2014-03-01

Development and Regional Application of Sub-Seasonal Remote- Sensing Chlorophyll Detection Models

Carly Adeline Hyatt
Brigham Young University - Provo

Development and Regional Application of Sub-Seasonal Remote-Sensing Chlorophyll Detection Models

Carly Adeline Hyatt

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Gustavious Paul Williams, Chair
A. Woodruff Miller
Everett James Nelson

Department of Civil and Environmental Engineering

Brigham Young University

March 2014

ABSTRACT

Development and Regional Application of Sub-Seasonal Remote-Sensing Chlorophyll Detection Models

Carly Adeline Hyatt
Department of Civil and Environmental Engineering, BYU
Master of Science

Remote sensing has been used as an effective chlorophyll-a detection method in inland lakes and reservoirs. Concentration estimates of chlorophyll-a approximate the amounts of algae and phytoplankton in a body of water, can indicate the existence of large blooms and high nutrient loading, and can be used as an indicator of water quality. These biomasses pose potential threats to the quality of the water and the local environment by depleting oxygen, influencing the taste of the drinking water and detrimentally affecting aesthetics and recreation. Deer Creek Reservoir exhibited eutrophic tendencies in the early 1990's, caused by phosphorus pollution. This was made evident by accelerated algae growth. Following remediation efforts, Deer Creek Reservoir, as well as nearby Jordanelle Reservoir have been closely monitored with regular field sampling. These field data have been used to develop remote sensing methods using Landsat images to provide supplementary information for reservoir management. These remote sensing methods allow for mapping of the distribution of chlorophyll-a, which provides spatial distribution average, and maximum estimates of chlorophyll-a concentrations, data and information that are not feasible with in-field sampling.

In this thesis, traditional methods for remote sensing models are discussed, and a novel sub-seasonal approach based on seasonal algal succession is proposed and demonstrated. Each seasonal model is created using a standard stepwise regression using historic field data and the associated Landsat images and is statistically tested for leverage to ensure unbiased model development. These sub-seasonal detection models are applied to 5 reservoirs in the central-Utah region to provide a more comprehensive description of reservoir behavior and water quality trends over the past 30 years. Historic trends of the average and maximum chlorophyll-a estimates are provided for each of the reservoirs. Example color maps are presented to demonstrate the ability of remote sensing to represent the spatial distribution of algae (using chlorophyll as an indicator). Limitations for this approach are discussed, as well as applications for remotely sensed water quality data on a regional scale.

Keywords: remote sensing, chlorophyll-a detection, regional chlorophyll trends, Landsat

ACKNOWLEDGEMENTS

First and foremost, I would like to acknowledge and thank past and current members of the Deer Creek Research Team, for their training, examples, and guidance as I began my research, especially Brent Hargis, Nicolas Gonzalez, Nathan Swain, Pablo Moreno, and Oliver Obregon. I would also like to specifically thank Zola Adjei for her assistance in calibration and preparation of the satellite images, as well as support and enthusiasm in preparing results.

Second, I would like to thank those who have supported the Deer Creek Research Team, with funding as well as data support, Reed Oberndorfer with the CUWCD and Nick Williams at the BOR. I am also extremely grateful to Dr. Gustavious Paul Williams for his words of wisdom, direction, and interest in cowboy poetry, as well as Dr. A. Woodruff Miller, and Dr. Everett James Nelson for their counsel, advice and support.

Finally, I would like to thank my husband, Bradley, and my family for their support and encouragement throughout my program.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

Various remote sensing methods have been used over the past 40 years in engineering and earth sciences to describe and evaluate many land and hydrologic features. The applications vary from identifying geologic formations to evaluating climate change or vegetative health. Remote sensing instruments can be passive, meaning that they measure the reflectance of an object from a source other than the sensor, or active, meaning that they provide the signal that is measured. Perhaps the most familiar passive remote sensing instrument is the human eye. Light is reflected at different wavelengths by different objects, and the eye can sense characteristics such as color, shape, and depth. The practice of using remote sensing in engineering and earth science research relies on this same principle of describing objects through their reflective properties.

The ability of remote sensing to provide useful information depends on accurate models that correlate reflectance values to useful measurements, such as elevation, temperature, or type of vegetation. In order to develop a model for practical applications, remote sensing measurements need to meet several resolution criteria: spectral, spatial, and temporal. Additionally, while certain resolutions may be more suited for an application based on its geographic or temporal scope, an organization may be restricted due to budget or time constraints in developing a model. These considerations are discussed later with respect to the use of NASA's Landsat Satellites as remote sensing instruments and suitability of their data to

support long-term, regional applications. Landsat measures reflectance data in 7 or 8 spectral bands depending on the satellite mission. These data can be correlated with features of interest to develop models or applications. Table 1-1 shows several common applications of remote sensing using Landsat data and the principal bands, or ranges of wavelength, involved in these applications (2013). (Band 6 is not included, as its resolution is different from other bands and is not commonly used in the model development).

Table 1-1 Landsat 5 TM and 7 ETM Band Distribution and Applications

Band Number	Description	Wavelength (μm)	Example Use
Band 1	Blue	0.45-0.52	Bathymetry, soil/vegetation distinction
Band 2	Green	0.52-0.60	Peak vegetation, plant vigor
Band 3	Red	0.63-0.69	Discriminates vegetation slopes
Band 4	Near Infrared	0.77-0.90	Emphasizes biomass content, distinguishing shorelines
Band 5	Short-wave Infrared	1.55-1.75	Moisture content
Band 7	Short-wave Infrared	2.09-2.35	Hydrothermally altered rocks, mineral deposits
Band 8 (Landsat 7 only)	Panchromatic	0.52-0.9	15 meter resolution

One of the most important remote sensing measurements in environmental engineering is with respect to water quality monitoring. There are several water quality parameters that can be measured through remote sensing, including Secchi Disk Transparency (SDT), and chlorophyll-a, a pigment found in algae. Measuring the chlorophyll levels helps describe algal blooms, and can be an indicator for lake and reservoir health due to the responsive relationship of algae and nutrient levels of phosphorus and nitrates (Healey and Hendzel 1980). Remote sensing is an especially effective method for quantifying chlorophyll concentrations when comparing lakes and monitoring over a long period of time (Dekker 1993). There are however, limitations in

developing and applying accurate chlorophyll detection models and correlating those measurements with algal concentrations, which include the resolution of the remote sensor (spectral, spatial, and temporal), and the quality of its products, as well as the quality of the field-measured data. The following study addresses each of these limitations as it explores a unique seasonal approach to develop and apply a regional chlorophyll detection model and use that model to estimate concentrations of various algal populations.

1.1 Background and Limitations of Traditional Methods

Traditional methods for chlorophyll-a detection and quantification have been developed for a range of satellites, each of which has unique spatial, spectral, and temporal resolutions, and data availability. Many traditional methods discussed in literature often use the ratio of the near-infrared and red spectral bands (NIR/R) to develop models which relate spectral responses to algae concentration (Han and Rundquist 1997; Mishra and Mishra 2011). These models typically focus only on the late growing season which is generally the period of highest impact (Falconer 1999; Kutser 2004; Kutser, Metsamaa et al. 2006). The reliance on the NIR/R can be problematic when chlorophyll concentrations are low (Han and Rundquist 1997) and when the remote sensors do not provide the spectral resolution necessary to make the NIR/R ratio useful. An illustration of this difficulty can be seen in Figure 1-1.

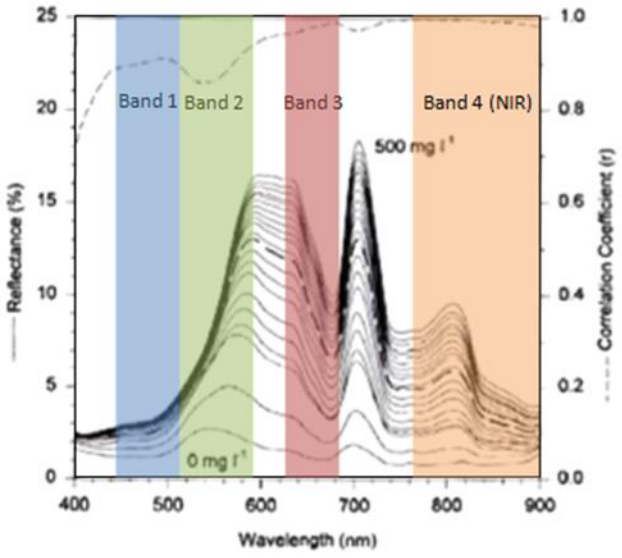

Figure 1-1 Reflectance of Algae-Laden Water Under Varying Suspended Solids Concentration (SSC) with Landsat Band Ranges (Han 1997)

This figure shows the characteristic reflectance values (or spectra) of algae-laden water as the suspended solids concentration (SSC) varies. The characteristic profile of this particular population of algae can be seen at each of the SSCs. There are consistent and distinct local minimum near 680 nm and local maximum at approximately 820 nm within the ranges of bands 3 (red) and 4 (near-infrared) respectively, as well as a maximum near 700 nm. The ratio of a maximum to minimum would serve as clear indicator for concentration if the discrete reflectance values of each were known. This correlation between the band ratios and measured data could then be used to develop a model. However, as noted by the colored overlays, the reflectance measured by the Landsat satellite covers a much larger range of wavelengths than just the minimum or maximum, and does not include the global maximum at all. This means the reflectance measured by Landsat may not provide an accurate representation of the relationship between the maxima and the minima. Also, suspended solids concentrations can significantly impact the measured spectra, and characteristic spectra can vary among populations of algae.

Other remote sensors which may provide the spectral resolution to better resolve the chlorophyll spectra are generally not feasible for long term monitoring applications due to high cost of acquisition and processing and limited data or coverage.

Traditional approaches to model development that only use the later portion of the growing season can also be problematic because they ignore the unique properties and processes that occur in connection with seasonal algal succession and unique algae populations (Cullen 1982; Morel and Berthon 1989), as discussed in the following section. Water quality monitoring agencies commonly focus on the late season because of the blooms of toxic cyanobacteria that occur in the late summer months (Falconer 1999; Kutser 2004; Kutser, Metsamaa et al. 2006). However, diatom blooms in the early growing season can have significant biomass and can negatively affect recreational activities, fish habitats and the overall ecosystem (Spaulding 2007). Attempting to fit a model using only the conventional spectral range ratios and concentration on the late summer ignores unique physical and spectral properties of the algae and characteristics of the body of water under observation during other parts of the growing season. This one-size-fits-all approach is limited and narrow in scope and ability to accurately describe and monitor the health of a lake or reservoir over the entire growing season.

Placing all of the focus on the end of the growing season, while focusing on a sensitive period, ignores the importance of monitoring water quality during other sub-seasons, which may require monitoring. For example, a study of over 145 lakes in Minnesota found that approximately 17% of the chlorophyll maxima occurred between May and June, and that early peaks occurred in all trophic and geometry classes. (Stadelmann, Brezonik et al. 2001). Field measured chlorophyll values were similarly analyzed for Deer Creek and Jordanelle Reservoirs between 1984-2012. Figure 1-2 shows the distribution and frequency of the maximum measured

chlorophyll values over the 28 year time record. Approximately 39% of the maxima occurred for these two lakes in the early summer months, and a total of 60% of the maxima occurred outside of the later months of August and September.

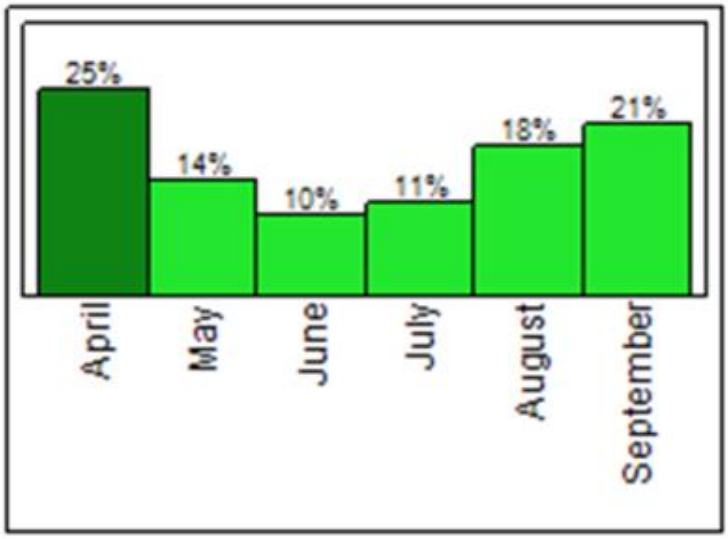

Figure 1-2 Frequency of Field-Measured Chlorophyll Maximums throughout the Growing Season in Deer Creek and Jordanelle Reservoirs from 1984-2012

As can be seen from Figure 1-2, the high percentage of chlorophyll values occurring throughout the early and middle months of the growing season demonstrate the importance and need for accurate monitoring throughout.

