two primary undergraduate helpers regarding GRSM work.

Some quotation...

Abstract

Abstract text goes here...

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Text and tables should show up.

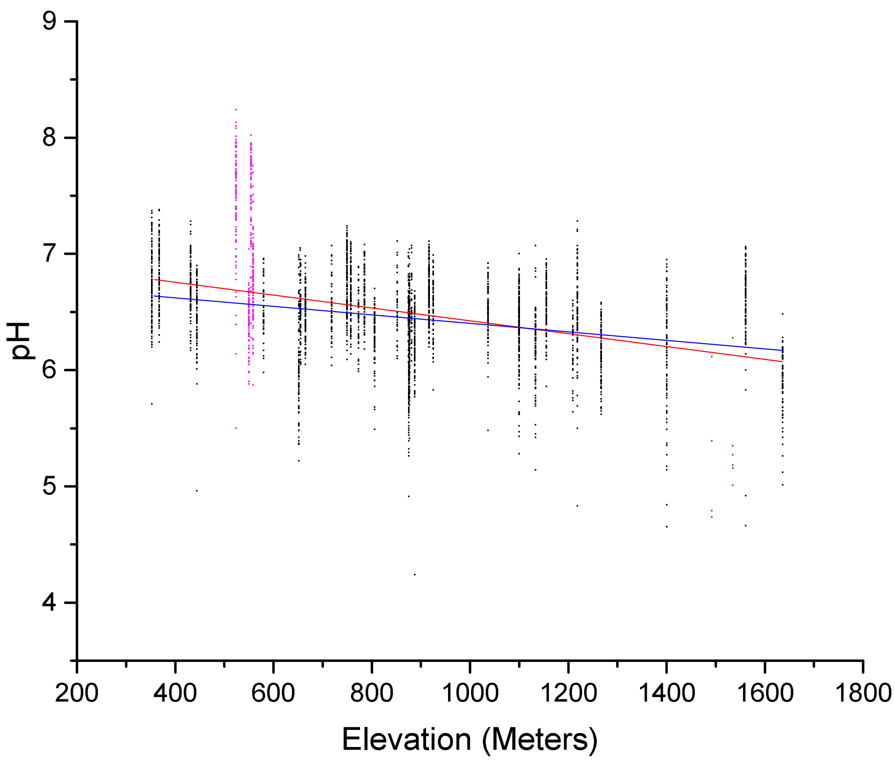


Figure 1.1: pH plotted vs. Elevation. With and without outliers.

Acid rain is believed to negatively affect The Great Smokey Mountain National Park

Acid Deposition, more commonly known as Acid Rain, is a constant problem for the i

Acid Deposition occurs when the emissions of sulfur oxides (SOx), and nitrogen oxi

Acid deposition greatly impacts surface water and the surrounding environment. The

Acidification of bodies of water can be either chronic or episodic. Chronic acidif

The Great Smoky Mountains National Park (GRSM) is located in the southern Appalach

In order to monitor acid deposition the park has a program called the Inventory an

Figure 1-1

This figure shows all pH data from 1993 to 2012 vs. Elevation (m). The red trend

Figure 1 is a graph of all measured pH values for Stream Survey between the years

In support of GRSM natural resource management, stream water quality has been moni

Table 1 shows the current historical elevation classes with the number of sites th

The National Park Service is currently developing a Vital Sign Monitoring Program

Objectives of this study were to:

\begin{itemize}

\item characterize time trends in stream pH and acidic anions among elevation ran

\item characterize sampling variance based on available water quality data, within

\end{itemize}

\begin{itemize}

\item Has stream pH and acid anion concentrations changed among three time periods

\begin{itemize}

\item ANOVA

\item Time trends

\end{itemize}

\item What is the statistical power for water quality parameters based on frequenc

```
\begin{itemize}
\item Post Hoc Analysis
\item A Priori Analysis
\end{itemize}
\end{itemize}
The thesis is organized into two separate chapters following the two above researc
```

Chapter 2

Trend Analysis

2.1 Methods

2.1.1 Introduction

Trend analysis is a great way to characterize the park's water quality using data collected through the stream survey. It is used to state the condition of the parks water bodies while trying to predict where the water quality is headed in the future. A trend analysis on the stream survey data was conducted in 2002 and published in (Robinson et al., 2008) and then again in 2009 for the Biotics Effects report (Cai and Schwartz, 2012). This statistical procedure is used to discover sudden and gradual trends over time. Of the ten elevation bands analyzed in (Robinson et al., 2008) six had negative julian date coefficients and the other four had no trend. Of the 67 sites studied in the biotics effects report most showed no trend, 22 showed an increase in pH, and 2 showed a decrease(Cai and Schwartz, 2012). The trend analysis will use stream survey data from 1993 to 2012 using the statistical programs JMP and SPSS for analysis.

2.1.2 Body

Water quality is an ongoing concern for the park. The acidification of the streams can have significant negative effects on wildlife and vegetation. The stream survey collects water samples all over the park to monitor the health of the water. The pH trends are used to indicate what condition the park is in.

Twenty years of data were available for this paper from the years 1993 to 2012. The data used in these analyses are collected through the park wide stream survey. Most samples are collected every two months and analyzed in a lab for many water quality variables including pH, ANC, nitrate, sulfate and some metals. A single trend line containing all 20 years is unrealistic. The difference in trends from Robinson et al. (2008) and Cai and Schwartz (2012) suggests an inflection point in the trend line somewhere between 2002 and 2009. For this reason and also to be able to easier compare results from this paper with those from Robinson et al. (2008), who used the years 1993 to 2002, a separate data set will be sectioned off from 1993 to 2002. A third data set will be created after the year 2008 because this is the year that the Kingston and Bull run power plants installed scrubbers onto their stacks. The hypothesis being the sulfate concentrations will be noticeably different and this difference could indicate a need for further study. These three time sets will be analyzed separately.

Two more factor divisions of the data include dividing the data by elevation classes and and by four dependent variables (pH, ANC, Nitrate, and Sulfate). The elevation classes used in this paper were set up to contain a minimum number of sites in order that the upper classes would not be too weak to be useful. The divisions are presented here in **??**. These are different from the historical eleven elevation bands which were separated by 500 foot intervals. Some of the upper bands only contained only one site. After years of collection this one site can describe its own characteristics but it cannot describe characteristics of the elevation band very well. Elevation is an important factor in water quality and because the upper elevations are most effected by acid rain there needs to be enough sites in each band to make them statistically

Table 2.1: Elevation Bands

Elevation Bands	Meters (Feet)	n	Site #
1	304.8-609.6 (1000-2000)	5	13 ,23, 24, 30, 479
2	609.6-762 (2000-2500)	9	4, 311, 268, 480, 310, 483, 147, 148, 484
3	762-914.4 (2500-3000)	13	114, 481, 482, 149, 66, 492, 137, 293, 270, 493, 485, 144, 224
4	914.4-1066.8 (3000-3500)	4	143, 142, 73, 71
5	1066.8-1371.6 (3500-4500)	4	74, 221, 251, 233
6	1371.6 < (4500 <)	2	253, 234

strong. Without adding sites the best way to do this is to reorganize the elevation bands. The dependent variables in this study are as mentioned before pH, ANC, Nitrate, and Sulfate. Dividing all the data into three different time sets, six elevation bands and studying four different dependents will create 72 different trend lines.

Instruments

All of the statistical analysis was completed in statistical software. Initial data cleaning and influential data points were found using JMP 10. The power analysis was done using G*power and all other statistical analyses was done using SPSS.

Several plots were created in order to reduce the variance of the data before any statistical analysis was attempted. JMP was chosen for this task based on its ease of use in plotting data. A plot of pH vs. time visually represents the existence of a positive trend in the pH data from 1993 to 2012. Figure 2.1 shows the pH vs. elevation plot. It shows two trend lines, one which represents the trend of all of the data points and the other represents the trend after the influential points are removed. Both of the trends are negative as elevation increases but the trend for all the data is steeper. pH was plotted against the month that the sample was collected to check for seasonality. Seasonality was expected and found for pH over one year. The variance caused by seasonality will be removed with sine and cosine functions.

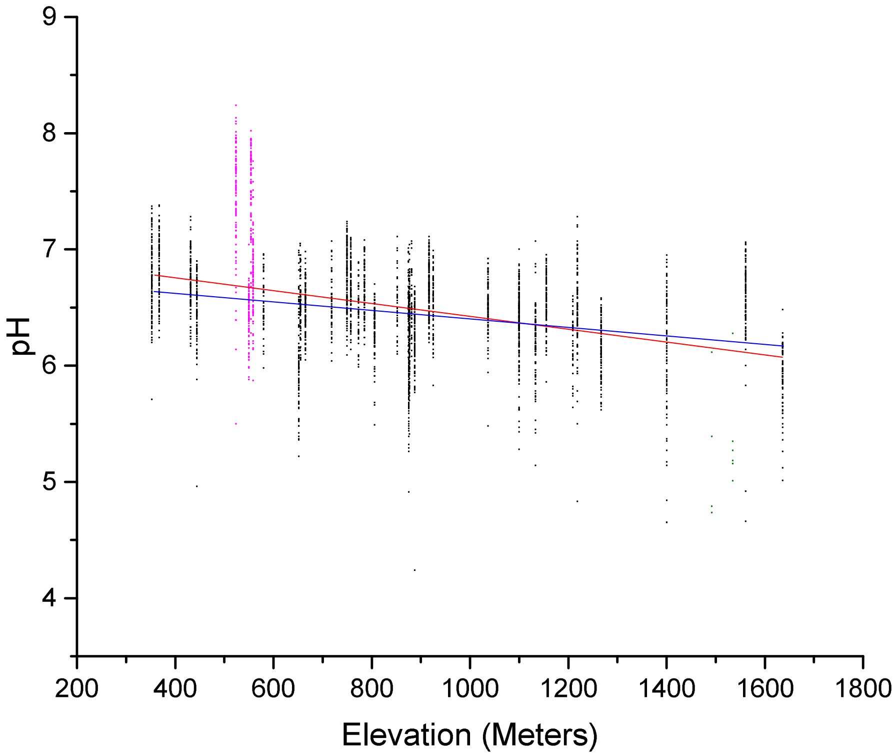


Figure 2.1: pH plotted vs. Elevation. With and without outliers.

Much of the variance in Figure 2.1 can be attributed to known influences of the Abrams creek watershed, sites that are affected by anakeesta geology , and storm flow. Abrams is a lower elevation, low slope area where the underlying geology is much limestone. This limestone contributes to high ANC values which in turn keep the pH levels higher than the res of the sites in the survey. Site numbers 237 and 252 are sites which are down slope of road cuts that have exposed the underlying anakeesta geology to runoff. The high sulfate? concentration from the anakeesta run into the streams and keep their pH values very low. Storm flow is also usually seen as an out lier in past GRSM water quality projects. Storms can bring high intensity rain fall which can very quickly raise the levels of nitrate and sulfate pollution in the streams. The runoff can also carry any pollutants that have previously settled on vegetation or the ground. The lowered pH of the streams caused by the storm flow can cause leeching of the surrounding mineral geology in affected areas. Healthy streams can rebound to normal pH values, unhealthy streams can have lowered ANC due to the leaching. Measurements taken from storm flow can show uncharacteristically low pH values and high amounts of metals. In this way storm flow is sometimes considered an influential group on the rest of the data, because the measurements are significantly different from the average. Dr. Cai characterized all of the available water quality data between 1993 and 2010 as storm flow or base flow, this work is summarized in Cai and Schwartz (2012). Water quality data after 2010 has not been characterized. If all storm flow observations are to be dismissed as influential, the years 2011 and 2012 would need to be characterized. Quick analyses were run to see how influential storm flow was on the data as a whole and it turned out that some were and many were not. Instead of throwing out all of storm flow observations, storm flow became a reason to throw out influential observations during regression analysis. They can be removed on a case by case basis during the regression procedure.