Another limitation of traditional remote sensing chlorophyll detection models is the development of a model based on a small number of images and data points. Many studies have made great effort to take the field samples as close as possible to the time of satellite image acquisition, but then base their entire model using the correlation between just a few points and the reflectance values of a few pixels, and often a single time and a limited number of locations. This approach often results in an artificially high correlation during model development and only describes the individual body of water well only under the conditions of that single image. These kinds of

models prove less useful in applications to other bodies of water, or the same body under different conditions (such as different sun elevation, atmospheric interferences, etc.), or in developing historical analyses that may use satellite images that do not have accompanying field data.

1.2 **Sub-Seasonal Approach**

It has been well documented that algae populations experience a phenomenon known as succession in temperate lakes and reservoirs as different types of algae dominate the bodies of water at different periods of the growing season (Prowse and Talling 1958; Castenholz 1960; Morel and Berthon 1989; Stadelmann, Brezonik et al. 2001). Diatoms, or yellow-brown algae, dominate during the early months of the growing season, while green algae and blue-green algae dominate during the middle and late months, respectively. Each algae population exhibits unique physical properties including buoyancy and pigmentation, making the visible or reflective properties (or spectra) unique to each type of algae (Cullen 1982; Morel and Berthon 1989). Studies show that diatoms for instance, have lower absorption coefficients than other phytoplankton, which means that the reflectance and relation to chlorophyll pigmentation will be different from other phytoplankton (Sathyendranath, Watts et al. 2004). More accurate recognition of the different types of algae may drastically improve the reliability of remote sensing chlorophyll detection models (Hansen, Swain et al. 2013). Also, correlation developed for a sub-season should be more accurate as there is less variance in the measured reflectance or spectra.

1.3 **Objective**

This study addresses the issues and limitations associated with traditional methods by developing three sub-seasonal chlorophyll-a detection algorithms. We also demonstrate the use of historic field data for model development even if these data were not specifically collected for remote sensing studies and may not be coincident in time. These three algorithms developed correspond with the three distinct types of algae present in the temperate reservoirs of north-central Utah, namely diatoms in late spring, green algae in summer, and cyanobacteria in late summer. The algorithms are developed using a stepwise linear regression to fit the Landsat-5 Thematic Mapper (TM) and Landsat-7 Enhanced Thematic Mapper (ETM) band reflectance to *in-situ* measured chlorophyll-a. This regression approach includes all band reflectance values (not just the red and near infrared bands) and every possible ratio of these values, which addresses the issue of spectral resolution, and does not limit the model to only the ratio between the bands containing the near-infrared and red values.

Once the algorithms were developed, they were applied to Landsat images from 1984 to 2012 to determine the historical trends of average and maximum chlorophyll levels for five different reservoirs. Between 51 and 77 measurements were used to the model for each reservoir depending on the quality of images available and the life-span of the reservoir. This use of historical data rather than one or two images coincident with field data dramatically increases the scope of application and amount of information gained from analysis. This also provides a large base of data for evaluating conditions over the years.

Nominally, Landsat revisits any given location every 16 days; however some of these images are unusable because of cloud cover or other issues. This study used every available image that had field data collected within either the same day (for early and late seasons) or within 24 hours

(middle season) of the satellite image. Sample dates are discussed in more detail in Section 2.1.2. Landsat images were collected at approximately 10:30 AM, MST. Generally the field data were collected in the morning, near the time of the satellite overpass. This is important, as the spatial algae distributions in these lakes can change significantly because of afternoon winds.

The reservoirs used in this study and corresponding time period used in the analysis are listed below in Table 1-2.

Table 1-2 Scope of Model Application for Each Reservoir

Reservoir	Date of Model Application
Deer Creek	1984-2012
Jordanelle	1993-2012
East Canyon	1984-2012
Echo	1984-2012
Rockport	1984-2012

Long-term model development and application for multiple reservoirs has the potential to provide useful information for enhanced analysis, both spatially (regional rather than single bodies of water) and temporally (over the life of a reservoir, not for a single day). Improving the accuracy and expanding the scope of chlorophyll detection models allows for better monitoring, understanding, and planning for reservoir water quality.

2 METHODS

2.1 Data Collection

Data used in the development of the chlorophyll detection models were field samples at locations shown in Figure 2-1 as well as 25 unique images corresponding with the dates and locations of field measurement. An index of images used for model development is included in Appendix A.

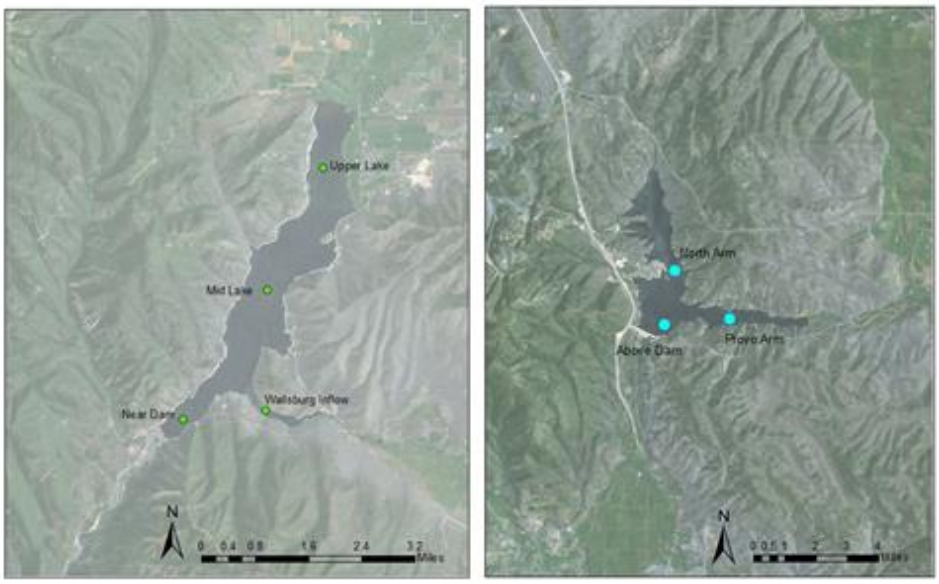

Figure 2-1 Sampling Locations for Deer Creek and Jordanelle Reservoirs

2.1.1 Field Measured Data

The Central Utah Water Conservancy District (CUWCD) collected *in-situ* water quality data from 1984-2012 for Deer Creek Reservoir and 1993-2012 for Jordanelle Reservoir (beginning with the completion of the dam). These measurements included many water quality parameters, such as pH, conductivity, nutrient levels, and chlorophyll-a content. Chlorophyll-a was reported in μg/L and was measured using spectrophotometric analysis and acetone extraction. It should be noted that non-uniform laboratory procedures and pheophytin-a correction procedures had been used by the CUWCD prior to 2013, resulting in possible error in the data. In addition, unpublished studies at BYU indicate that correlations between measured chlorophyll content and algal mass can exhibit significant uncertainty.

The chlorophyll-a measurements used in the development of the detection models were limited to those within 2 meters of the surface. Though other deeper measurements existed, they were not included; the Landsat satellites data could not accurately detect concentrations at greater depths. A total of 84 samples between Deer Creek and Jordanelle reservoirs were used in the development of the detection models.

2.1.2 Remotely Sensed Data

Images in the form of GeoTiffs were downloaded from the USGS Earth Explorer website, from both the Landsat 5 TM and 7 ETM satellites which have 30x30 meter pixels. The Landsat satellites were chosen due to the appropriate spatial and temporal resolution for this application. Given the size of the reservoirs in question, the 30x30 m pixel resolution was suitable for chlorophyll detection applications (Sawaya, Olmanson et al. 2003). Images used in development of the algorithms were limited temporally to the same day as field-data collection for early and late seasons, and within 24 hours of the field-data collection time for the mid-

summer. Spatially the images were limited to those in paths 37 and 38, row 32 which puts the reservoirs away from the edges of the image, where data are distorted because of the slant angle. The temporal limitations allowed for an accurate match of the field conditions by matching sample collection and satellite image acquisition and provided a suitable number of samples for model development. While the Landsat series has a 16-day return period for a particular path and row, the overlapping nature of image acquisition provide a 7-9-7 image availability for features located on the edges of a path, as is the case for Deer Creek Reservoir. A total of 26 unique images (same day images from early and late, and within 24 hour images for the mid-summer season) were used in development of the sub-seasonal models. The temporal limitation was expanded for the mid-summer season in order to increase the number of sample matches. The numbers for each season are shown in Table 2-1. After the model was developed, it was applied to additional images to analyze the time history of the reservoirs.

Table 2-1 Temporal Limitations by Season

Season	Temporal Limitation	Number of Images	Number of Sample Matches
Early	Same day	6	18
Mid	Same day	2	6
Mid	Within 24 hours	14	41
Late	Same day	6	25

2.2 Calibration and Processing

Prior to use in statistical analysis, Landsat images required calibration and processing to accurately represent reflective properties of the water surface. Using the image processing software, ENVI, the raw image files were calibrated based on the sensor (Landsat 5 or 7), the

date of image acquisition, and the sun elevation. The images were then atmospherically corrected using a simple dark subtraction algorithm. Atmospheric correction is necessary to account for interferences such as water vapor or pollution in the atmospheric column. The simple dark subtraction atmospheric correction algorithm locates the pixel with least amount of reflection and assumes that amount of reflection in that pixel is attributed to interferences from the atmospheric column. That amount is then subtracted from all other pixels, thus correcting or removing the reflectance caused by atmospheric interference.

The next step in processing the data was to apply a filter each of the 3x3 pixel grids. This was accomplished by using a common threshold based on the standard deviation of the reflectance values (Bailey and Werdell 2006; Johnson, Strutton et al. 2013). Using the following equation, a new mean was found to represent the reflectance values for each of the six bands in the 90 meter square area surrounding the sample site.

$$Filtered\ Mean = \frac{\Sigma_i(1.5*\sigma-\bar{X})<X_i<(1.5*\sigma+\bar{X})}{N} \quad (2\text{-}1)$$

where $\bar{X}$=unfiltered mean, σ=unfiltered standard deviation, and N=number of values within ±1.5* σ. This approach calculates the mean value for a 3x3 set of pixels, while ignoring pixels with large noise or speckle values.

2.3 **Statistical Analysis**

The first step in statistical analysis was determining whether the field-measured data were normally distributed, as normal distribution is a requirement for regression analyses. The results of a lilliefor distribution test (shown in Table 2-2) found the data for the middle and late seasons to be not normally distributed, while the early season data were normally distributed.

The distributions for the natural log transform of the middle and late season chlorophyll concentrations, however, were normally distributed.

Table 2-2 Distributions for Each Sub-season

Season and measurement	p-value	Distribution
Early; Chlorophyll	0.0535	Normal
Early; Natural Log of chlorophyll	0	Not normal
Mid; Chlorophyll	0.0027	Not normal
Mid; Natural Log of Chlorophyll	0.5729	Normal
Late; Chlorophyll	0	Not normal
Late; Natural Log of Chlorophyll	0.2009	Normal

Using the appropriate data, a forward stepwise regression was run on each of the sub-seasons to determine which parameters should be included in the predictive model and their respective coefficients. A p-value of 0.25 was used as a threshold for the amount of effect that a parameter must have to be included in the model on a forward step, while a threshold p-value of 0.1 was used to determine whether a parameter should be removed during a backward step. Initial regression results are reported in Appendix B.

Following the initial regression, a series of tests of influence were applied to determine whether a sample point had large influence on the model, even if it were not obvious graphically. A sample was categorized as having large leverage if it exceeded the following criteria for 2 or more tests of influence.

Hats test: $h_i > \frac{2p}{n}$

where,

$$h_i = \frac{1}{(n-1)}\left(\frac{X_i - X}{s_x}\right)^2 + \frac{1}{n} \quad (2\text{-}2)$$

p= number of parameters in the model and n=number of samples included

Studentized Residual: $abs(studres_i) > \pm 2$

where,

$$studres_i = \frac{res_i}{\sigma\sqrt{1-h_i}} \quad (2\text{-}3)$$

Res= residual of a case

Cook's Distance: $D_i > 1$

where,

$$D_i = \frac{1}{p}(studres_i)^2\left(\frac{h_i}{1-h_i}\right) \quad (2\text{-}4)$$

One case of high leverage was identified in both the early and late sub-seasons, resulting in a new stepwise regression and updated model after the case was removed from the regression. An example of the leverage test results is shown in Figure 2-2 for the early sub-season data. The remainder of the leverage test results can be found in Appendix B. Figure 2-2a shows the points which have high leverage; one data point has high leverage for two of the tests, while no points are flagged as having high leverage in Figure 2-2b, after the data point is excluded from the regression. Difference in scales should be noted when comparing the results of the leverage tests.

The results of the final regressions are shown in the plots of actual vs. predicted chlorophyll or log(chlorophyll) values in Figure 2-3. The original Actual v. Predicted plots for the models prior to excluding points that failed the leverage test are included in Appendix B. By examining the leverage of individual cases within the regression, the models were improved by a 10% increase in the correlation coefficient value for the early model and a 2% increase for the late model.

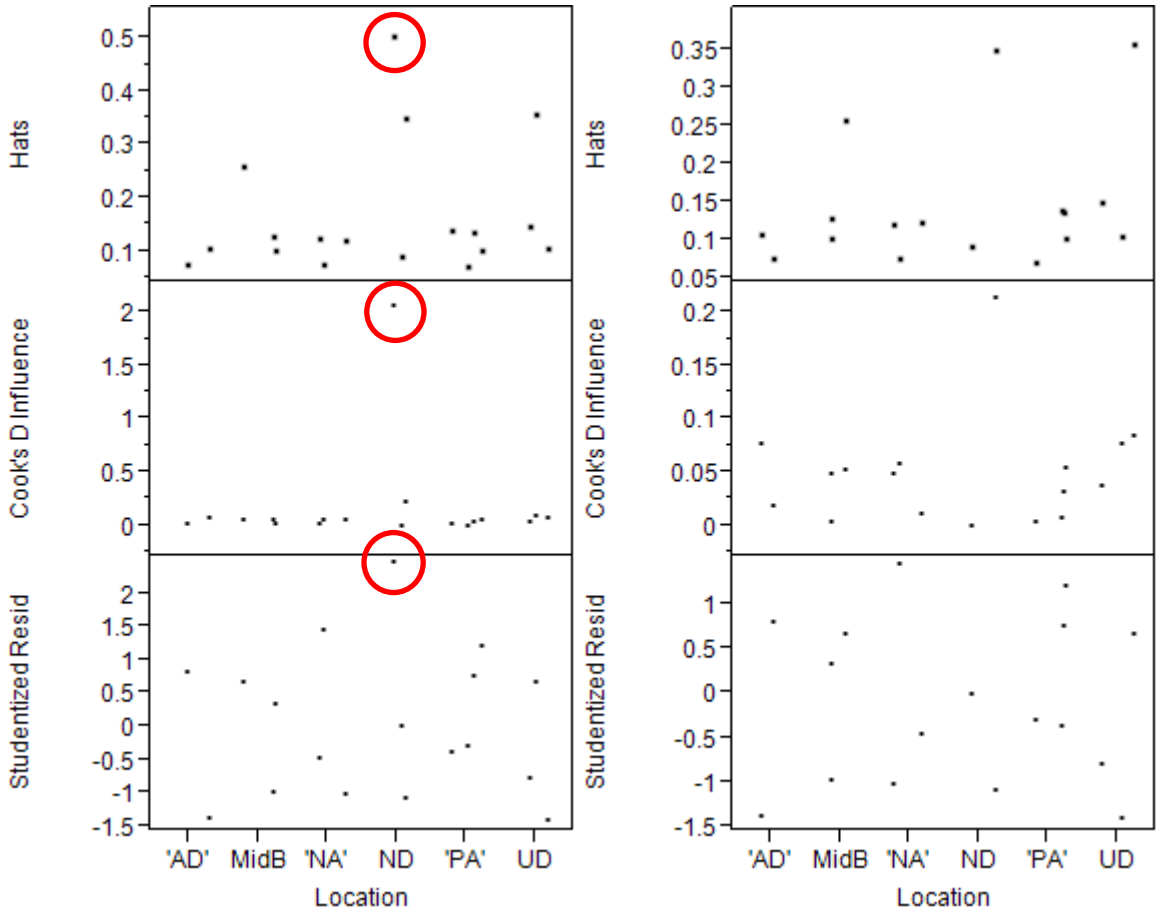

Figure 2-2 Example of Leverage Tests for Early Sub- Seasonal Model a) Including all data points, and highlighting the point which fails the leverage tests b) Excluding points that failed at least two of the leverage test

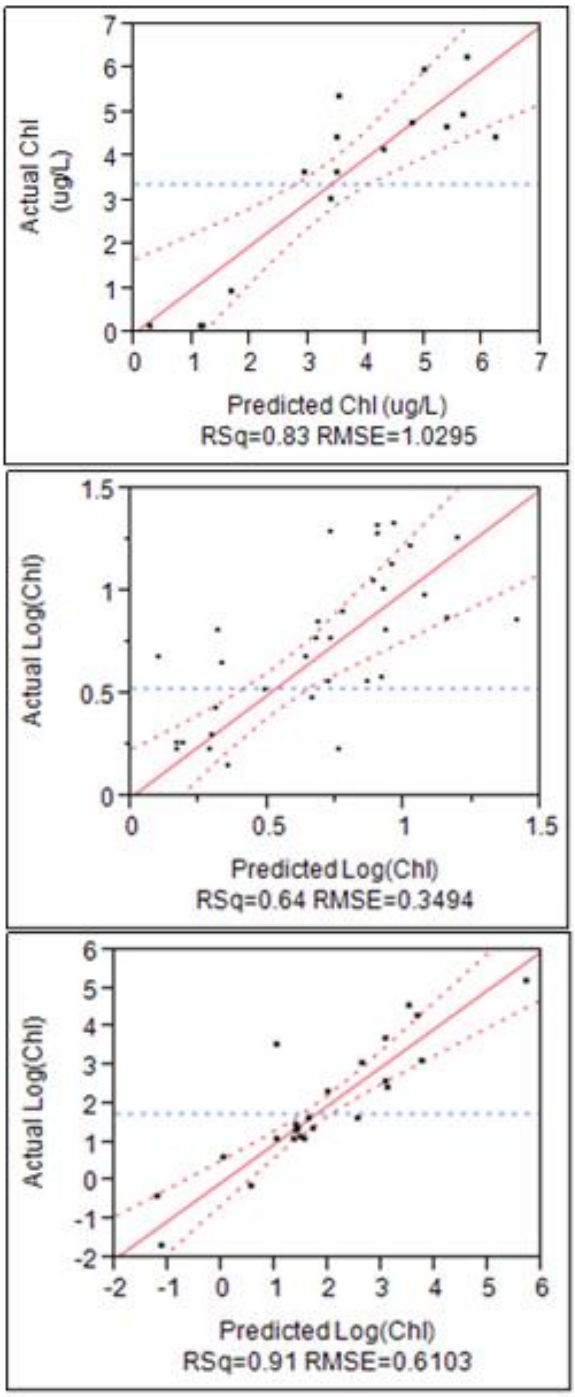

Figure 2-3 Actual v. Predicted Values for Early, Mid, and Late Sub-seasons

The algorithms for each of the sub-seasons are included below.

Early summer (April – May)

$$Chlorophyll\ \left(\frac{\mu g}{L}\right) = 10.8633 + 0.8396 * \left(\frac{B3}{B1}\right) - 0.82 * \left(\frac{B4}{B7}\right) - 9.0696 * \left(\frac{B5}{B4}\right) + 0.9616 * \left(\frac{B7}{B5}\right) \quad (2\text{-}5)$$

Mid-summer (June – July)

$$Log\left(Chlorophyll\ \left(\frac{\mu g}{L}\right)\right) = 1.5504 - 98.1679 * (B5) + 102.9181 * (B7) - 0.435 * \left(\frac{B2}{B4}\right) - 0.2639 * \left(\frac{B5}{B3}\right) \quad (2\text{-}6)$$

Late summer (August – September)

$$Log\left(Chlorophyll\left(\frac{\mu g}{L}\right)\right) = -2.1451 - 36.32 * (B1) + 1.226 * \left(\frac{B2}{B5}\right) - 1.4604 * \left(\frac{B3}{B5}\right) + 0.3133 * \left(\frac{B4}{B7}\right) + 6.3214 * \left(\frac{B5}{B4}\right) \quad (2\text{-}7)$$

3 RESULTS AND DISCUSSION

3.1 Model Comparisons

As can be seen in Figure 2-3, the regression coefficients for each of the sub-seasonal models are relatively high, at 0.83, 0.64, and 0.91 for the early, mid, and late sub-seasons respectively. When compared to the regression coefficients for the individual sub-seasons, the application of the late algorithm to early and mid-summer months were significantly lower. Figure 3-1 and Figure 3-2 illustrate the inability of a model developed using only data from the end of the growing season to accurately estimate the chlorophyll during other months of the growing season. This demonstrates how the reflectance properties of the algae change with the season and how sub-seasonal models can provide more accurate estimates.

Figure 3-3 shows the results of a model developed using stepwise regression and fitting all of the data for the entire growing season – rather than for individual seasons. Though it is a better fit than the late algorithm applications to other sub-seasons, the correlation coefficient is only R^2=0.42 compared to 0.83, 0.64, or 0.91 from the sub-seasonal algorithms.

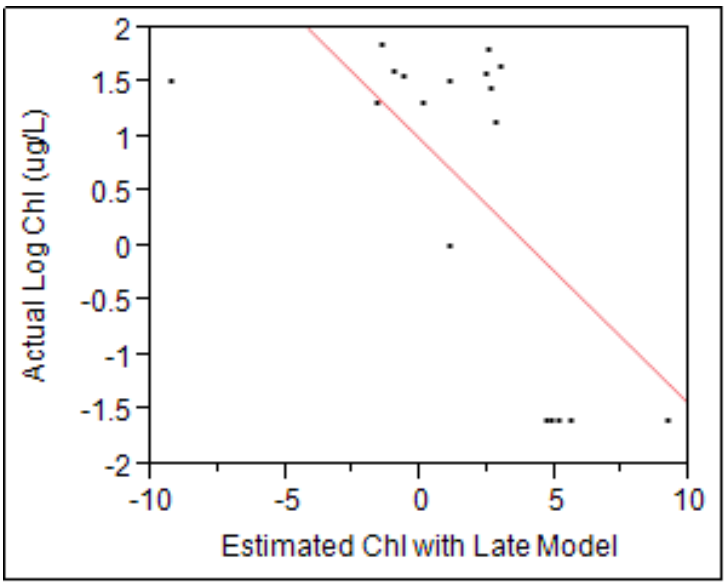

Figure 3-1 Actual v. Predicted for Early Data with the Late Model Applied. R^2 = 0.443 Compared to 0.83 of the Early Data with the Early Model Applied.

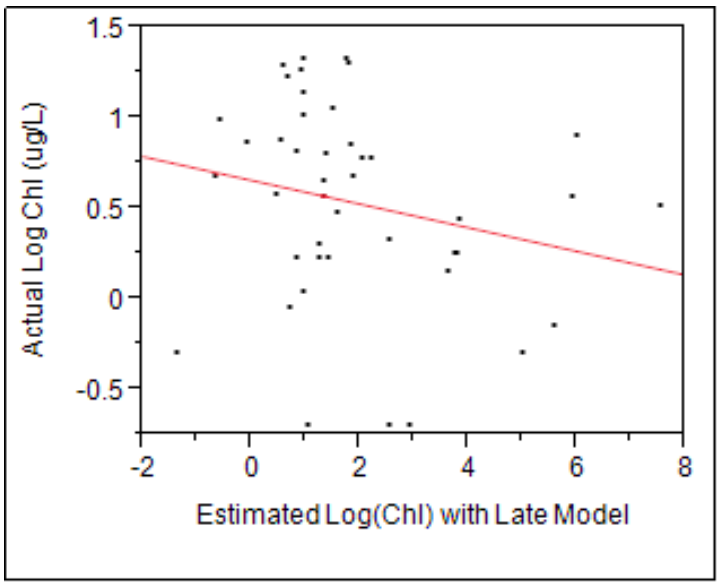

Figure 3-2 Actual v. Predicted for Mid-summer Data with the Late Model Applied. R^2=0.049 Compared to 0.64 of the Mid-summer Data with the Mid-summer Model Applied.

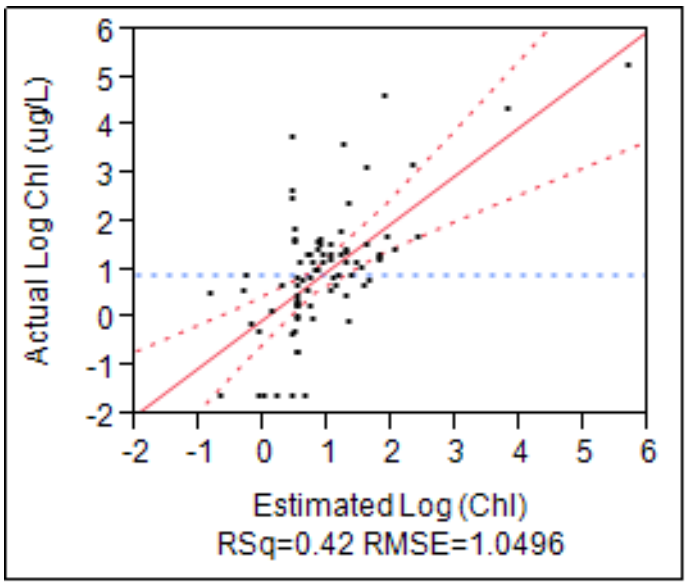

Figure 3-3 Actual v. Predicted for the Entire Growing Season Using a Single Model

It should also be noted that the data for the early sub-season was normally distributed, but the mid and late seasons required log transformation before they could be considered normally distributed. When all of the data was combined, a log transformation was required to be normally distributed.