Part of the regression procedure includes preparing the data and identifying influential observations. The output of a step-wise regression analysis done in SPSS can be configured to complete many different analyses in order to clean the data. The different tests employed for this paper include tests for normality, heteroscedasticity, cook's D, DFBETAS, and DFFITS. As observations were identified by cook's d, DFBETAS, and or DFFITS as influential they were then checked in Excel to determine why. Modification of removal of an influential observation had to be justified or it would remain an outlier. An example of modification of the data included a pH value that read 16.47 was changed to 6.47. Another example is that some conductivity values were obvious copies of the ANC column which was right next to it. These conductivity values were removed. Some influential observations were not as obvious and if they could not be labeled as storm flow or bad data keeping they would be left in. After sufficient attention was given to the influential observations the analysis was re-run and more influential observations could be found and attention would need be given to these also. This process was completed for all four of the dependent variables, (pH , ANC, Nitrate, and Sulfate)

The step-wise variable selection process requires a list of variables to choose from. The original list used and the variables chosen are reported in Table C.1. The variables chosen for this list were chosen from variables available within the stream survey dataset. The variables selected were used to create the fixed models in Table 2.2. If any of the time variablse were chosen by the step-wise method then the others were added. This was done to ensure the julian date coefficient was present along with $\sin(\theta)$ and $\cos(\theta)$ for seasonality. The problem with these equations is still the high

Table 2.2: Equations created through step-wise variable selection.

Dependent (n)	Model	Adjusted r^2	Model p
pH (3116)	$.673 \times \log_2(\text{Sum Base Cations}) + (-.368 \times NO_3) + (.262 \times \text{Julian Day}) + (-.266 \times SO_4) + (-.050 \times \cos(\theta))$	0.630	<0.001
ANC (3116)	$(.415 \times \text{Sum Base Cations}) + (-.185 \times SO_4) + (.595 \times \text{Conductivity}) + (-.102 \times NO_3) + (.019 \times \text{Julian Date}) + (.005 \times \text{Cl}) + (.005 \times \sin(\theta))$	0.984	0.049
NO_3 (3116)	$(-.295 \times SO_4) + (-3.183 \times \text{ANC}) + (2.19 \times \text{Conductivity}) + (.923 \times \text{Sum Base Cations}) + (.120 \times \text{Julian Date}) + (.051 \times \text{Cl}) + (.047 \times \sin(\theta)) + (.031 \times \cos(\theta))$	0.498	0.017
SO_4 (3116)	$(-.166 \times NO_3) + (2.318 \times \text{Conductivity}) + (-3.229 \times \text{ANC}) + (1.033 \times \text{Sum Base Cations}) + (.042 \times \text{Julian Date})$	0.720	<0.001

amount of variation within them while trying to determine a time trend. All of the equations contain the time variables (julian date, $\sin(\theta)$, and $\cos(\theta)$ along with some other chemical variables. Because of the difficulty of explaining what the julian date coefficient really means along side all of this other variation a second set of equations was created. Theses equations contain only the three time variables.

2.2 Results

In Robinson et al. (2008) julian date coefficients are are reported for each of the eleven historical elevation classes and across each of the dependent variables (pH, ANC, NO_3, and SO_4). Julian date coefficients for this paper were reported in similar tables. There are 144 different julian date coefficients in two tables. Table D.1 records the julian date coefficients calculated using equations with step-wise variables and Table D.2 records the julian date coefficients for equations containing only the time variables.

- Each trend line is represented by its Julian date coefficient, the r^2 value for the trend line, and it's statistical significance.
- 2 of the 72 trend lines in Table D.1 are insignificant. In contrast only 20 of the trend lines in Table D.2 are significant.
- Insignificance is determined by receiving a p-value greater than .05, the α of the trend line. A p-value greater than .05 rejects the hypothesis that $\beta \neq 0$. There is greater than a 5% chance that β=0. There is to much chance of no trend line for the scientific community.
- Repeat trends from previous thesis
- Compare to (Robinson et al., 2008), his trends are negative.

2.3 Discussion

- Why were the trends insignificant?
- Why are the water quality trends trending the way they are at separate time sets (discuss comparisons between sets in the ANOVA bonferoni section).
- How should the water quality variables have behaved based on known properties and (Robinson et al., 2008).

- Very generally speaking these results are different than (Robinson et al., 2008) predicted.
- Water quality will continue to get better.
  - because pollution is being regulated
- there are still unknowns and prediction is still hard.

Chapter 3

Means Comparison

3.1 Methods

3.1.1 Introduction

- In the year 2008 scrubbers were installed into the Bullrun and kingston power plants
- These scrubbers significantly reduced the amount of SO_4 emitted by the smoke stacks of the power plants by **how much**
- A the same time an obvious decrease in measured SO_4 was discovered in the Stream Survey samples (UTK, 2012).
- The amount of SO_4 in the streams is thought to be (correlated with?) to the pH index of the streams when the SO_4 goes up the pH goes down.
- The hypothesis is that of the three sets of data containing water quality measurements from 1993 to 2012, if the data is broken at 2002 and 2008, that because of the obvious measured decrease in SO_4, there will be an obvious difference of means in the sets before and after 2008.
- This can be tested using an Analysis of Variance procedure.

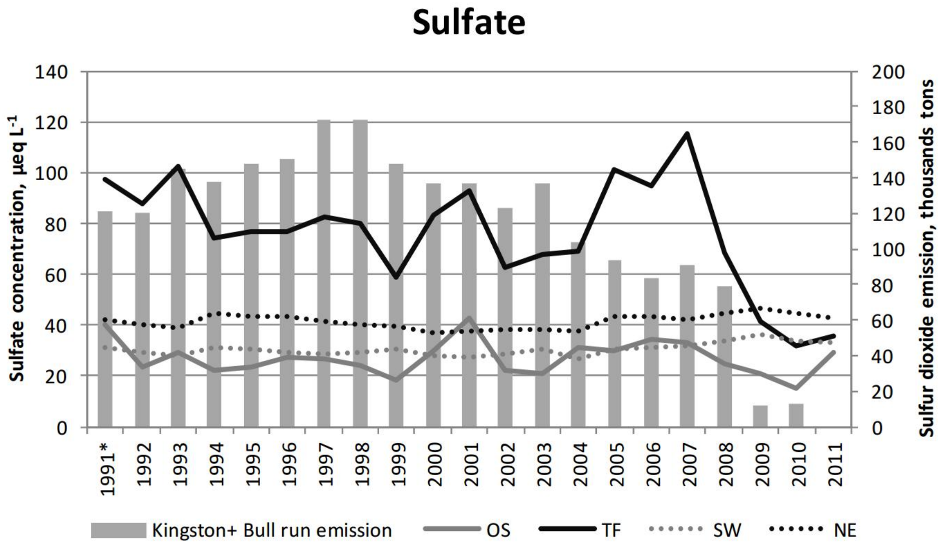


Figure 3.1: Sulfate emmisions of Kingstion and Bull run against those measured in Noland high elevation site.

- The data is only pH measurements for the three sets

Instruments

- The program used for this procedure was (probably SAS).
- Heterscedasticity can be a problem a brown-forsythe test was employed to test for this.
- If three groups are analyzed using ANOVA the only two outcomes may be "they are different" or "they are not different".
- If they are not different then the analysis of the data is over.
- If they are different then it would be nice to know which sets are different.

- This is accomplished with a Bonferoni analysis

3.1.2 Bonferoni Introduction

- Introduction from text book.

- rank-sum

  instruments

  - Bonferoni can output a graph presenting the means of each group in order to visually check for a difference in means. It will also output 95% confidence intervals between each pair of groups. This way definitive answers can be found for the question of "are they or are they not the same?"
  - Bonferoni assumptions
  - SAS

3.2 Results

- Background info?

- The output of the Bonferoni method includes 95% confidence intervals that represent definitive comparisons of the means of two groups of data. If the C.I. includes zero then the means are not statistically different.

- Table 3.1 reports the Bonferoni comparison means between the four water quality variables(pH, ANC, NO_3, SO_4) in one time set against the same water quality variable in another time set by elevation bands.

Table 3.1: Bonferoni comparisons between multiple groups

Elevation Classes	pH			ANC			Nitrate			Sulfate		
	1-2	1-3	2-3	1-2	1-3	2-3	1-2	1-3	2-3	1-2	1-3	2-3
1	$\neq$	$\neq$	$\neq$	$=$	$=$	$=$	$\neq$	$=$	$=$	$=$	$=$	$=$
2	$=$	$=$	$=$	$=$	$\neq$	$=$	$\neq$	$\neq$	$=$	$\neq$	$\neq$	$=$
3	$\neq$	$\neq$	$\neq$	$=$	$\neq$	$=$	$=$	$\neq$	$\neq$	$=$	$=$	$=$
4	$=$	$\neq$	$\neq$	$=$	$=$	$=$	$=$	$=$	$=$	$=$	$=$	$=$
5	$\neq$	$\neq$	$\neq$	$=$	$\neq$	$\neq$	$\neq$	$=$	$\neq$	$=$	$=$	$=$
6	$=$	$\neq$	$\neq$	$=$	$=$	$=$	$=$	$=$	$=$	$=$	$=$	$=$

- There are three groups compared in Table 3.1, they are the three time sets: 93-02, 03-08, 09-12. The table uses $=$ and $\neq$ to represent equality or inequality between the means of the groups compared.
- stuff from first draft
- the bonferoni analysis also outputs the results in figure form. These figures visually represent the group means and are presented in Figure **??** through Figure **??**.
- The negative trend of ANC is something to take note of.

3.3 Discussion

- Are these results a special case?
- Do these differences show up in other water quality analyses, not in the S.S.?
- What are the reasons that the means are higher or lower than expected?
- Differences between sets 2 and 3 were expected due to the scrubbers, did this occur?

- Why aren't the results as clear as the chart in the (Cai and Schwartz, 2012)?
  - probably math
- General hypothosis about what the results suggest
  - The results suggest that a larger difference is needed to see a sulfate difference between sets 2 and 3.

Chapter 4

Power Analysis

4.1 Methods

4.1.1 Introduction

- Statistics come with an inherent amount of error.
- The trend lines created in the trend analysis chapter have a defined error called type II error or β.
- β describes failure to reject a false null hypothesis or failure to detect a trend in the data when there really is one.
- β is usually described in terms of probability and it's opposite is called power(1-β)
- The power of a statistical test describes the probability that the test is true.
- The statistical test is the hypothesis test which tests if the coefficients of a regression line are zero. So whether or not a trend exists.
- The power of the trend lines will state the "truth" of the slope of the trend. A trend lilne with a power of 1.00 means that there is a 100% chance that the calculated slope is not zero.