3.2 Sub-Seasonal Model Results

Models were then applied to images from throughout the growing season for each of the five reservoirs. The temporal analysis covered the period of satellite or reservoir record, with images varied from reservoir to reservoir due to image quality, cloud cover, ice cover, or construction of the reservoir; however, the range of dates included in the application span from 1984 to 2012. The appropriate seasonal model was applied to the each image, and then using unsupervised classification, the image was masked to only include the reservoir and exclude the land areas. This provided a data set covering the entire reservoir that consisted of the estimated

chlorophyll concentrations. I then calculated measures of central tendency and variation on these reservoir pixels. Table 3-1 summarizes the results of these statistical measures for each of the sub-seasons and the entire growing season for each reservoir. The mode was chosen as a measure of the average chlorophyll value instead of the mean as it is less influenced by outliers, and provides a more accurate description for the most commonly occurring value in the reservoir. The maximum value was calculated as follows:

$$Max\ Chl = \bar{x} + 1.96(\sigma) \quad (3\text{-}1)$$

where,

$\bar{x} = mean\ of\ estimated\ chlorophyll$ and σ = standard deviation.

The purpose of calculating the maximum using this approach is to avoid using a single pixel with a high value which could be contributed to noise in the image rather than a true high chlorophyll value. The value located approximately two standard deviations from the mean represents the 95th percentile, or a value that is greater than 95 percent of all of the observations. The standard deviation is used as a measure of the variability of chlorophyll levels in the entire body of water.

The following plots show the results by season of the mode and maximum estimated chlorophyll levels for the entire reservoir for Jordanelle over a 19 year period (1993-2012). Figure 3-4 and Figure 3-5 show the trends for the early summer months, Figure 3-6 and Figure 3-7 show the trends for the mid-summer months, and Figure 3-8 and Figure 3-9 show the trends for the late summer months. A collection of historical trends for the other reservoirs for each season is provided in a similar manner in Appendix D.

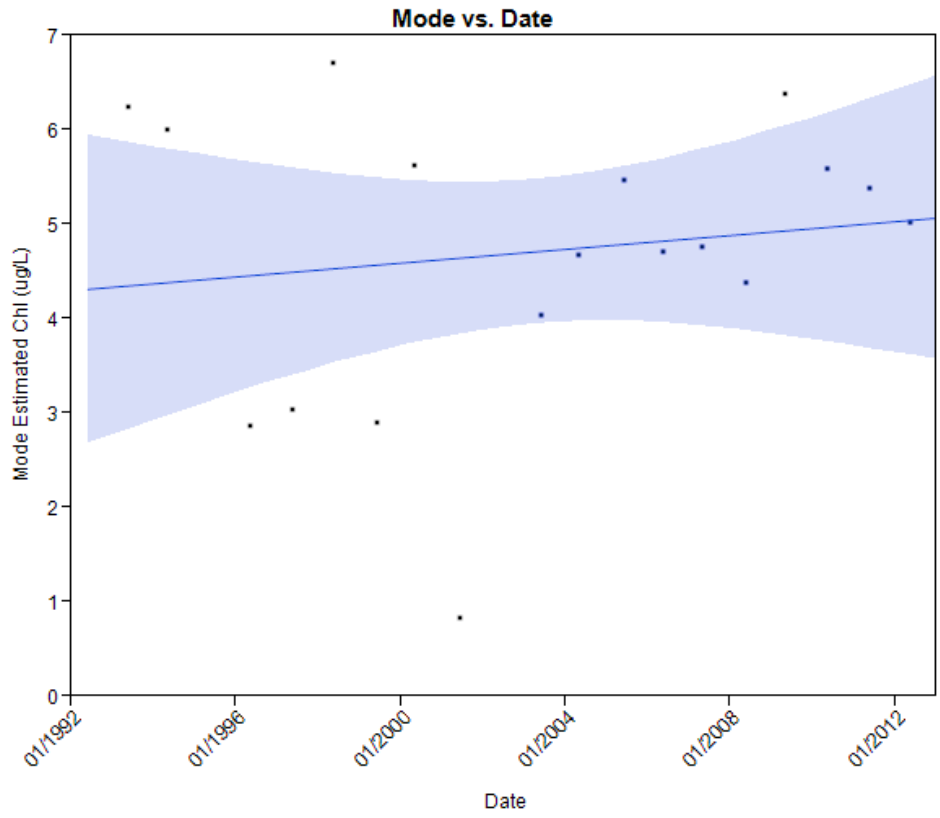

Figure 3-4 Average (mode) Estimated Chlorophyll Levels in Jordanelle for the Early Summer Months from 1993-2012

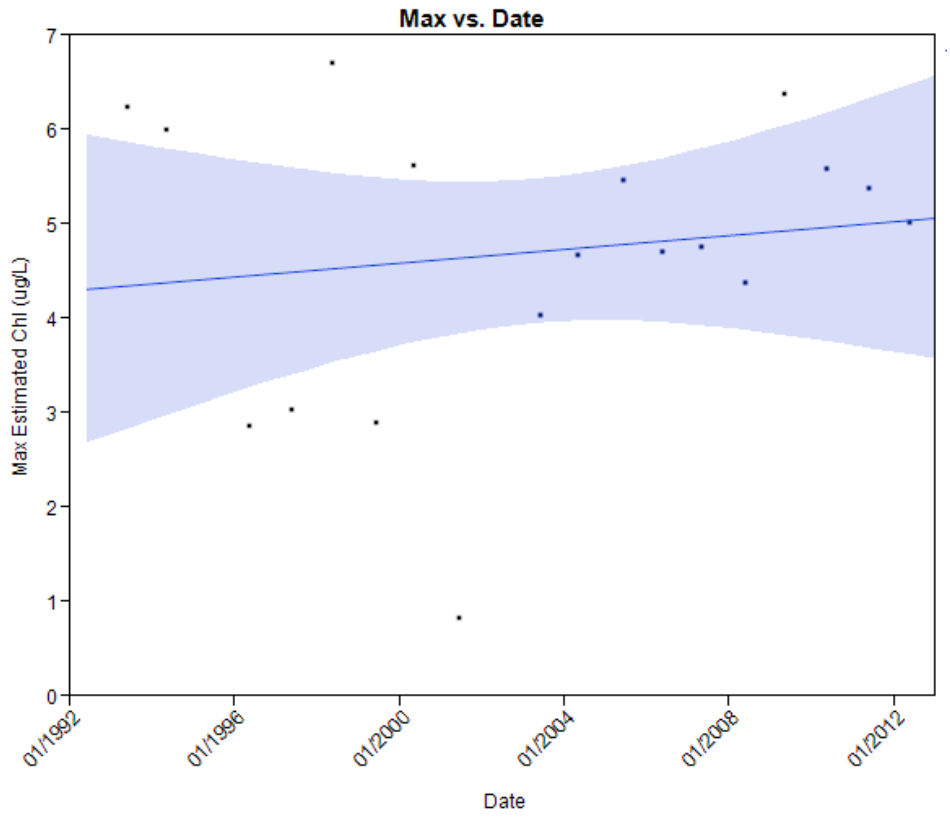

Figure 3-5 Maximum Estimated Chlorophyll Levels in Jordanelle for the Early Summer Months from 1993-2012

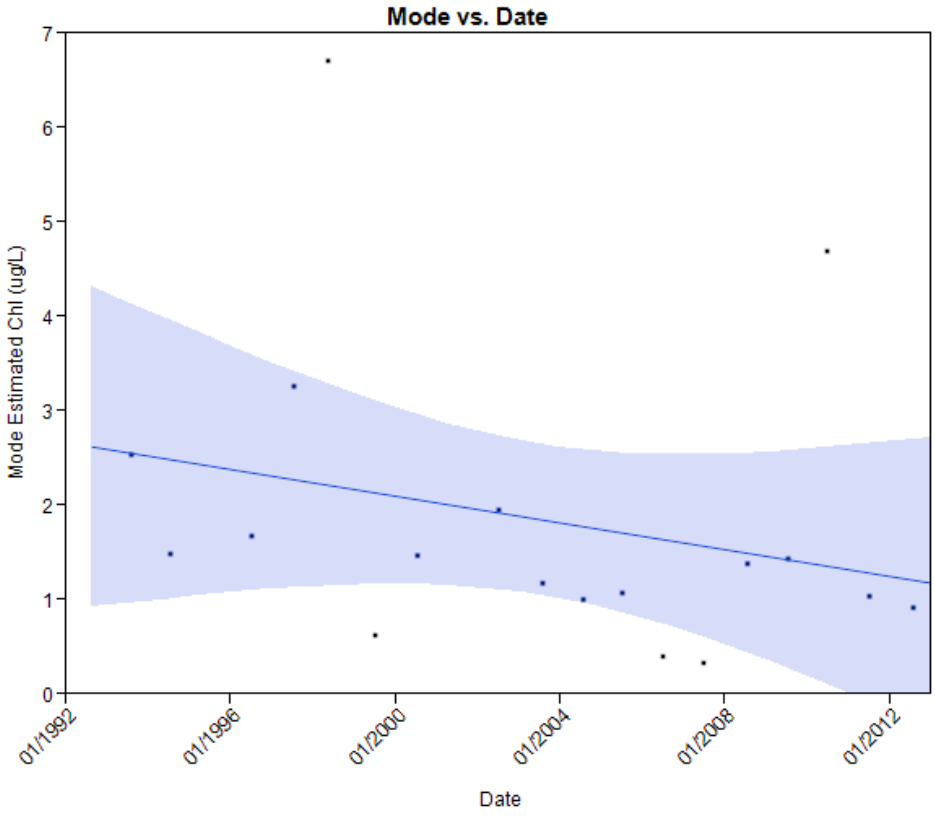

Figure 3-6 Average (mode) Estimated Chlorophyll Levels in Jordanelle for the Mid-summer Months from 1993-2012

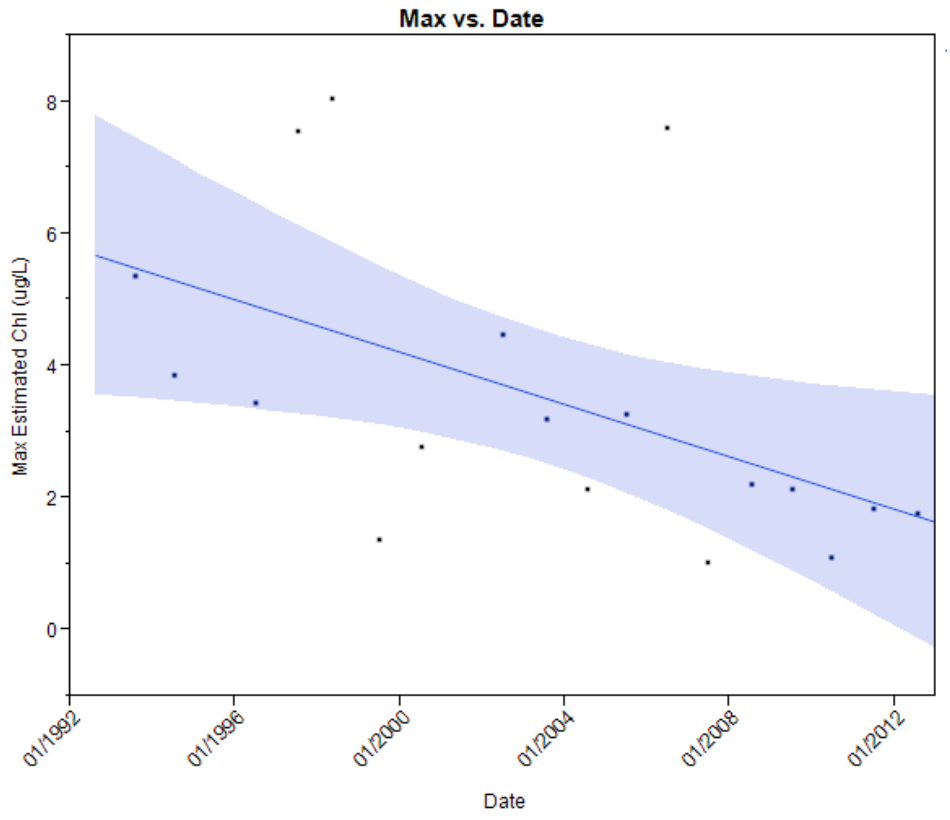

Figure 3-7 Max Estimated Chlorophyll Levels in Jordanelle for the Mid-summer Months from 1993-2012

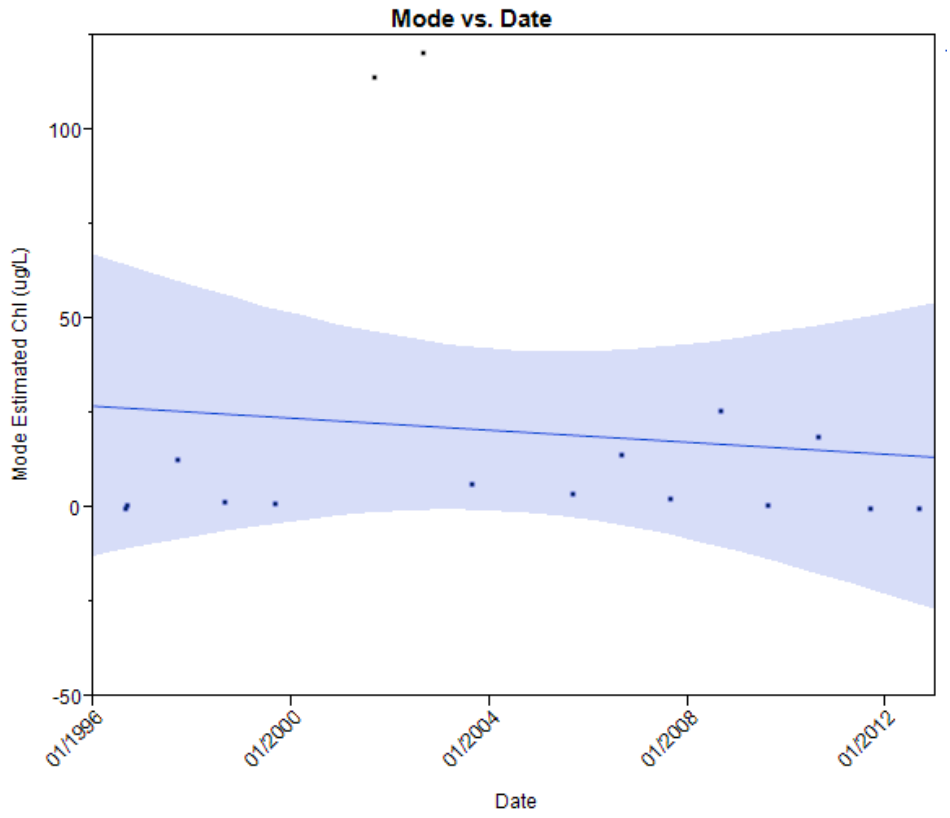

Figure 3-8 Average (mode) Estimated Chlorophyll Levels in Jordanelle for the Late Summer Months from 1993-2012

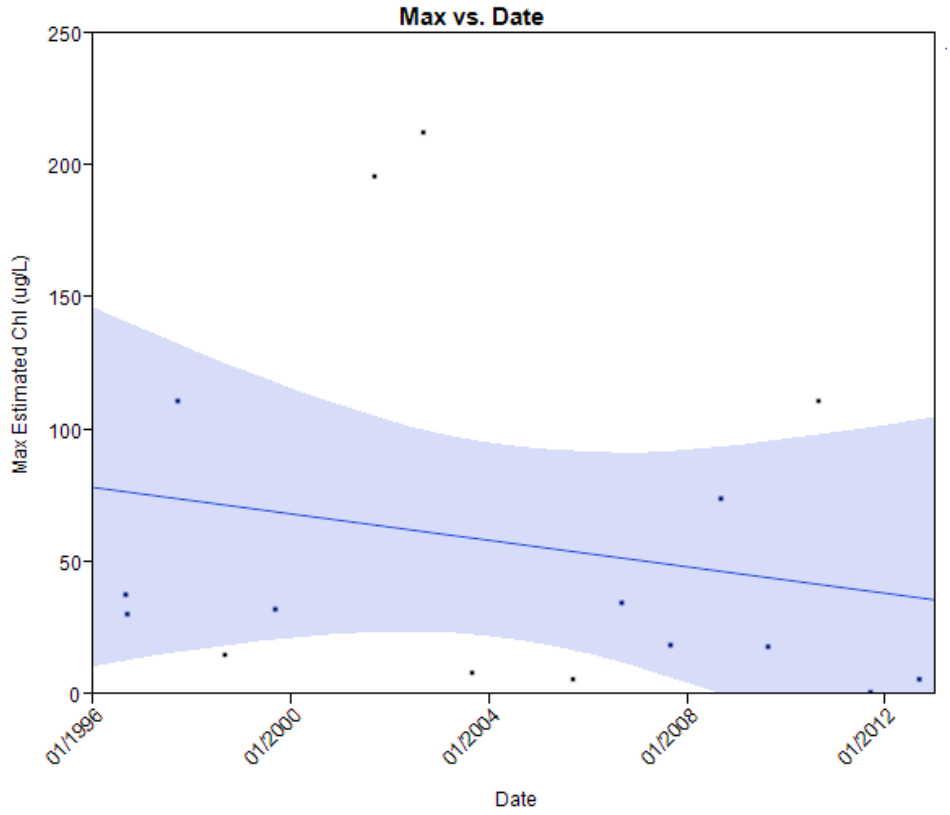

Figure 3-9 Max Estimated Chlorophyll Levels in Jordanelle for the Late Summer Months from 1993-2012

Jordanelle was chosen to demonstrate the benefits of a seasonal analysis and reinforce the importance of monitoring reservoirs throughout the growing season, rather than only during the late summer months. These trends show change in the reservoir since the dam was completed until the present. As can be seen in Figure 3-4 and Figure 3-5, the trend for the early growing season over 19 years is increasing, indicating that diatom concentrations are increasing, while Figure 3-6 through Figure 3-9 show decreasing trends for the mid and late summer months indicating decreasing concentrations of green and blue-green algae. It should be noted that the trends differ in degree for each reservoir and each sub-season. This indicates that each reservoir is unique, and each sub-season is unique and requires monitoring.

The results in Table 3-1 present the slope of the linear trend in degree of change per year for each trend line for the average, maximum, and standard deviations of estimated chlorophyll levels in each of the five reservoirs. Values are highlighted to emphasize which reservoirs and seasons are increasing in the average, maximum or variation of chlorophyll levels over time. Three of the five reservoirs show positive trends for the early months in both maximum levels, albeit by less than 0.1 μg/L per year. The most drastic changes are occurring in the late summer months, as the trends for maximum levels are decreasing (with the exception of Echo Reservoir) by an average of 2.56 μg/L per year.

Table 3-1 Slope of Trend Lines for Each Sub-Season and Reservoir

Reservoir and Season	Mode (ug/L/year)	Max (ug/L/year)	Standard Deviation (ug/L/year)
Deer Creek – Early	-0.05	-0.10	-0.04
Deer Creek – Mid	-0.06	-0.12	-0.03
Deer Creek – Late	-2.14	-3.48	-1.33
Jordanelle – Early	0.06	0.09	0.02
Jordanelle – Mid	-0.07	-0.19	-0.05
Jordanelle – Late	-0.79	-2.50	-0.81
East Canyon - Early	0	-0.12	-0.05
East Canyon – Mid	-0.04	-0.25	-0.10
East Canyon – Late	0.17	-2.08	-0.65
Echo – Early	0.06	0.04	0
Echo – Mid	-0.04	-0.17	-0.07
Echo – Late	0.30	0.18	-0.01
Rockport – Early	-0.04	0.02	0.02
Rockport – Mid	-0.05	-0.10	-0.03
Rockport - Late	0.18	-2.19	-0.86

Another way of using the model to examine the quality of water on a regional basis can be demonstrated by applying the algorithm to a reservoir image to show the spatial distribution of the algae. All of the reservoirs used in this study are within Landsat path 38, row 32. The spatial distribution of the chlorophyll within each reservoir on May 21, 2003 can be seen in the color maps of Figure 3-10.

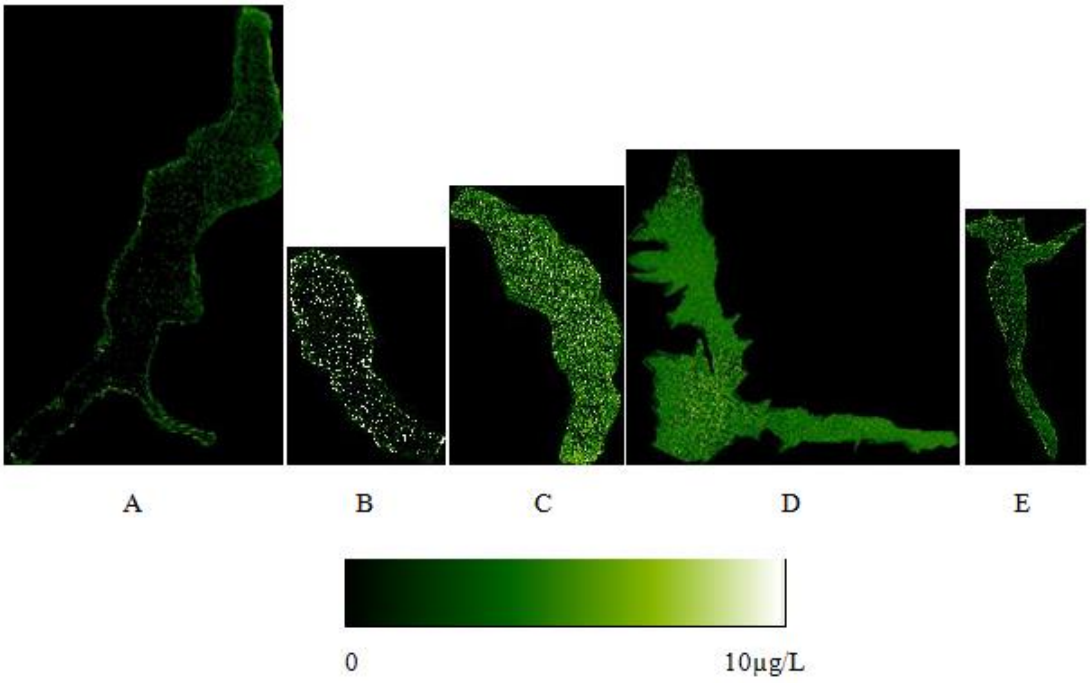

Figure 3-10 Spatial Distribution Maps of Each Reservoir from May 21, 2003 Landsat Image (A. Deer Creek, B. Rockport, C. Echo, D. Jordanelle, E. East Canyon)

This example of spatial distribution color maps shows that each of the reservoirs behaves quite differently. For instance, Deer Creek has relatively low values throughout the reservoir, with high levels concentrated along the shoreline and upper reservoir (due to both wind and nutrient inputs), while Rockport has highly variable levels throughout, Echo has relatively

constant levels, Jordanelle has a slightly higher concentration just above the dam, and lower concentrations in the upper arm, and East Canyon is fairly constant throughout. These spatial distribution maps have the potential to identify reservoirs that might require additional mitigation efforts or locations within reservoirs that should have increased monitoring.

4 CONCLUSIONS

The following discussion summarizes many of the benefits and the limitations that were overcome by developing a remote sensing chlorophyll detection model from a long-term, sub-seasonal approach. These include improvements made in the developing process, improvements in the model accuracy and reliability, and enhanced information and insights provided by this unique approach to developing and applying the model.

The first obstacle overcome in this approach was choosing the most appropriate sensor that would fit the resolution requirements, as well as the time and economic constraints of the study. The use of the Landsat series as the source of remote sensing data met these requirements. In a single image, all five reservoirs of interest could be examined. This improved the efficiency of model development as the time spent calibrating and processing the images was reduced because fewer images needed processing. The pixel size of 30 meters was also adequate for each of the reservoirs of interest. Additionally, the Landsat images provided a low-cost option for remote sensing model development as the images were completely free to download. This economic factor is important for agencies or organizations looking to enhance their water quality monitoring efforts on a tight budget.

Second, the stepwise regression used to develop the models, considering all possible bands and band ratios, overcame a major limitation of the spectral resolution of the Landsat images. This technique of statistical analysis provided algorithms with much more accurate

estimation capabilities than the traditional band ratios, and was much more appropriate to the use of Landsat reflectance values. The further statistical analysis, using tests of leverage and filtering the data for noise, also improved the accuracy of the model.

The main improvement for accuracy and reliability comes from the unique sub-seasonal division of data and development of three algorithms for the growing season. As shown in Figure 3-1 and Figure 3-2, the algorithm developed for the late summer months does not have accurate predictive capabilities for the early or mid-summer months, even though it has a high correlation for the chlorophyll levels in the late months. The relatively high correlations (R^2 = 0.83, 0.64, and 0.91) for each of the sub-seasons indicates that the model using unique algorithms is much more appropriate for describing the water quality with respect to chlorophyll throughout the growing season than a single model for the growing season(R^2=0.42). As noted in the single model, the distribution of data required a log transform before it could be considered normally distributed, however, the data from the early months is already normally distributed. Splitting up the growing season and fitting the reflectance values to the true measured amount or log transformation of this value depending on the actual distribution of the sub-season allows for a better representation of the data and provides a much better fit.

In addition to improvements in model development and accuracy, this kind of regional analysis can provide valuable information and insights that is not available from point sampling. The figures in Appendix D show average, maximum, and standard deviations of the estimated chlorophyll values for the entire reservoir for each reservoir included in the study. None of these measurements of water quality can be accurately obtained from practical point sampling campaigns but this remote sensing approach allows a regional analysis. Through the application of the algorithms, these measures of water quality can be determined for the entire region, and

comparisons can easily be made. For instance, the average values of chlorophyll in the late summer months throughout the entire reservoirs of Deer Creek and Jordanelle are decreasing, while Rockport, East Canyon, and Echo Reservoirs are increasing each year. This helps to identify which reservoirs require further monitoring and mitigation efforts, as well as determines when those efforts should be focused. The spatial maps shown in Figure 3-10 also illustrate how an agency might use these model applications to identify problem areas and evaluate efforts.