- Using the earlier calculated trend lives as input, the power of each regression line was calculated with the help of G*power. An a priori analysis was calculated to help determine the number of samples needed for desired levels of power.

4.1.2 Body

- The objectives of the power analysis are to determine the power of the trend lines calculated from past observations and to determine an adaquate number of samples needed for different levels of power for the future.

- The inputs needed in the G*power program for a post hoc analysis are: number of observations (N), adjusted r^2, number of predictors, and Effect size. N and adj.r^2 are outputs from the trend analysis and effect size is calculated using G*power. These values are reported in Table F.1 and Table F.2.

- A post hoc analysis of the trend line data from Table F.2 is not useful. This is because most of the lines have terribly low r^2 values and are insignificant trend lines. The power of an insignificant trend line is also insignificant.

- Post hoc analysis and a priori were run on both methods for trend lines

- G*power is a free power analysis program written by four germen psychology professors.

- It runs the gamut in power analysis options and uses methods stated in (Cohen, 1992).

- G*power was used to calculate powers in the post hoc analysis and sample sizes for the a priori analysis.

- Excel was used along with results provided by G*power to create scenarios to finish up the a priori analysis.

4.1.3 Procedures

Post hoc

- Data compiled in Table F.1 and Table F.2 give the inputs required for a post hoc power analysis on th previously created trend lines.
- required inputs for G*power include, ES(Effect Size), α(alpha), number of observations, and number of predictors.
  - ES is calculated in G*power by the Cohen method stated in (Cohen, 1992) "A Power Primer".
  - Alpha refers to the α of the trend lines (.05).
  - Number of observations is given in trend line output from SPSS.
  - Number of predictors is also stated in trend line output.
- The calculate button will calculate the power
  - This is all that is needed for a post hoc analysis. It answers the question "What was the power of the survey ?" or "How strong are the trend lines that were computed?"
- The calculated powers are reported along side their trend line inputs in Table F.1 and Table F.2.

A priori

- The a priori analysis can help survey planners to create a sampling survey that will produce trend lines with certain ES values and powers.
- There are two objectives to this analysis which are to create "power graphs", which are plots of power vs. sample size. The other is to plan out an actual scenario for which samples can be added or subtracted to elevation bands for a desired power of .80 and an ES of .15.

- The power graphs are created in G*power using the "x-y plot for a range of values" button next to the "calculate" button.

- They-axis has the power values while the x-axis contains the number of observations or samples. The power will increase with number of samples until it reaches 100.

- Four power graphs were created, one for each water quality variable. If ES and power are set to .15 and .80 respectively for every presumed trend line then the only variable is number of predictors. The number of predictors is set for each water quality variable (pH, ANC, NO_3,SO_4). Taken from the earlier step-wise selection method (link to step-wise table). Therefore only one "power graph" is needed for every trend line in each variable.

- ES and power can be chosen or kept constant based on reports by Cohen.

- Cohen's standardizations

- While the "power graphs" are useful in planning for the future of the stream survey, it can be shown that if ES and power are chosen, exact numbers of samples and sites can be added and subtracted from elevation bands.

Table 4.1: A priori calculation in G*power when alpha, ES, and power are set to .05, .15, and .80 respectively.

	Number of predictors	N
pH	6	98
ANC	8	109
Nitrate	8	109
Sulfate	7	103
Time	3	77

- The rest was done in excel

- This calculated number of observations can be divided by the number of samples collected in one year to get the number of years required to reach a power of .80.

- The analysis can be further conducted by calculating the number of samples per year to achieve a power of .80. For this calculation all water quality variables were given the highest number of samples of 110 and 77 was used for the trends using only time variables.

Table 4.2: samples/year to achieve a power .80

Years	1	2	3	4
Water Quality Variables	110	55	37	28
Time Variables	77	39	26	19

- Table 4.2 is needed to calculate number of samples needed per elevation band to achieve a power of .80. This number of samples can then be further divided to get a number of sites needed to achieve a power of .80. If a trend line with a power of .80 is desired after one year ,for all water quality variables to be satisfied, 110 samples need to be collected. If four years are waited then only 28 samples need to be collected per year.

- To create this final table the number of samples per elevation band was subtracted from the number of samples to achieve a power of .80 which gives us the number of samples needed in addition to what is currently collected to receive a power of .80. These results are organized into samples needed per elevation band to achieve a power of .80 and seperated by years depending of how many years of data go into the trend lines.

4.2 Results

4.2.1 Post hoc

- A post hoc power analysis was conducted for each of the two methods of trend analysis.
- Table F.1 and Table F.2 record the results of the post hoc analysis on the trend lines with variables created through the step-wise method and the trend lines created using only time variables respectively. Included in these tables are the number of samples and r^2 variables from the trend analysis and effect size and power from the post hoc analysis.
- Table F.1 and Table F.2 are broken into the four analyzed water quality variables (pH, ANC, NO_3,SO_4) and divided into the tree time sets (93-02, 03-08, 09-12), and then further divided into the six elevation classes.
- use results from previous draft
- any similar power analysis?

4.2.2 A priori

Power graphs

- The results of the a priori power analysis will be the most important for planning.
- The usual output is the "power graph" which plots power on the y-axis and total sample size on the x-axis.
- G*power outputs some very nice power graphs. The power graphs created from the a priori power analysis are presented in Figure G.1, Figure G.2, Figure G.3, and Figure G.4.

- There were four power graphs created, three for the water quality variables and one for the time variables. ANC and Nitrate both have the same number of predictors from the step-wise variable selection method and therefore create the same power graph.

- each graph contains 3 lines representing 3 different ES choices: .15, .25, and .35. These were chosen to mimic the choices of small, medium, and large effects standardized by Cohen in (Cohen, 1992). Limitations of the G*power program left the best choices to be .15, .25, and .35. A small effect of .02 was ignored because preliminary graph results showed it to be not useful.

Planning with power analysis

- Using the ability of the a priori power analysis to compute a number of samples needed for a certain power, a scenario was played out to see how many sites needed to be added or could be removed from an elevation band in the stream survey.

Table 4.3: Years to acheive a power of .80

Elevation Bands	Site #	Current n/yr	pH	ANC NO_3	SO_4	Time variables
1	13 ,23, 24, 30, 479	26	3.77	4.19	3.96	2.96
2	4, 311, 268, 480, 310, 483, 147, 148, 484	34	2.88	3.21	3.03	2.26
3	114, 481, 482, 149, 66, 492, 137, 293, 270, 493, 485, 144, 224	62	1.58	1.76	1.66	1.24
4	143, 142, 73, 71	24	4.08	4.54	4.29	3.21
5	74, 221, 251, 233	22	4.45	4.95	4.68	3.50
6	253, 234	12	8.17	9.08	8.58	6.42

- This scenario was followed through with both methods of trend lines.

- Table 4.3 records the six elevation bands along with the site numbers that belong to them. In the column labeled ,current n per year, the amount of samples collected per elevation band in the year 2012 was tabulated. The values in the remaining columns were calculated by dividing the number of samples given in Table 4.1 by the current samples per year column in Table 4.3.

- Looking at the table there are 26 samples collected in elevation band one in one year. In order to compute a trend line that receives a power of .80 with pH as the dependent samples would need to be collected for 3.77 years before the trend line is computed. The larges is elevation class for a trend line in ANC or NO_3 which requires 9.08 years.

Table 4.4: Necesary sites scenario for water quality variables

	#Samples required				# sites required			
Elevation Bands	1 yr	2 yrs	3 yrs	4 yrs	1 yr	2 yrs	3 yrs	4 yrs
1	84	29	11	2	14	5	2	0
2	76	21	3	-7	13	4	0	-1
3	48	-7	-25	-35	8	-1	-4	-6
4	86	31	13	4	14	5	2	1
5	88	33	15	6	15	6	2	1
6	98	43	25	16	16	7	4	3

Table 4.5: Necesary sites scenario for time variables

	#Samples required				# sites required			
Elevation Bands	1 yr	2 yrs	3 yrs	4 yrs	1 yr	2 yrs	3 yrs	4 yrs
1	51	13	0	-7	9	2	0	-1
2	43	5	-8	-15	7	1	-1	-2
3	15	-24	-36	-43	3	-4	-6	-7
4	53	15	2	-5	9	2	0	-1
5	55	17	4	-3	9	3	1	0
6	65	27	14	7	11	4	2	1

- The left side of both **??** and **??** show how many more samples are required to get a trend line with a power of .80.

- In **??** for elevation class 3, 48 more samples need to be collected if a trend line with a power of .80 is to be created after one year. But if a trend line can wait to be created after two years, then there is a surplus of seven samples per year. If four years can be waited there is a surplus of 35 samples which on the right side of the table translates into a surplus of 6 whole site locations per year.

- **??** works the same way as **??** but of course it uses different variables for the trend lines.
- results from previous draft
- any other papers like this?

4.3 Discussion

4.3.1 Post hoc

- The results presented in Table F.1 and Table F.2 show how the calculated power is highly affected by number of observations more than anything else.
- In Table F.1, even when the r^2 and ES values are relatively low if the N is greater than 100 then the power is excellent.
- Table F.2 show the effect of the ES on power. Other than these lines being insignificant, many of the ES values are small according to Cohen and when compared to Table F.1. Low ES values and low observations create low powers. Low ES values com from low r^2 values. The low r^2 values can be blamed for the insignificance of the lines and the poor powers.
- Some lines are just not well described by Julian Date,$\sin(\theta)$, and $\cos(\theta)$ only.

4.3.2 A priori

- How can these results be used?
- How can these results be manipulated?

- The results in Table 4.4 and Table 4.5 can help with both of the problems of The park wanting a cheaper survey and researchers wanting more high elevation sites.
- The table can be used to re-organize sites across bands.
  - In the current SS scheme there is a surplus of sites in lower elevation bands and a deficit for sites in higher elevations.
  - Looking at the right side of Table 4.4, if trends are desired after four years of data with a power of .80 and an ES of .15, seven sites may be taken from elevation bands 2 and 3 and 5 would need to be added to elevation bands 4,5, and 6.
  - After this re-arrangement two sites may be completely discontinued.
  - This saves time, effort, and money, but it is a very specific scenario.
- The downside of an a priori power analysis is that once you pick all the variables that go into it, you can' t change them in the future
  - Variables that can change include how you divide the sites into elevation bands
  - Trend line creation (alpha, variable selection)
  - Power analysis (power, and ES)
- If during the hypothetical situation in which four years are waited to do another trend analysis, a better model is found, then the survey would need to be re-evaluated to reflect the new model.
  - the model could require a different number of sites
- Choices for power and ES could change
- planning with the a priori power analysis requires guessing the trends for the future.

- This guess will probably be based on the past , such as this one.
- This guess assumes that trends of the past will continue into the future
- The ANOVA/Bonferoni and the comparison between (Robinson et al., 2008) and the current trends shows that this is difficult.
- better understanding is needed
- At the end of the day the trends are positive!

Bibliography

Bibliography

Cai, M. C. and Schwartz, J. S. (2012). Biological effects of stream water quality on aquatic macroinvertebrates and fish communities within great smoky mountains national park. 4, 5, 8, 16

Cohen, J. (1992). A power primer. *Psychological bulletin*, 112(1):155. 18, 19, 23

Robinson, R. B., Barnett, T. W., Harwell, G. R., Moore, S. E., Kulp, M., and Schwartz, J. S. (2008). ph and acid anion time trends in different elevation ranges in the great smoky mountains national park. *Journal of Environmental Engineering*, 134(9):800–808. 4, 5, 10, 11, 27

UTK (2012). 2011 water quality annual report. Technical report. 12

Appendix

Appendix A

A.1 Site Data

A.2 Site data

	Site ID	Site Description	Watershed
1	173	Mill Creek above Abrams Creek	Abrams
2	174	Abrams Creek below Cades Cove	Abrams
3	488	Mill Creek at Pumphouse on Forge Creek Road	Abrams
4	489	Abrams Creek 300 m below trailhead bridge	Abrams
5	142	Beech Creek above Lost Bottom Creek	Cataloochee
6	143	Lost Bottom Creek (Cataloochee Creek)	Cataloochee
7	144	Palmer Creek above Pretty Hollow Creek	Cataloochee
8	147	Lower Cataloochee Creek	Cataloochee
9	148	Lower Little Cataloochee Creek	Cataloochee
10	149	Middle Cataloochee Creek at bridge	Cataloochee
11	293	Rough Fork at Caldwell House	Cataloochee
12	493	Palmer Creek at Davidson Branch Trail	Cataloochee
13	4	Lower Rock Creek	Cosby
14	114	Cosby Creek at log bridge	Cosby
15	137	Upper Rock Creek (Cosby Creek)	Cosby
16	492	Camel Hump Creek off Low Gap Trail	Cosby
17	221	Hazel Creek above cascades	Hazel
18	224	Hazel Creek just below Proctor Creek Confluence	Hazel
19	310	Bone Valley Creek (Hazel Creek)	Hazel
20	311	Hazel Creek below Haw Gap Creek	Hazel
21	479	Hazel Creek at Campsite 86	Hazel
22	480	Haw Gap Creek at bridge near Campsite 84	Hazel
23	481	Little Fork above Sugar Fork Trail	Hazel
24	482	Sugar Fork above Little Fork	Hazel
25	483	Sugar Fork above Haw Gap Creek	Hazel
26	484	Hazel Creek at Cold Spring Gap Trail	Hazel
27	485	Walker Creek above Hazel Creek Trail	Hazel
28	13	Little River at boundary	Little
29	23	Lower Middle Prong Little River	Little
30	24	Lower West Prong Little River	Little
31	30	West Prong Little Pigeon at Headquarters	Little
32	66	West Prong Little Pigeon at Chimneys Picnic Area	Little
33	71	Road Prong above barrier cascade	Little
34	73	Walker Camp Prong above Road Prong	Little
35	74	Walker Camp Prong above Alum Cave Creek	Little
36	233	Walker Camp Prong above Alum Cave	Little
37	234	Upper Road Prong	Little
38	237	Walker Camp Prong at last bridge	Little
39	251	Beech Flats above US 441 loop	Oconaluftee
40	252	Beech Flats below roadcut	Oconaluftee
41	253	Beech Flats above roadcut	Oconaluftee
42	268	Oconaluftee River below Smokemont	Oconaluftee
43	270	Beech Flats at Kephart Footbridge	Oconaluftee

Table A.1: GRSM Stream Survey site descriptions

	Site ID	Elevation (ft)	Elevation (m)	slope	Latitude	Longitude	Historical Elevation Classes	New elevation classes
1	173	1715	522.73	35.68	35.59104	-83.85361	3	3
2	174	1715	522.73	10.27	35.59186	-83.85308	3	3
3	488	1790	545.59	4.04	35.58349	-83.83446	4	1
4	489	1710	521.21	32.78	35.59145	-83.85397	4	1
5	142	3300	1005.84	32.42	35.63565	-83.14537	5	2
6	143	3280	999.74	35.69	35.63625	-83.14481	6	2
7	144	2990	911.35	35.66	35.63900	-83.13078	5	2
8	147	2460	749.81	16.84	35.66688	-83.07277	4	3
9	148	2475	754.38	7.58	35.66913	-83.07283	4	3
10	149	2550	777.24	4.45	35.64627	-83.07554	5	3
11	293	2755	839.72	18.73	35.62442	-83.11391	5	4
12	493	2840	865.63	33.10	35.63462	-83.11943	6	6
13	4	2080	633.98	6.11	35.76133	-83.21044	3	1
14	114	2510	765.05	13.71	35.74863	-83.20066	5	2
15	137	2750	838.20	22.92	35.74616	-83.21630	5	2
16	492	2730	832.10	25.86	35.74457	-83.19876	5	6
17	221	4000	1219.20	30.02	35.54632	-83.58283	8	3
18	224	2999	914.00	17.92	35.53212	-83.62234	6	3
19	310	2240	682.75	19.63	35.49994	-83.68014	4	4
20	311	2155	656.84	26.20	35.49377	-83.68852	4	5
21	479	1740	530.35	39.70	35.47233	-83.71933	3	5
22	480	2201	671.00	10.07	35.49474	-83.68873	4	5
23	481	2540	774.19	30.90	35.50256	-83.70835	5	5
24	482	2540	774.19	38.66	35.50236	-83.70859	5	6
25	483	2320	707.14	34.29	35.49947	-83.69494	4	6
26	484	2475	754.38	9.11	35.50331	-83.65930	5	1
27	485	2860	871.73	5.17	35.52249	-83.63101	6	1
28	13	1100	335.28	44.21	35.66763	-83.71450	2	1
29	23	1150	350.52	5.96	35.65724	-83.70979	2	1
30	24	1150	350.52	31.60	35.65682	-83.71017	2	1
31	30	1430	435.86	2.17	35.68819	-83.53672	2	1
32	66	2680	816.86	17.92	35.63723	-83.49484	5	2
33	71	3400	1036.32	31.28	35.63440	-83.47032	6	2
34	73	3360	1024.13	28.98	35.63476	-83.46931	6	2
35	74	3820	1164.34	18.07	35.62912	-83.45102	7	2
36	233	4255	1296.92	21.86	35.61830	-83.42718	8	3
37	234	5000	1524.00	23.93	35.60975	-83.45043	10	3
38	237	4520	1377.70	30.21	35.62409	-83.41692	9	3
39	251	4010	1222.25	19.03	35.60226	-83.41533	8	3
40	252	4680	1426.46	33.32	35.60666	-83.43391	9	3
41	253	4760	1450.85	26.42	35.60682	-83.43510	9	3
42	268	2169	661.00	3.31	35.55293	-83.30937	4	4
43	270	2799	853.00	22.92	35.58641	-83.36400	5	4

Table A.2: Site Data

Appendix B

Descriptive Statistics

Table B.1: Descriptive statistics of Water Quality in the GRSM

Set	Class	pH				ANC meql				Nitrate meql				Sulfate meql			
		N	Minimum	Maximum	Mean	N	Minimum	Maximum	Mean	N	Minimum	Maximum	Mean	N	Minimum	Maximum	Mean
1993-2002																	
	1	327	4.96	7.90	6.57	327	-20.74	1534.47	149.76	275	0.00	49.94	12.04	325	12.32	85.01	36.09
	2	393	5.32	7.00	6.25	392	-7.43	182.95	40.75	377	1.37	73.76	26.62	390	0.00	159.51	51.68
	3	400	4.65	8.24	6.44	398	-19.97	1624.49	158.44	365	0.00	96.13	26.14	391	0.00	262.37	54.00
	4	121	6.18	7.11	6.50	120	24.45	178.00	75.84	105	2.16	28.29	11.90	119	12.34	77.74	25.16
	5	116	6.07	7.05	6.50	116	41.34	162.76	77.06	66	1.23	10.55	4.35	116	7.51	79.98	26.14
	6	110	5.77	7.06	6.41	110	15.64	165.02	68.01	81	1.56	60.46	21.13	110	14.71	61.16	28.35
2003-2008																	
	1	255	5.22	7.95	6.65	255	-37.09	1314.56	173.48	252	0.50	62.75	16.56	261	10.00	93.23	38.85
	2	289	4.83	7.07	6.32	289	-1.88	145.95	42.20	296	0.62	67.12	29.20	298	11.64	152.55	48.19
	3	299	4.65	8.10	6.55	299	-26.45	1591.06	172.82	297	0.13	95.72	27.69	308	10.44	490.01	54.25
	4	119	5.95	7.06	6.58	119	23.36	128.28	69.90	121	1.87	55.67	17.51	123	13.88	61.31	29.04
	5	35	5.98	7.03	6.50	35	36.37	115.80	77.84	30	1.45	26.48	7.59	37	12.18	117.46	30.54
	6	97	5.79	7.05	6.44	97	6.73	130.63	55.68	98	1.09	72.79	24.88	101	10.02	65.53	34.31
209-2012																	
	1	191	5.42	8.02	6.77	191	-0.02	1377.93	164.72	191	0.22	62.14	16.31	190	14.61	113.83	39.63
	2	212	4.91	7.28	6.47	212	-11.74	174.52	44.45	212	4.43	72.17	30.08	212	13.45	125.36	47.41
	3	228	4.73	7.96	6.68	228	-18.28	1535.69	160.14	228	1.04	72.16	26.23	228	13.59	317.63	58.15
	4	97	6.20	7.08	6.68	97	25.70	107.58	64.13	97	0.54	34.67	18.72	97	19.89	46.66	29.33
	5	29	6.30	7.11	6.77	29	40.10	115.94	73.55	29	0.21	83.68	6.44	29	16.78	109.18	36.16
	6	76	4.24	7.09	6.52	76	-3.92	114.28	46.15	76	0.16	79.04	32.17	76	15.72	63.32	37.05

Appendix C

Variable selection

Table C.1: List of variables used for step-wise variable selection. X's for variables selected by the step-wise method, O's if variable was added after the step-wise process.

		Independents for step-wise regression			
Available Variables	comments	pH	ANC	NO_3	SO_4
pH	Dependent				
ANC	Dependent			X	X
NO_3	Dependent	X	X		X
SO_4	Dependent	X	X	X	
Site ID					
Julian Date		X	X	X	X
Month					
Year					
Julian Date Days	Seasonality				
Fraction of the year	Seasonality				
θ	Seasonality				
$\sin(\theta)$	Seasonality	X	X	X	O
$\cos(\theta)$	Seasonality	X	O	X	O
Sum Base Cations			X	X	X
Conductivity			X	X	X
Chloride			X	X	
Sodium					
Potassium					
Magnesium					
Calcium					
Elevation (m)					
Slope					
Elevation class	1-6				
Time set	1-3				
$\log_2$ (ANC)					
$\log_2$ (Base Cations)		X			
Number of predictors		6	8	8	7

Appendix D

Julian Date Coefficients

D.1 Step-wise Method

D.2 Temporal Variables

Table D.1: Time trend results for specific elevation classes using variables from step-wise regression. **Bold** results are insignificant.