In conclusion, the described methods for developing a sub-seasonal model can greatly improve the accuracy of remote sensing chlorophyll detection models throughout the entire growing season. Utilization of historical data and the large spatial coverage of Landsat images results in more accurate historical analyses and comparisons of an entire region through historical trends and spatial distribution maps. The products of these kinds of analyses provide valuable information to water resource managers, and have the potential to supplement and greatly enhance current monitoring and mitigation practices.

REFERENCES

(2013, July 18, 2013). "Ask Landsat." Retrieved December 11, 2013, from http://landsat.usgs.gov.

Bailey, S. W. and P. J. Werdell (2006). "A multi-sensor approach for the on-orbit validation of ocean color satellite data products." Remote Sensing of Environment **102**(1): 12-23.

Castenholz, R. W. (1960). "Seasonal changes in the attached algae of freshwater and saline lakes in the Lower Grand Coulee, Washington." Limnology and oceanography: 1-28.

Cullen, J. J. (1982). "The deep chlorophyll maximum: comparing vertical profiles of chlorophyll a." Canadian Journal of Fisheries and Aquatic Sciences **39**(5): 791-803.

Dekker, A. G. (1993). "Detection of optical water quality parameters for eutrophic waters by high resolution remote sensing."

Falconer, I. R. (1999). "An overview of problems caused by toxic blue–green algae (cyanobacteria) in drinking and recreational water." Environmental Toxicology **14**(1): 5-12.

Han, L. and D. C. Rundquist (1997). "Comparison of NIR/RED ratio and first derivative of reflectance in estimating algal-chlorophyll concentration: a case study in a turbid reservoir." Remote Sensing of Environment **62**(3): 253-261.

Hansen, C., N. Swain, et al. (2013). "Development of Sub-Seasonal Remote Sensing Chlorophyll-A Detection Models." American Journal of Plant Sciences **4**: 21.

Healey, F. and L. Hendzel (1980). "Physiological indicators of nutrient deficiency in lake phytoplankton." Canadian Journal of Fisheries and Aquatic Sciences **37**(3): 442-453.

Johnson, R., P. G. Strutton, et al. (2013). "Three improved satellite chlorophyll algorithms for the Southern Ocean." Journal of Geophysical Research: Oceans **118**(7): 3694-3703.

Kutser, T. (2004). "Quantitative detection of chlorophyll in cyanobacterial blooms by satellite remote sensing." Limnology and oceanography: 2179-2189.

Kutser, T., L. Metsamaa, et al. (2006). "Monitoring cyanobacterial blooms by satellite remote sensing." Estuarine, Coastal and Shelf Science **67**(1-2): 303-312.

Mishra, S. and D. R. Mishra (2011). "Normalized difference chlorophyll index: A novel model for remote estimation of chlorophyll-< i> a</i> concentration in turbid productive waters." Remote Sensing of Environment.

Morel, A. and J. F. Berthon (1989). "Surface pigments, algal biomass profiles, and potential production of the euphotic layer: Relationships reinvestigated in view of remote-sensing applications." Limnology and oceanography: 1545-1562.

Prowse, G. and J. Talling (1958). "The seasonal growth and succession of plankton algae in the White Nile." Limnology and oceanography: 222-238.

Sathyendranath, S., L. Watts, et al. (2004). "Discrimination of diatoms from other phytoplankton using ocean-colour data." Marine ecology progress series **272**: 59-68.

Sawaya, K. E., L. G. Olmanson, et al. (2003). "Extending satellite remote sensing to local scales: land and water resource monitoring using high-resolution imagery." Remote Sensing of Environment **88**(1-2): 144-156.

Spaulding, S. A., and Elwell, L (2007). Increase in Nuisance Blooms and Geographic Expansion of the Freshwater Diatom *Didymosphenia geminate*. US Geological Survey Open-File Report. **38:** 4-17.

Stadelmann, T. H., P. L. Brezonik, et al. (2001). "Seasonal patterns of chlorophyll a and Secchi disk transparency in lakes of East-Central Minnesota: Implications for design of ground-and satellite-based monitoring programs." Lake and Reservoir Management **17**(4): 299-314.

APPENDIX A. INDEX OF DATA

Table A-1 provides a list of the images used in developing the three sub-seasonal models. The number of usable sample points refers to the number of samples from sites in both Deer Creek and Jordanelle Reservoirs that both a recorded field measurement and were free of cloud cover.

Table A- 1 Index of Landsat 5 and 7 Images Used in Model Development

Date of Image Acquisition	Number of Usable Sample Points
6/18/1984	3
6/12/1985	4
8/15/1985	4
4/21/1986	3
7/1/1986	4
9/22/1987	3
7/18/1989	3
7/16/1991	3
9/17/1991	3
6/29/1994	3
8/17/1994	6
7/19/1995	3
6/19/1996	1
8/29/1996	6
5/28/1997	3
4/22/1998	2
5/18/1999	4
6/22/2000	3
8/1/2000	3
5/31/2001	3
6/12/2002	2
6/19/2002	6
5/21/2003	3
7/17/2003	1
6/21/2005	3
6/13/2012	2

APPENDIX B. INITIAL MODELS AND TESTS OF LEVERAGE

The following figures show results from initial stepwise regressions and the corresponding tests of leverage for each of the sub-seasons. The early and late sub-seasons indicate one data point has high leverage. The final stepwise regressions and leverage tests for these sub-seasons are shown following the initial results.

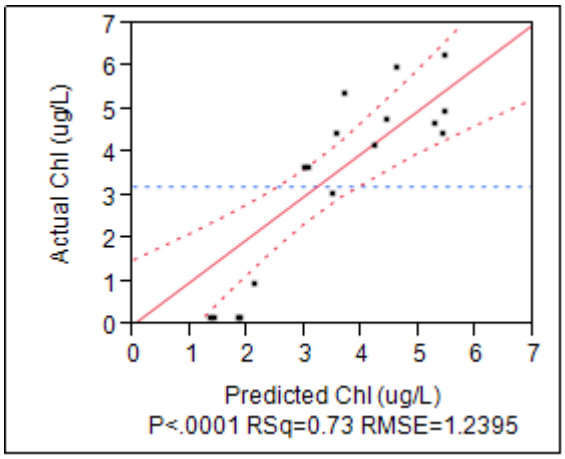

Figure B- 1 Measured Chlorophyll v. Estimated Chlorophyll using Early Model and all Points

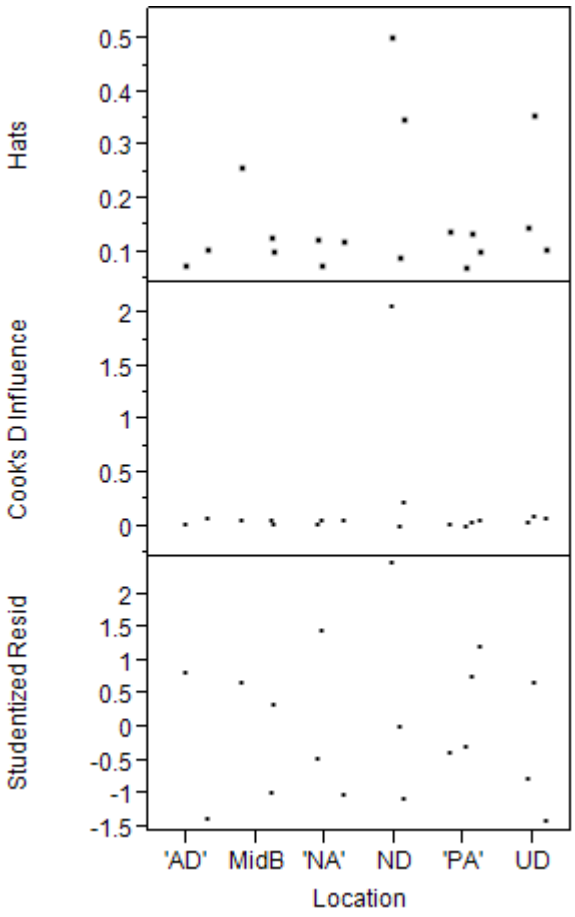

Figure B- 2 Tests of Leverage for Initial Early Model Indicate High Leverage for One Point

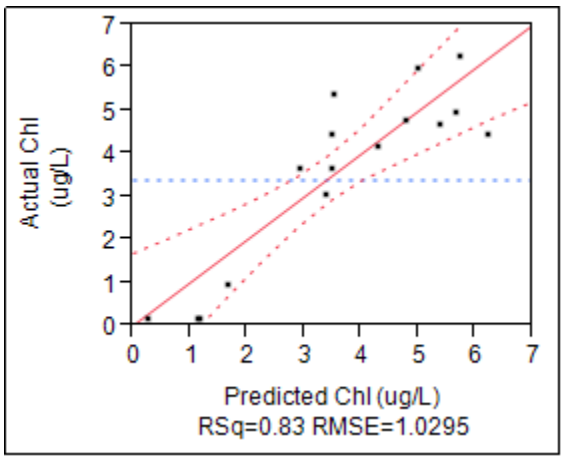

Figure B- 3 Measured Chlorophyll v. Estimated Chlorophyll after Exclusion of High Leverage Points using Early Model

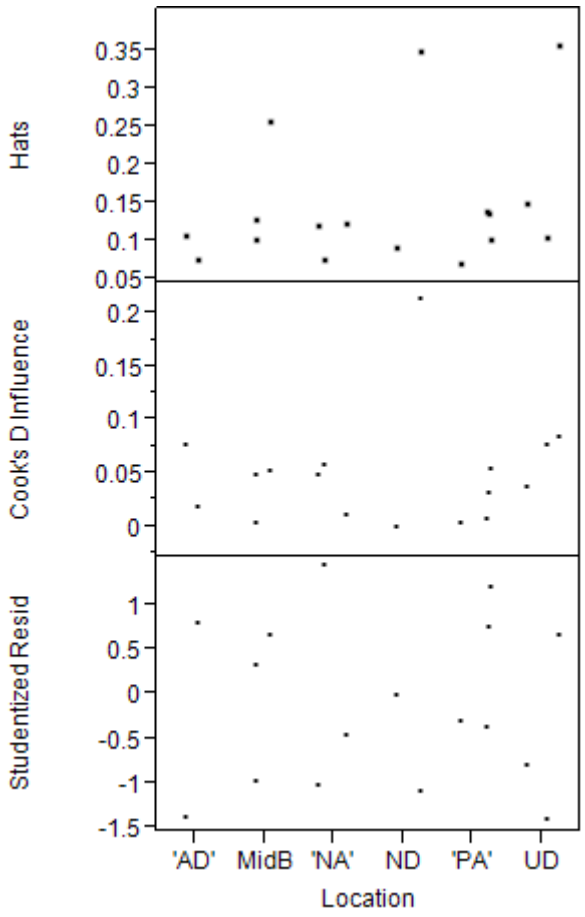

Figure B- 4 Tests of Leverage for Early Model after Exclusion of High Leverage Point

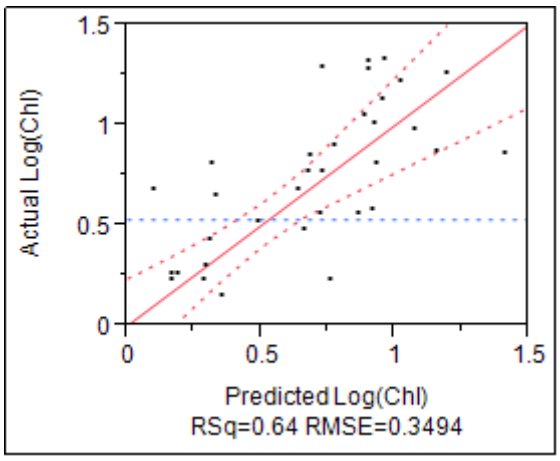

Figure B- 5 Log Measured Chlorophyll v. Log Estimated Chlorophyll using Mid-Summer Model