Time set	Elevation class	Elevation range m (ft)	Number of sites	Julian date coefficient, eq/L or pH units (model adjusted r^2) (p-value)			
				pH	ANC	NO_3	SO_4
1993-2002	1	304.8-609.6	5	0.069	0.007	0.034	-0.096
		(1000-2000)		0.712	0.985	0.503	0.569
				0.000	0.000	0.000	0.000
	2	609.6-762	9	-0.091	-0.036	-0.037	0.019
		(2000-2500)		0.388	0.603	0.699	0.766
				0.000	0.000	0.000	0.000
	3	762-914.4	13	-0.010	0.008	-0.013	0.024
		(2500-3000)		0.693	0.971	0.359	0.590
				0.000	0.000	0.000	0.000
	4	914.4-1066.8	4	0.019	0.015	0.058	0.061
		(3500-3500)		0.205	0.709	0.410	0.402
				0.000	0.000	0.000	0.000
	5	1066.8-1371.6	4	-0.157	-0.082	0.288	-0.133
		(3500-4500)		0.165	0.760	0.328	0.566
				0.010	0.000	0.000	0.000
	6	1371.6<	2	0.218	0.067	-0.011	0.092
		(4500<)		0.505	0.802	0.871	0.716
				0.000	0.000	0.000	0.000
2003-2008	1	304.8-609.6	5	0.150	-0.004	0.038	0.039
		(1000-2000)		0.781	0.996	0.551	0.673
				0.000	0.000	0.000	0.000
	2	609.6-762	9	0.275	0.033	0.044	0.044
		(2000-2500)		0.348	0.779	0.816	0.893
				0.000	0.000	0.000	0.000
	3	762-914.4	13	0.156	0.005	0.072	0.034
		(2500-3000)		0.663	0.996	0.637	0.923
				0.000	0.000	0.000	0.000
	4	914.4-1066.8	4	0.249	-0.028	0.092	0.110
		(3500-3500)		0.400	0.779	0.405	0.343
				0.000	0.000	0.000	0.000
	5	1066.8-1371.6	4	0.137	-0.020	0.204	0.135
		(3500-4500)		0.300	0.739	0.562	0.884
				0.027	0.000	0.001	0.000
	6	1371.6<	2	0.359	0.127	0.074	0.161
		(4500<)		0.317	0.812	0.832	0.844
				0.000	0.000	0.000	0.000
2009-2012	1	304.8-609.6	5	0.106	-0.002	0.026	-0.052
		(1000-2000)		0.894	0.989	0.376	0.536
				0.000	0.000	0.000	0.000
	2	609.6-762	9	0.218	0.069	0.121	0.039
		(2000-2500)		0.606	0.862	0.735	0.887
				0.000	0.000	0.000	0.000
	3	762-914.4	13	0.056	0.007	0.019	0.050
		(2500-3000)		0.766	0.997	0.598	0.915
				0.000	0.000	0.000	0.000
	4	914.4-1066.8	4	0.413	-0.006	-0.013	-0.068
		(3500-3500)		0.593	0.772	0.635	0.529
				0.000	0.000	0.000	0.000
	5	1066.8-1371.6	4	**-0.115**	0.901	**0.098**	0.015
		(3500-4500)		**0.158**	0.540	**-0.272**	0.658
				0.130	0.001	**0.975**	0.000
	6	1371.6<	2	0.289	0.059	0.097	-0.059
		(4500<)		0.286	0.809	0.881	0.861
				0.000	0.000	0.000	0.000

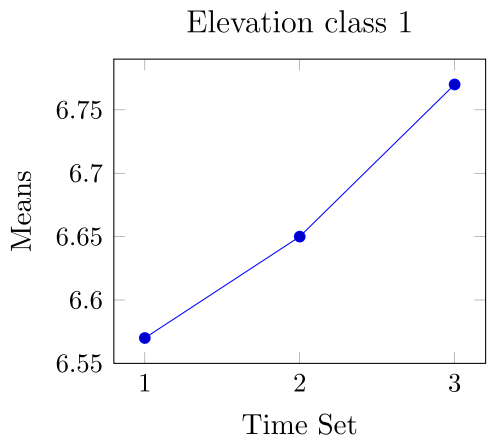
Table D.2: Time trend results for specific elevation classes using julian date, cosine(θ), and sine(θ) only. **Bold** results are insignificant.

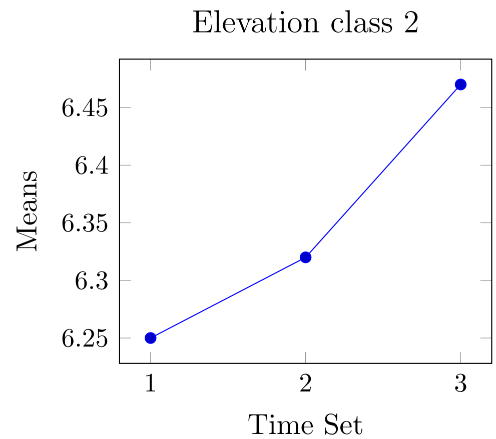
Time range	Elevation class	Elevation range m (ft)	Number of sites	Julian date coefficient, eq/L or pH units (model adjusted r^2) (p-value)			
				pH	ANC	Nitrate	Sulfate
1993-2002	1	304.8-609.6	5	**0.054**	**0.089**	-0.138	-0.190
		(1000-2000)		**0.047**	**0.024**	0.016	0.045
				0.321	**0.106**	0.022	0.001
	2	609.6-762	9	**-0.090**	**-0.060**	**-0.060**	**-0.075**
		(2000-2500)		**0.128**	**0.189**	**0.017**	**0.009**
				0.060	**0.195**	**0.248**	**0.142**
	3	762-914.4	13	**-0.012**	**-0.030**	**-0.048**	**-0.047**
		(2500-3000)		**0.013**	**0.000**	**-0.004**	**-0.004**
				0.817	**0.550**	**0.365**	**0.355**
	4	914.4-1066.8	4	**-0.047**	**-0.151**	**-0.009**	**0.095**
		(3500-3500)		**0.059**	**0.294**	**-0.027**	**-0.016**
				.597	**0.055**	**0.926**	**0.313**
	5	1066.8-1371.6	4	**-0.151**	-0.148	0.330	**0.092**
		(3500-4500)		**0.051**	0.381	0.120	**-0.010**
				.100	0.047	0.006	**0.331**
	6	1371.6<	2	**.156**	**-0.016**	**-0.208**	**-0.036**
		(4500<)		**.096**	**0.075**	**0.092**	**-0.009**
				.092	**0.863**	**0.058**	**0.707**
2003-2008	1	304.8-609.6)	5	.139	**0.009**	0.155	0.192
		(1000-2000		0.040	**0.001**	0.061	0.043
				0.025	**0.888**	0.012	0.002
	2	609.6-762	9	0.145	**-0.090**	0.178	0.138
		(2000-2500)		0.061	**0.081**	0.043	0.014
				0.012	**0.114**	0.002	0.017
	3	762-914.4	13	**0.103**	**-0.006**	**0.047**	**0.099**
		(2500-3000)		**0.020**	**-0.003**	**-0.003**	**0.006**
				0.075	**0.925**	**0.418**	**0.085**
	4	914.4-1066.8	4	0.235	**-0.029**	0.193	0.192
		(3500-3500)		0.148	**0.180**	0.086	0.023
				0.007	**0.728**	0.030	0.035
	5	1066.8-1371.6	4	**0.135**	**-0.112**	**-0.176**	**0.067**
		(3500-4500)		**-0.069**	**0.337**	**-0.082**	**-0.024**
				0.466	**0.443**	**0.401**	**0.701**
	6	1371.6<	2	0.204	**-0.108**	0.236	0.307
		(4500<)		0.081	**0.094**	0.046	0.074
				0.041	**0.274**	0.020	0.002
2009-2012	1	304.8-609.6	5	**0.111**	**0.026**	**-0.036**	**-0.092**
		(1000-2000)		**0.028**	**0.000**	**0.018**	**0.005**
				0.122	**0.718**	**0.619**	**0.207**
	2	609.6-762	9	0.141	**0.017**	**0.020**	**-0.062**
		(2000-2500)		0.052	**0.056**	**0.011**	**-0.010**
				0.037	**0.800**	**0.767**	**0.376**
	3	762-914.4	13	**-0.034**	**-0.027**	**-0.036**	**0.078**
		(2500-3000)		**-0.009**	**-0.002**	**-0.004**	**-0.007**
				0.611	**0.684**	**0.592**	**0.246**
	4	914.4-1066.8	4	0.405	**0.032**	**-0.067**	**-0.129**
		(3500-3500)		0.200	**0.161**	**-0.016**	**-0.011**
				0.000	**0.733**	**0.518**	**0.215**
	5	1066.8-1371.6	4	**-0.031**	0.891	**0.052**	**-0.414**
		(3500-4500)		**0.218**	0.466	**-0.039**	**-0.076**
				0.934	0.007	**0.904**	**0.347**
	6	1371.6<	2	0.264	**0.083**	**-0.021**	**-0.214**
		(4500<)		0.039	**0.058**	**-0.016**	**0.007**
				0.023	**0.462**	**0.859**	**0.068**