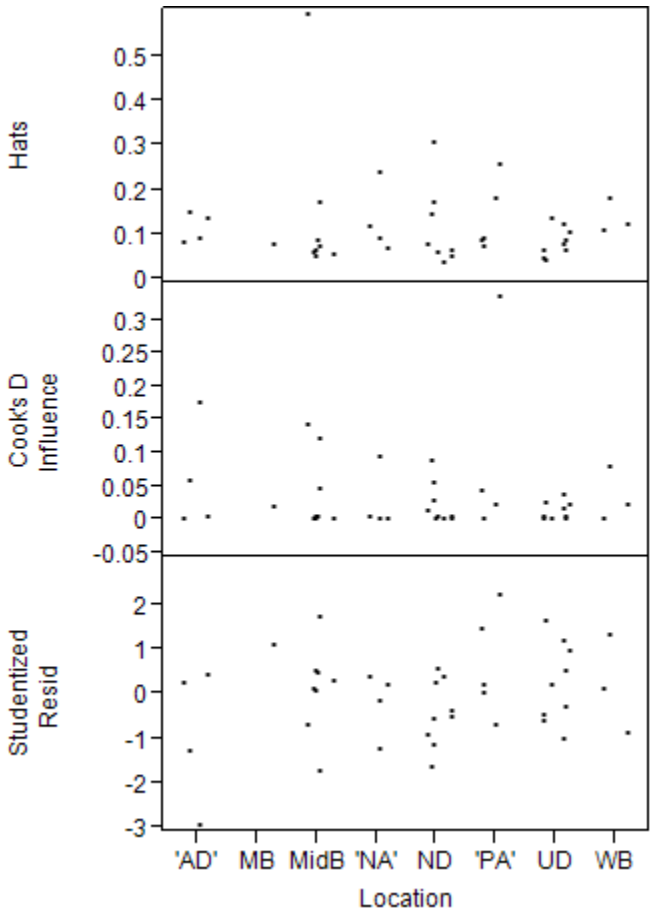

Figure B- 6 Tests of Leverage for Mid-Summer Model Indicating No Points with High Leverage

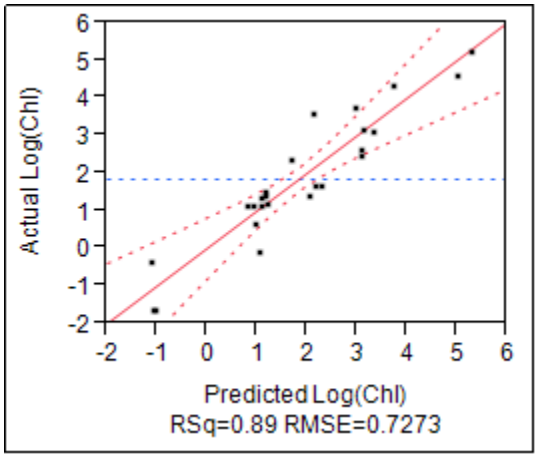

Figure B- 7 Log Measured Chlorophyll v. Log Estimated Chlorophyll using Initial Late Model and All Points

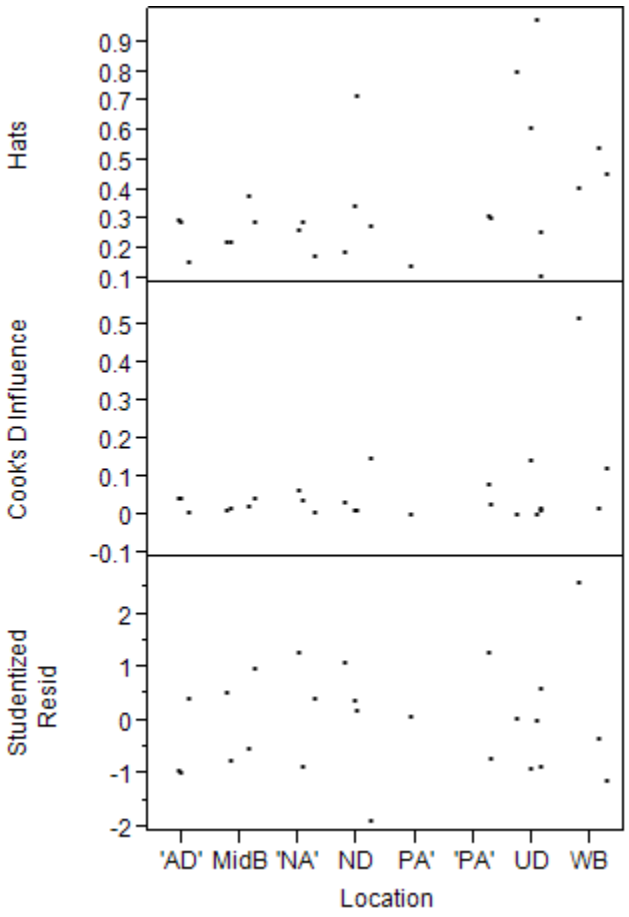

Figure B- 8 Tests of Leverage for Late Model Indicate High Leverage for One Point

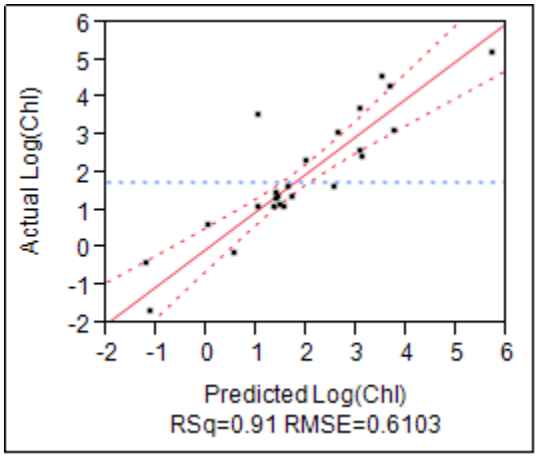

Figure B- 9 Log Measured Chlorophyll v Log Estimated Chlorophyll after Exclusion of High Leverage Point using Late Model

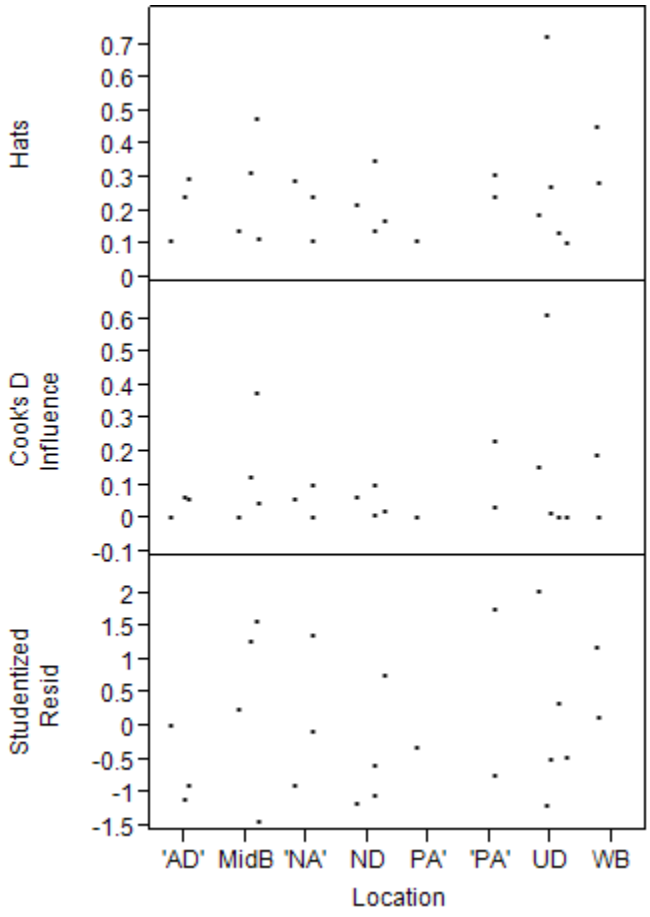

Figure B- 10 Tests of Leverage for Late Model after Exclusion of High Leverage Point

APPENDIX C. PARAMETERS USED IN STEPWISE REGRESSION

Table C-1 shows the final parameters chosen by the stepwise regression for each sub-season. The order listed indicates the order in which they were selected by the regression to be included in the model.

Table C- 1 Model Parameters Used By Season

Sub-season	Dependent Variable	Independent Variables
Early	Chlorophyll	B3/B1, B4/B7, B5/B4, B7/B5
Mid	Log(Chlorophyll)	B5, B7, B2/B4, B5/B3
Late	Log(Chlorophyll)	B1, B2/B5, B3/B5, B4/B7, B5/B4

APPENDIX D. HISTORICAL TRENDS BY SEASON AND RESERVOIR

This appendix contains the trends for the mode, maximum, and standard deviation of each reservoir by sub-season.

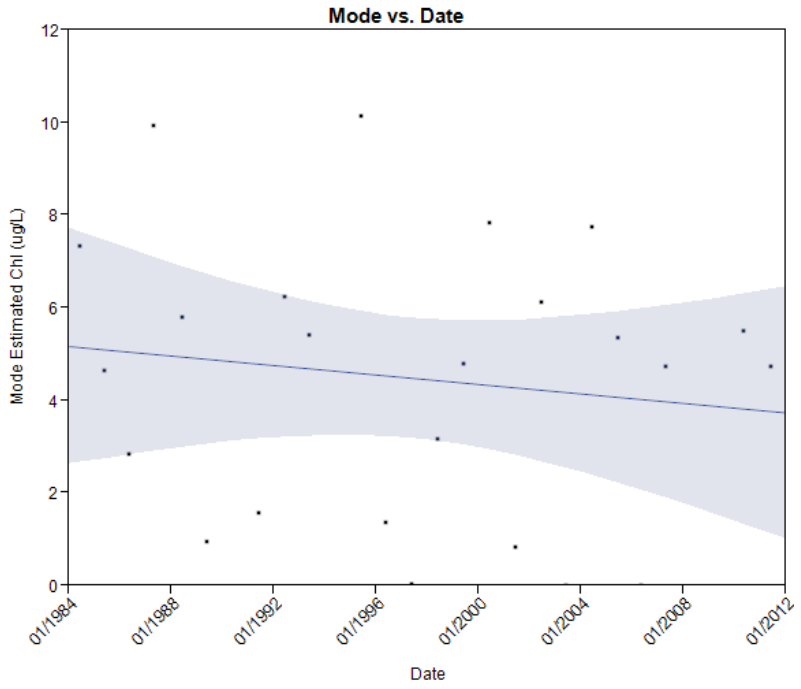

Figure D- 1 Average Estimated Chlorophyll from 1984-2012 for Early Summer Months in Deer Creek Reservoir

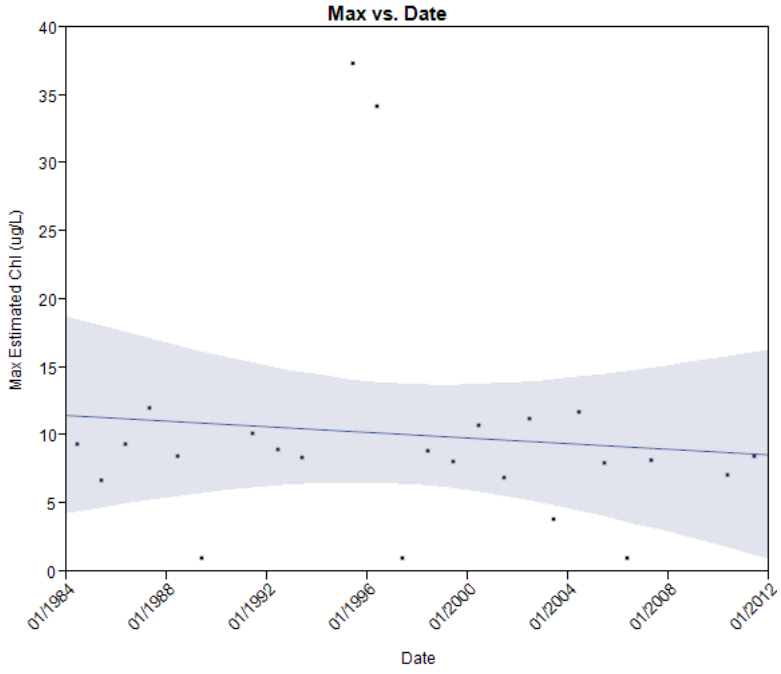

Figure D- 2 Max Estimated Chlorophyll from 1984-2012 for Early Summer Months in Deer Creek Reservoir

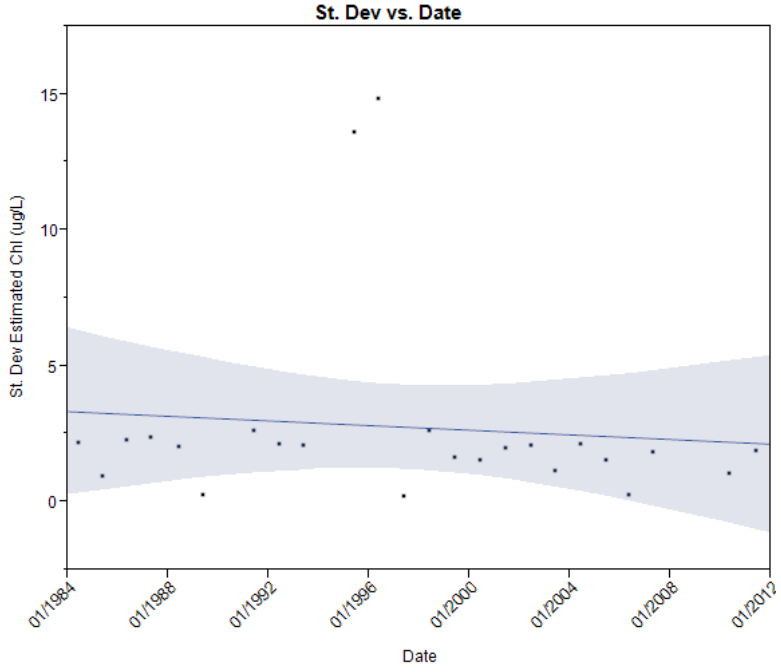

Figure D- 3 Standard Deviation of Estimated Chlorophyll from 1984-2012 for Early Summer Months in Deer Creek Reservoir