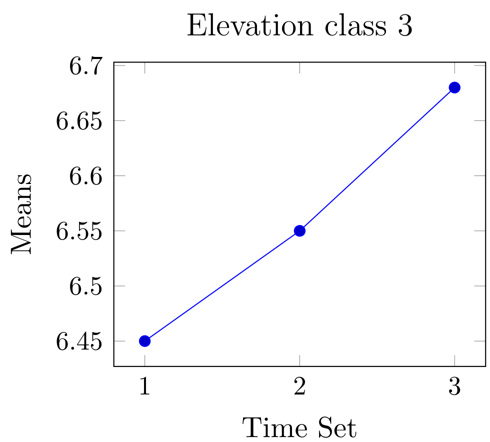
Appendix E

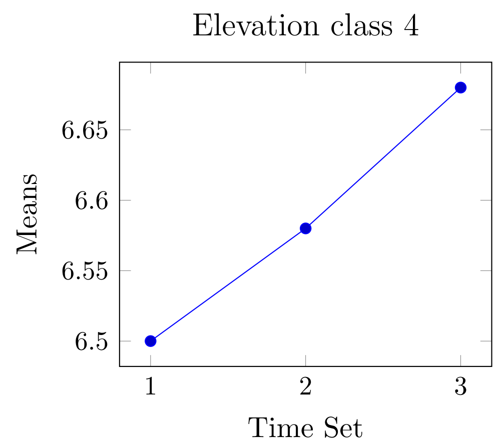
ANOVA/Bonferoni

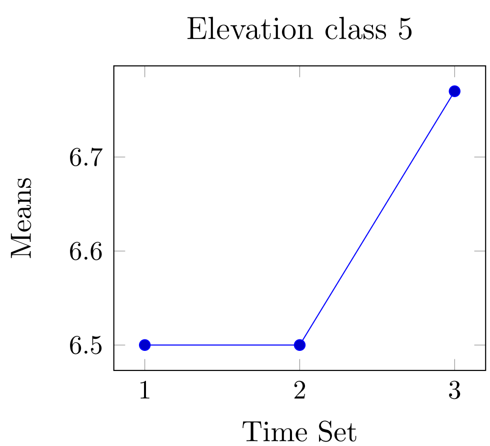
E.1 pH

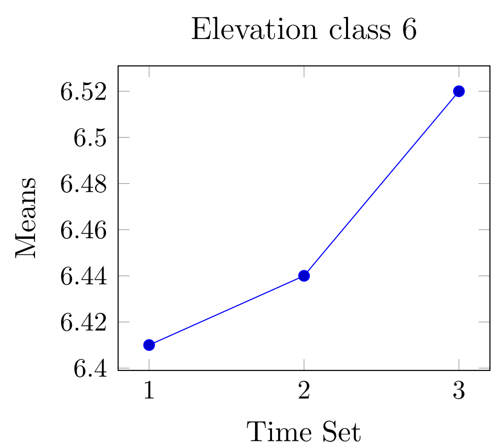

Elevation class 1
6.75
6.7
6.65
6.6
6.55
Means
1
2
3
Time Set


Elevation class 2
6.45
6.4
6.35
6.3
6.25
Means
1
2
3
Time Set

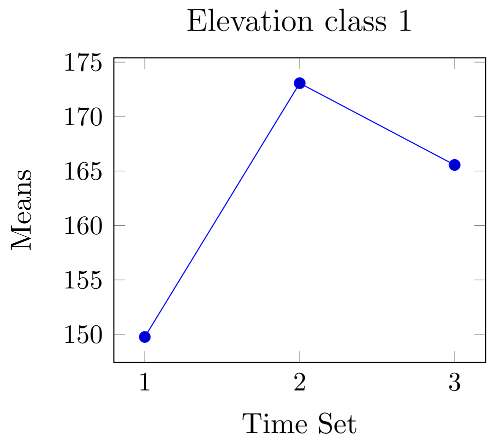

Elevation class 3
6.7
6.65
6.6
6.55
6.5
6.45
Means
1
2
3
Time Set

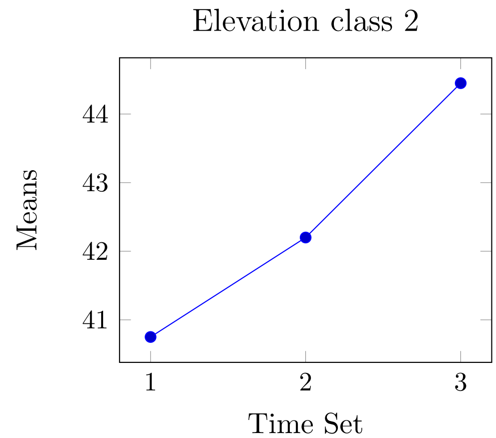

Elevation class 4
6.65
6.6
6.55
6.5
Means
1
2
3
Time Set

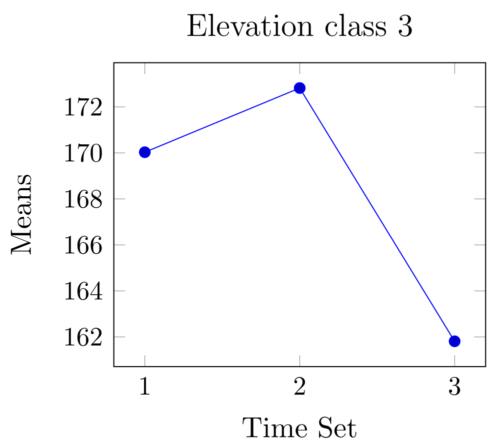

Elevation class 5
6.7
6.6
6.5
Means
1
2
3
Time Set

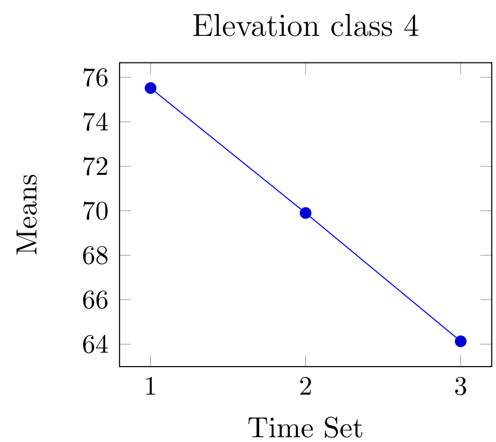

Elevation class 6
6.52
6.5
6.48
6.46
6.44
6.42
6.4
Means
1
2
3
Time Set

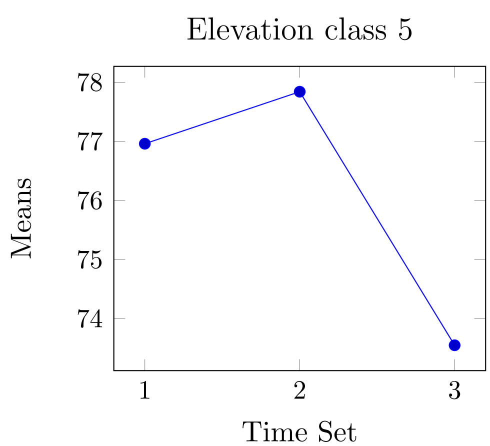
E.2 ANC

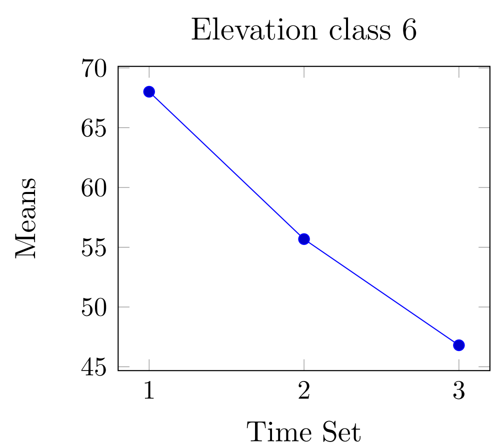

Elevation class 1
Means
175
170
165
160
155
150
1
2
3
Time Set


Elevation class 2
Means
44
43
42
41
1
2
3
Time Set

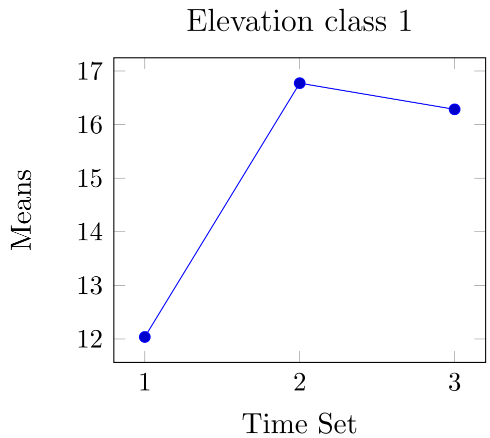

Elevation class 3
Means
172
170
168
166
164
162
1
2
3
Time Set

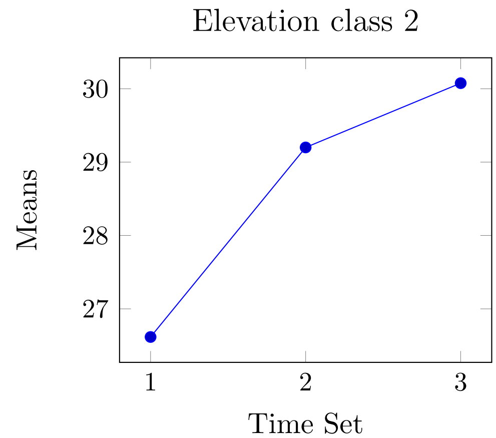

Elevation class 4
Means
76
74
72
70
68
66
64
1
2
3
Time Set

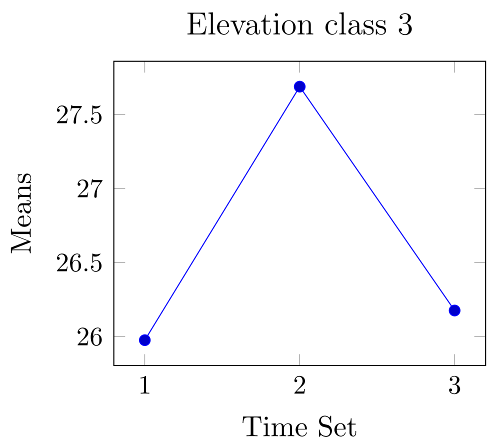

Elevation class 5
Means
78
77
76
75
74
1
2
3
Time Set

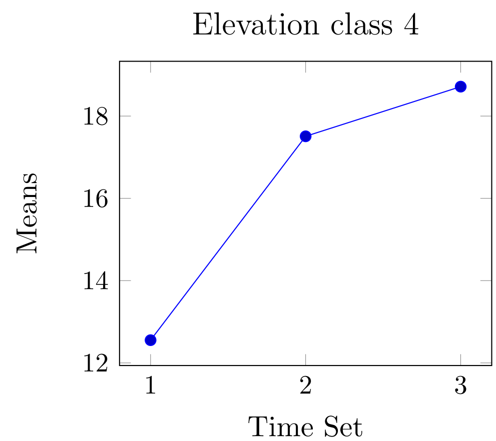

Elevation class 6
Means
70
65
60
55
50
45
1
2
3
Time Set

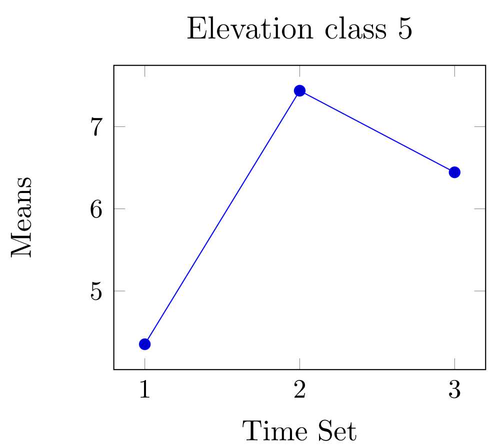
E.3 Nitrate

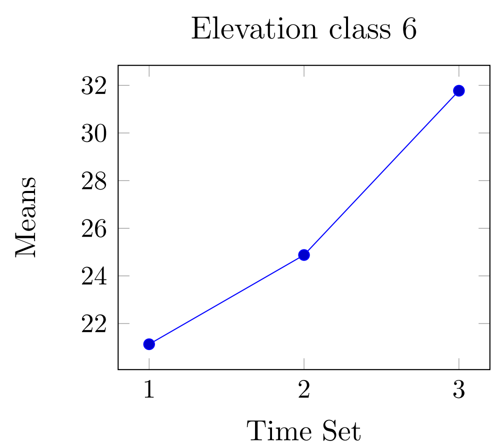

Elevation class 1
Means
17
16
15
14
13
12
1
2
3
Time Set


Elevation class 2
Means
30
29
28
27
1
2
3
Time Set

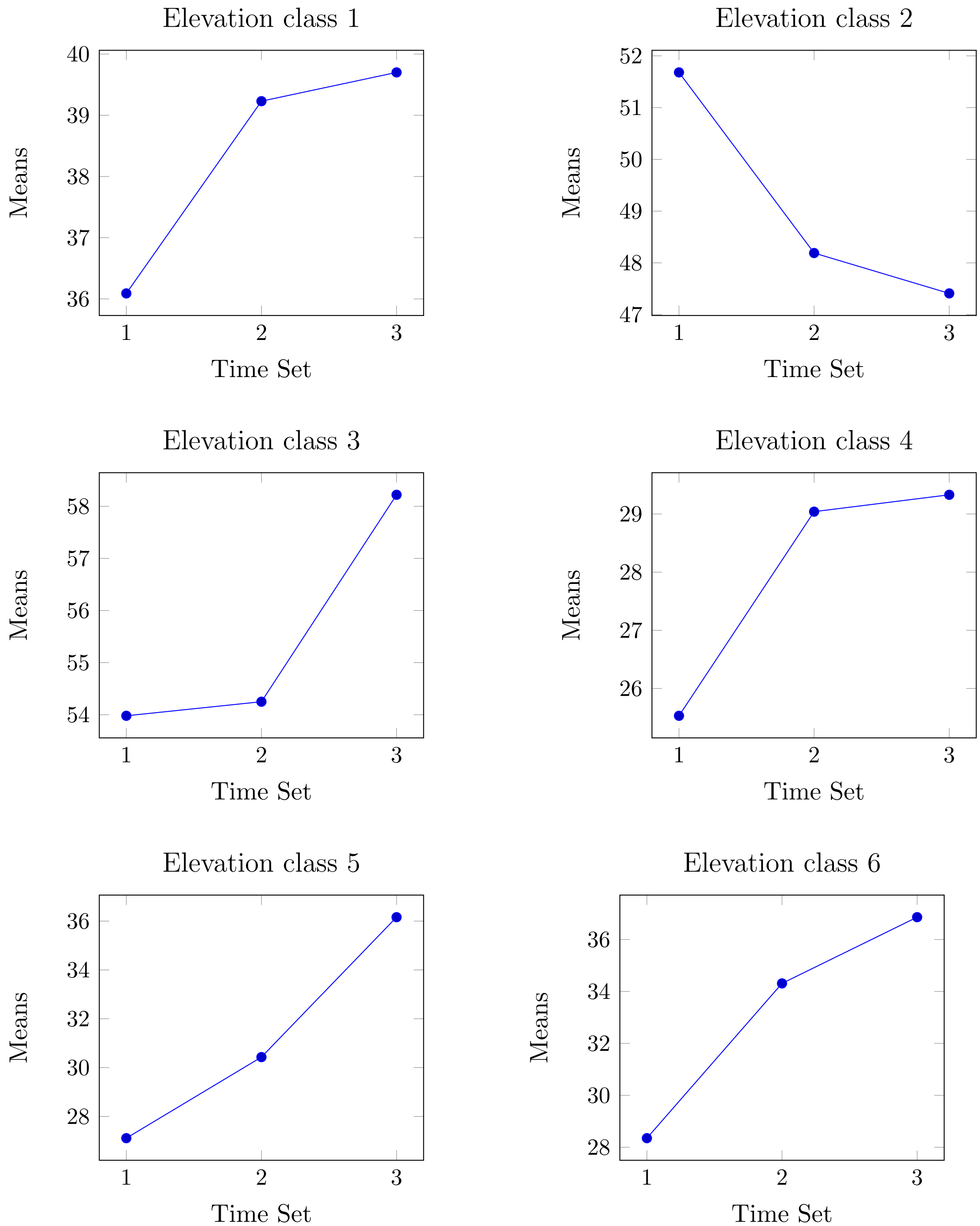

Elevation class 3
Means
27.5
27
26.5
26
1
2
3
Time Set


Elevation class 4
Means
18
16
14
12
1
2
3
Time Set


Elevation class 5
Means
7
6
5
1
2
3
Time Set


Elevation class 6
Means
32
30
28
26
24
22
1
2
3
Time Set

E.4 Sulfate



Elevation class 1
Elevation class 2
Elevation class 3
Elevation class 4
Elevation class 5
Elevation class 6
Means
Time Set

Appendix F

Post Hoc Power Analsysis

F.1 Step-Wise Variables

F.2 Temperol variables

Table F.1: Post hoc power analysis using G*power and a calculated ES, alpha is .05. **Bold** results are insignificant.

		pH				ANCmeqL				NitratemeqL				SulfatemeqL			
Set	Class	N	Adjusted r^2	Effect Size	Actual Power	N	Adjusted r^2	Effect Size	Actual Power	N	Adjusted r^2	Effect Size	Actual Power	N	Adjusted r^2	Effect Size	Actual Power
1993-2002																	
	1	327	0.712	2.47	1.00	327	0.985	65.67	1.00	275	0.503	1.01	1.00	325	0.569	1.32	1.00
	2	393	0.388	0.63	1.00	392	0.603	1.52	1.00	377	0.699	2.32	1.00	390	0.766	3.27	1.00
	3	400	0.693	2.26	1.00	398	0.971	33.48	1.00	365	0.359	0.56	1.00	391	0.590	1.44	1.00
	4	121	0.205	0.26	0.99	120	0.709	2.44	1.00	105	0.410	0.69	1.00	119	0.402	0.67	1.00
	5	116	0.165	0.20	0.96	116	0.760	3.17	1.00	66	0.328	0.49	0.98	116	0.566	1.30	1.00
	6	110	0.505	1.02	1.00	110	0.802	4.05	1.00	81	0.871	6.75	1.00	110	0.716	2.52	1.00
2003-2008																	
	1	255	0.781	3.57	1.00	255	0.996	249.00	1.00	252	0.551	1.23	1.00	261	0.673	2.06	1.00
	2	289	0.348	0.53	1.00	289	0.779	3.52	1.00	296	0.816	4.43	1.00	298	0.893	8.35	1.00
	3	299	0.663	1.97	1.00	299	0.996	249.00	1.00	297	0.637	1.75	1.00	308	0.923	11.99	1.00
	4	119	0.400	0.67	1.00	119	0.779	3.52	1.00	121	0.405	0.68	1.00	123	0.343	0.52	1.00
	5	35	0.300	0.43	0.74	35	0.739	2.83	1.00	30	0.562	1.28	0.98	37	0.884	7.62	1.00
	6	97	0.317	0.46	1.00	97	0.812	4.32	1.00	98	0.832	4.95	1.00	101	0.844	5.41	1.00
2009-2012																	
	1	191	0.894	8.43	1.00	191	0.989	89.91	1.00	191	0.376	0.60	1.00	190	0.536	1.16	1.00
	2	212	0.606	1.54	1.00	212	0.862	6.25	1.00	212	0.735	2.77	1.00	212	0.887	7.85	1.00
	3	228	0.766	3.27	1.00	228	0.997	332.33	1.00	228	0.598	1.49	1.00	228	0.915	10.76	1.00
	4	97	0.593	1.46	1.00	97	0.772	3.39	1.00	97	0.635	1.74	1.00	97	0.529	1.12	1.00
	5	29	**0.158**	0.19	0.28	29	0.540	1.17	0.96	29	**-0.272**	NA	NA	29	0.658	1.92	1.00
	6	76	0.286	0.40	0.99	76	0.809	4.24	1.00	76	0.881	7.40	1.00	76	0.861	6.19	1.00

Table F.2: Post hoc power analysis using G*power a calculated ES, an alpha of .05 with the variables: sine(θ), cosine(θ), and julian date only. **Bold** results are insignificant.

		pH				ANCmeqL				NitratemeqL				SulfatemeqL			
Set	Class	N	Adjusted r^2	Effect Size	Actual Power	N	Adjusted r^2	Effect Size	Actual Power	N	Adjusted r^2	Effect Size	Actual Power	N	Adjusted r^2	Effect Size	Actual Power
1993-2002																	
	1	327	**0.047**	0.049	0.93	327	**0.024**	0.02	0.65	275	0.016	0.02	0.39	325	0.045	0.05	0.92
	2	393	**0.128**	0.15	1.00	392	**0.189**	0.23	1.00	377	**0.017**	0.02	0.55	390	**0.009**	0.01	0.32
	3	400	**0.013**	0.01	0.46	398	**0.000**	0.00	0.06	365	**-0.004**	NA	NA	391	**-0.004**	NA	NA
	4	121	**0.059**	0.06	0.61	120	**0.294**	0.42	1.00	105	**-0.027**	NA	NA	119	**-0.016**	NA	NA
	5	116	**0.051**	0.05	0.52	116	0.381	0.62	1.00	66	0.120	0.14	0.68	116	**-0.010**	NA	NA
	6	110	**0.096**	0.11	0.81	110	**0.075**	0.08	0.69	81	**0.092**	0.10	0.64	110	**-0.009**	NA	NA
2003-2008																	
	1	255	0.040	0.04	0.78	255	**0.001**	0.00	0.07	252	0.061	0.06	0.94	261	0.043	0.04	0.82
	2	289	0.061	0.06	0.96	289	**0.081**	0.09	0.99	296	0.043	0.04	0.87	298	0.014	0.01	0.37
	3	299	**0.020**	0.02	0.52	299	**-0.003**	NA	NA	297	**-0.003**	NA	NA	308	**0.006**	0.01	0.18
	4	119	0.148	0.17	0.97	119	**0.180**	0.22	0.99	121	0.086	0.09	0.80	123	0.023	0.02	0.26
	5	35	**-0.069**	NA	NA	35	**0.337**	0.51	0.93	30	**-0.082**	NA	NA	37	**-0.024**	NA	NA
	6	97	0.081	0.09	0.67	97	**0.094**	0.10	0.74	98	0.046	0.05	0.40	101	0.074	0.08	0.64
2009-2012																	
	1	191	**0.028**	0.03	0.47	191	**0.000**	0.00	0.05	191	**0.018**	0.02	0.31	190	**0.005**	0.01	0.11
	2	212	0.052	0.05	0.82	212	**0.056**	0.06	0.85	212	**0.011**	0.01	0.22	212	**-0.010**	NA	NA
	3	228	**-0.009**	NA	NA	228	**-0.002**	NA	NA	228	**-0.004**	NA	NA	228	**-0.007**	NA	NA
	4	97	0.200	0.25	0.99	97	**0.161**	0.19	0.96	97	**-0.016**	NA	NA	97	**-0.011**	NA	NA
	5	29	**0.218**	0.28	0.58	29	0.466	0.87	0.98	29	**-0.039**	NA	NA	29	**-0.076**	NA	NA
	6	76	0.039	0.04	0.27	76	**0.058**	0.06	0.39	76	**-0.016**	NA	NA	76	**0.007**	0.01	0.08

Appendix G

A priori analysis

G.1 Power graphs

G.1.1 pH

G.1.2 ANC and Nitrate

G.1.3 Sulfate

G.1.4 Time Variables

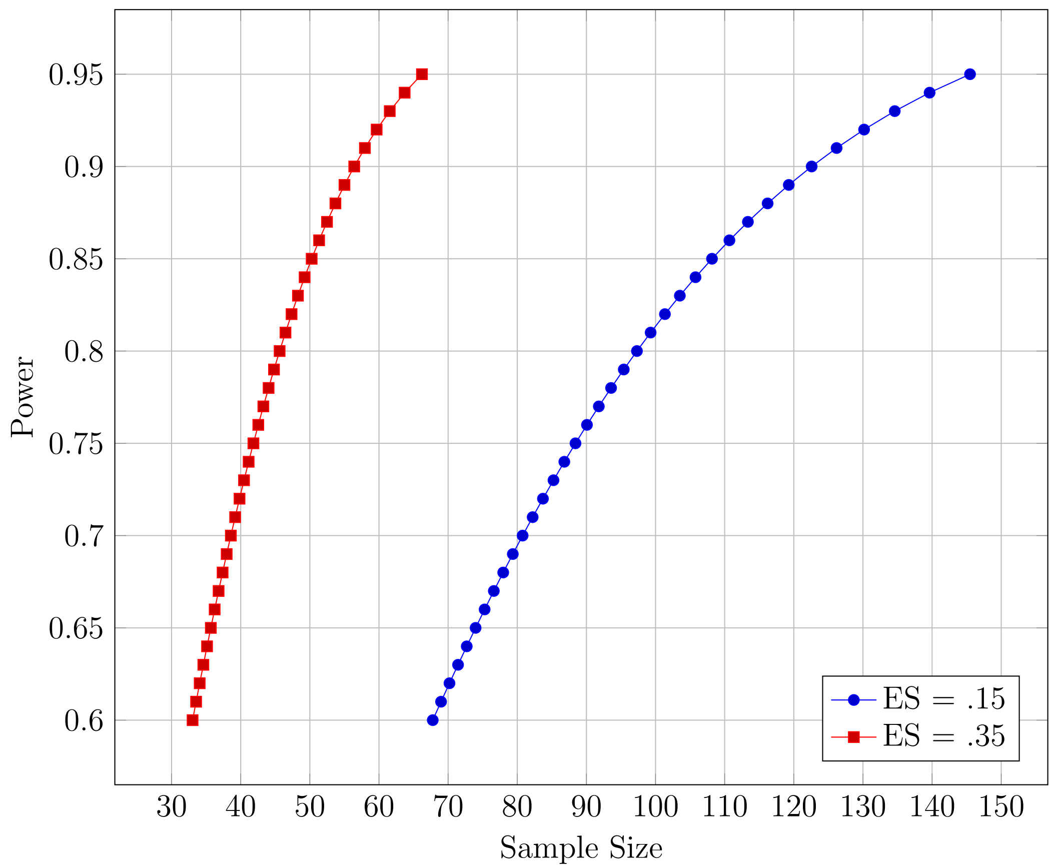


Figure G.1: pH Power Graph. The power is shown as a function of pH

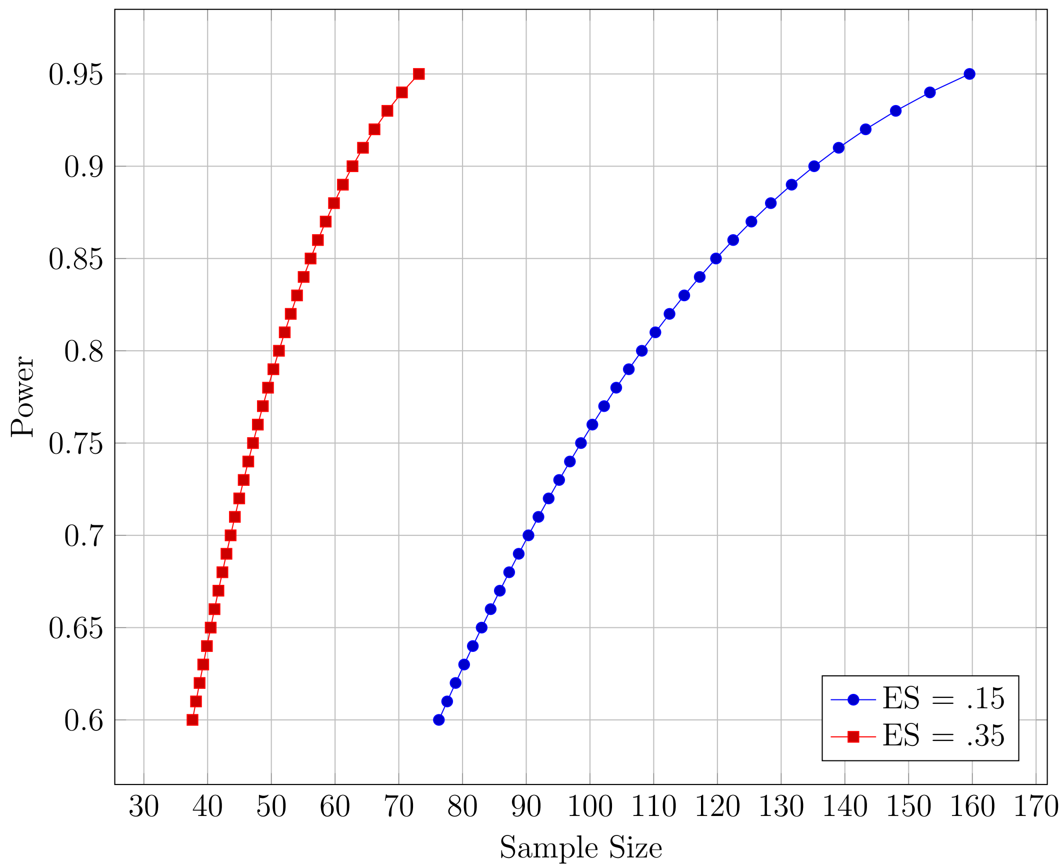


Figure G.2: ANC and Nitrate Power Graphs. The power graphs for ANC and Nitrate are the same because they both have the same number of predictors.

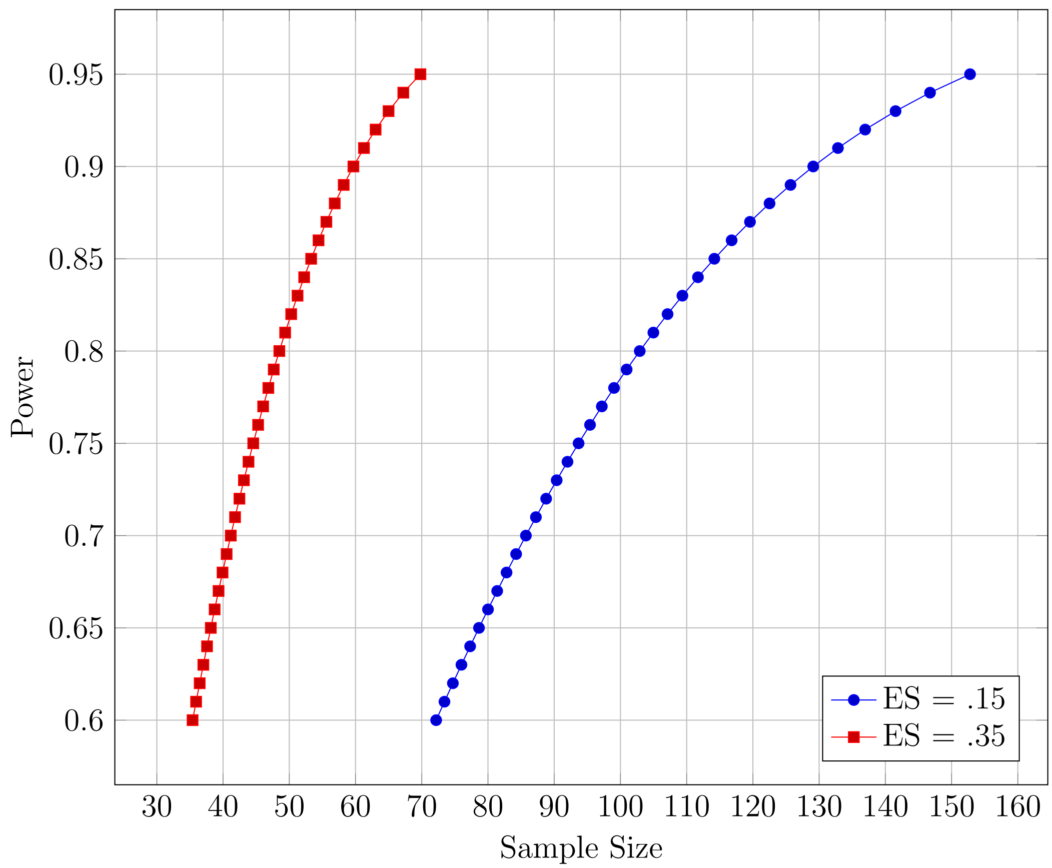


Figure G.3: Sulfate Power Graph

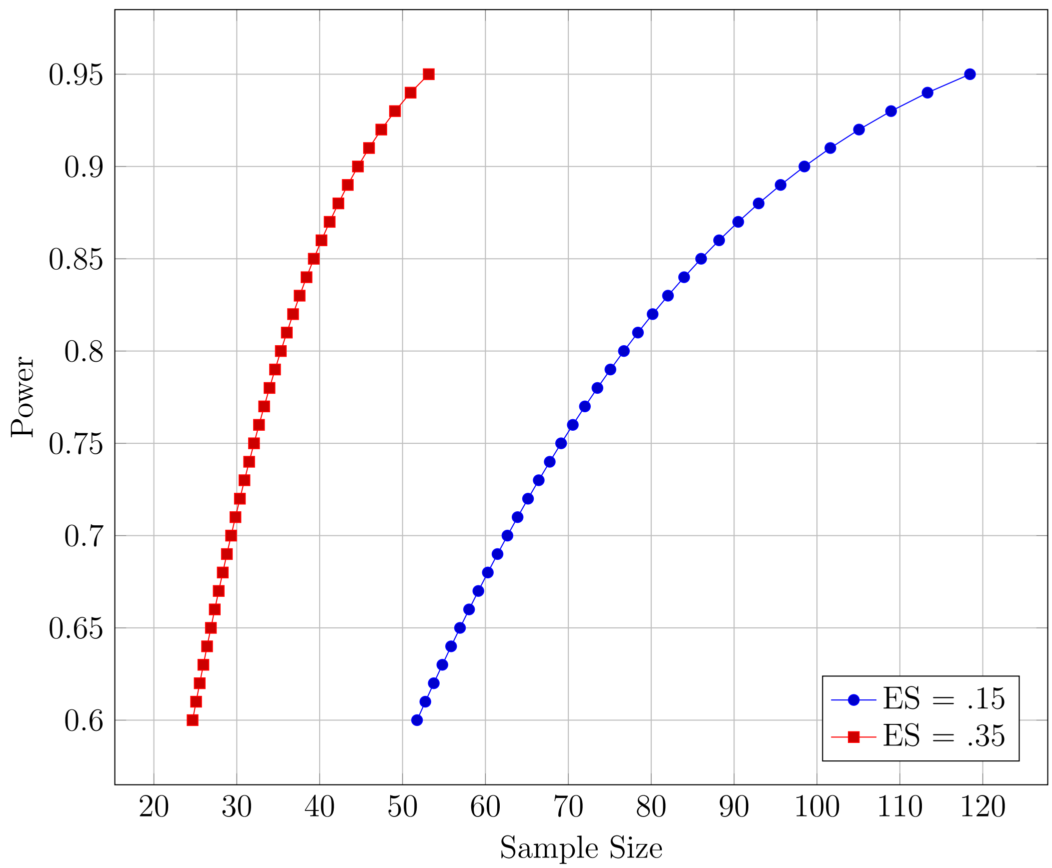


Figure G.4: Time Variables Power Graph

Vita

Tim Pobst was born in Nashville, TN on June 1st 1985 to George and Peggy Pobst. He graduated from Centennial High School near Franklin, TN and was accepted to the University of Tennessee immediately after. He was undecided for three years before deciding to try for a civil engineering degree and he finished it in spring of 2011. He stayed at the University of Tennessee to get a masters degree in environmental engineering under Dr. Schwartz.