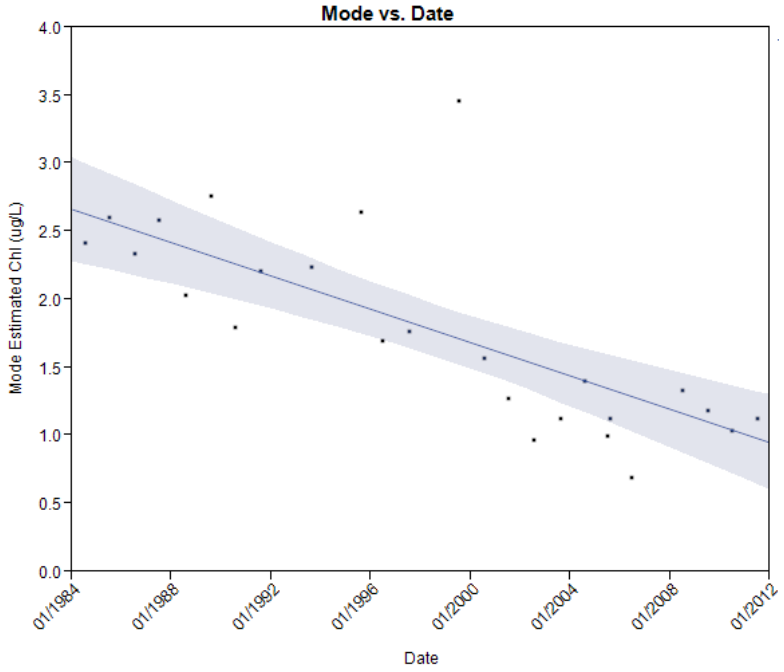

Figure D- 4 Average Estimated Chlorophyll from 1984-2012 for Mid-Summer Months in Deer Creek Reservoir

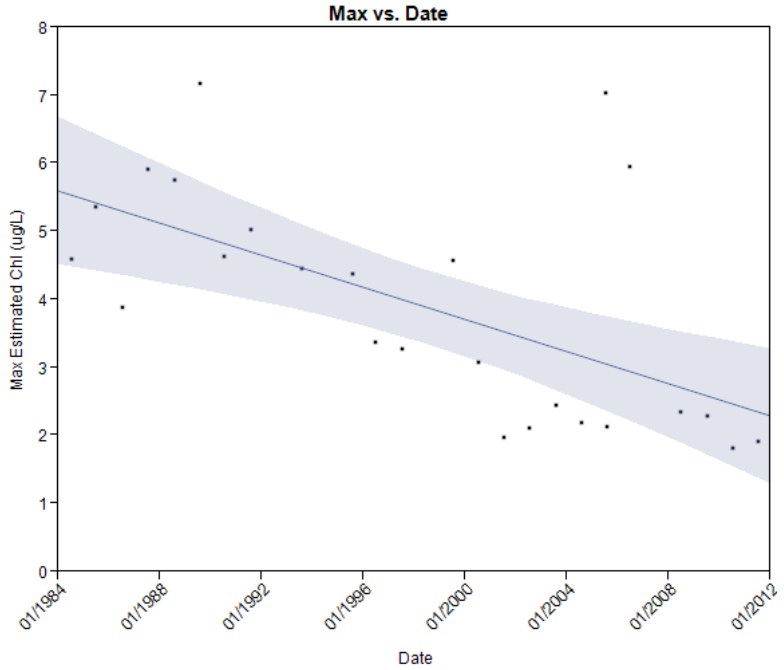

Figure D- 5 Max Estimated Chlorophyll from 1984-2012 for Mid-Summer Months in Deer Creek Reservoir

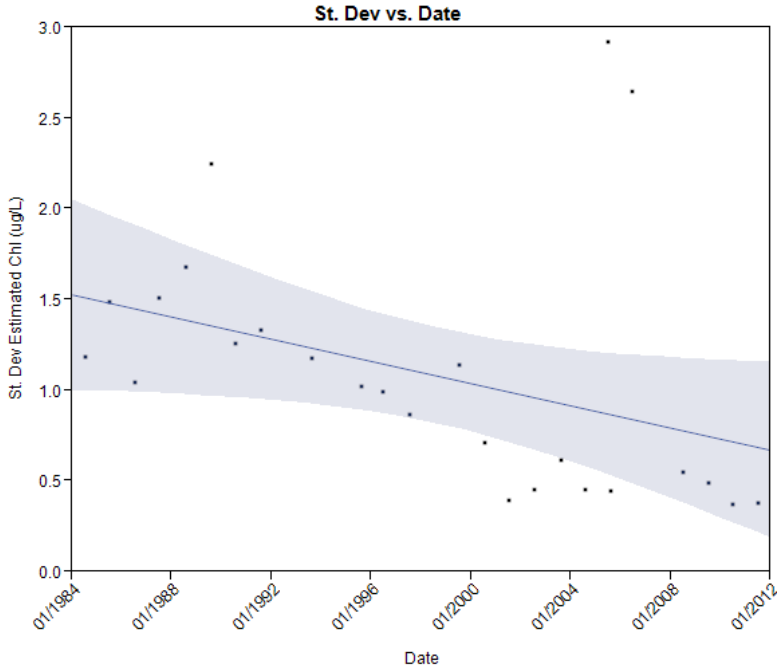

Figure D- 6 Standard Deviation of Estimated Chlorophyll from 1984-2012 for Mid-Summer Months in Deer Creek Reservoir

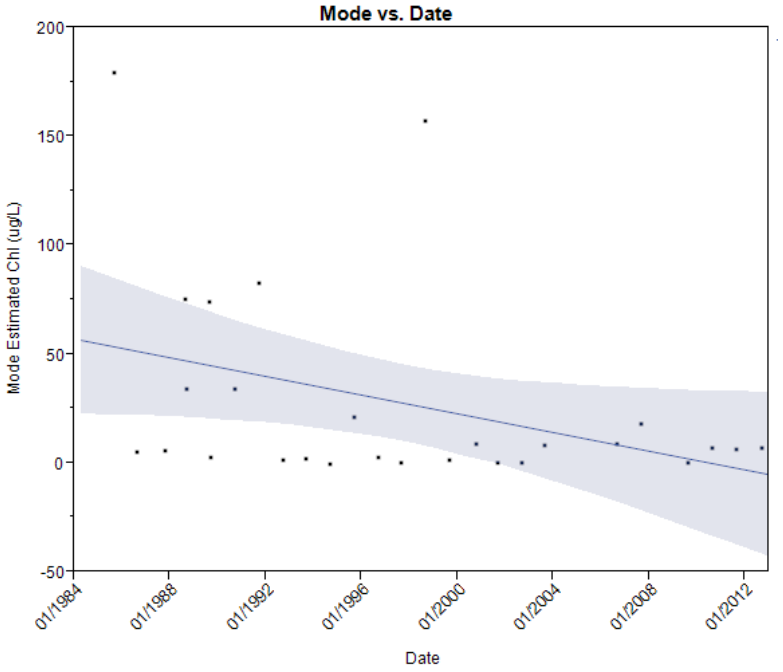

Figure D- 7 Average Estimated Chlorophyll from 1984-2012 for Late-Summer Months in Deer Creek Reservoir

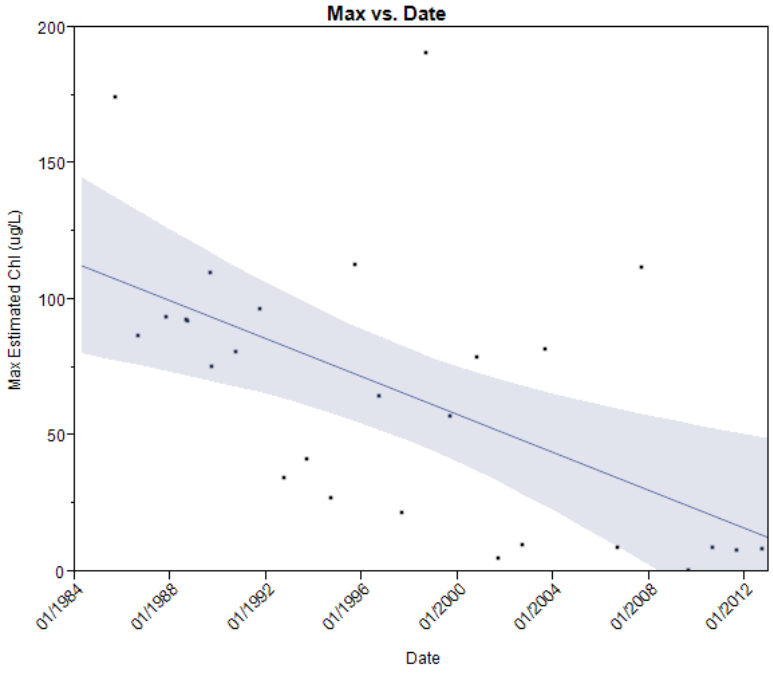

Figure D- 8 Max Estimated Chlorophyll from 1984-2012 for Late-Summer Months in Deer Creek Reservoir

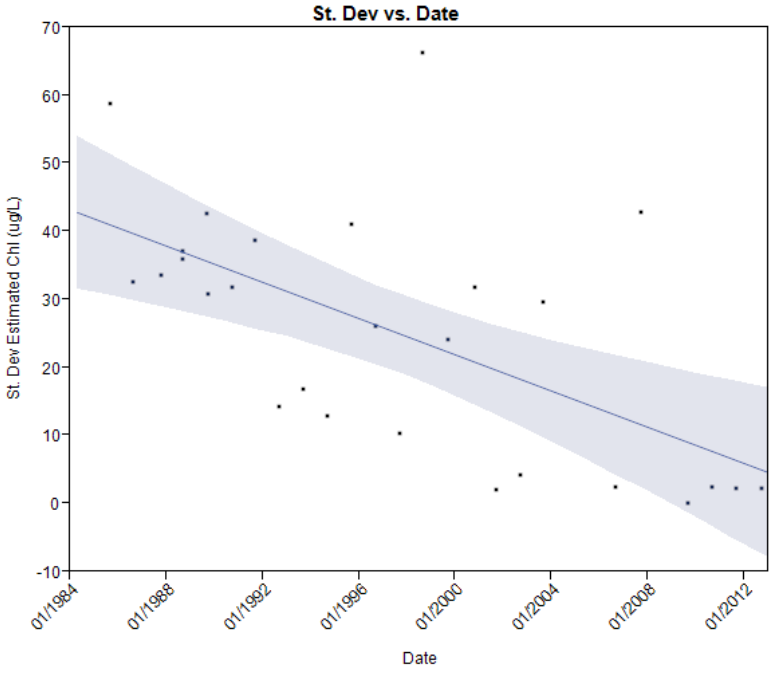

Figure D- 9 Standard Deviation of Estimated Chlorophyll from 1984-2012 for Late-Summer Months in Deer Creek Reservoir

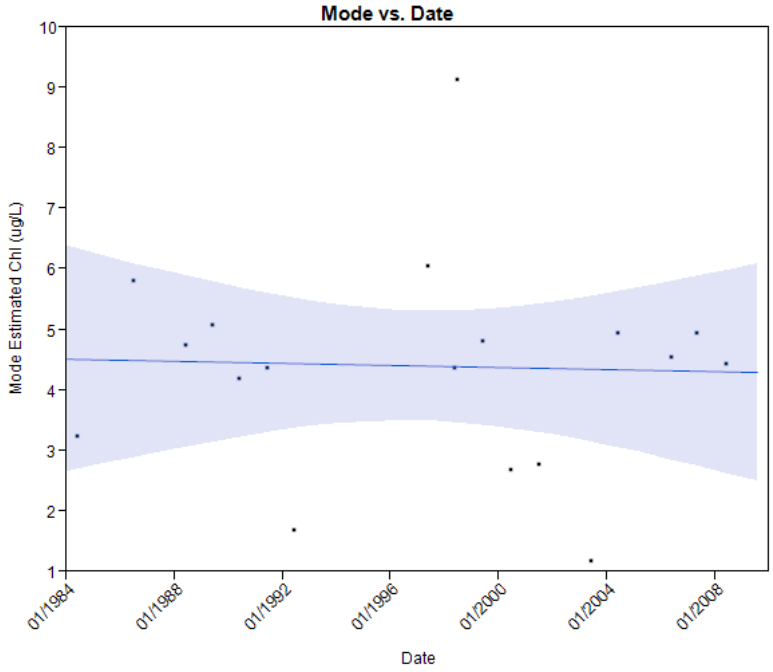

Figure D- 10 Average Chlorophyll from 1984-2012 for Early Summer Months in East Canyon Reservoir

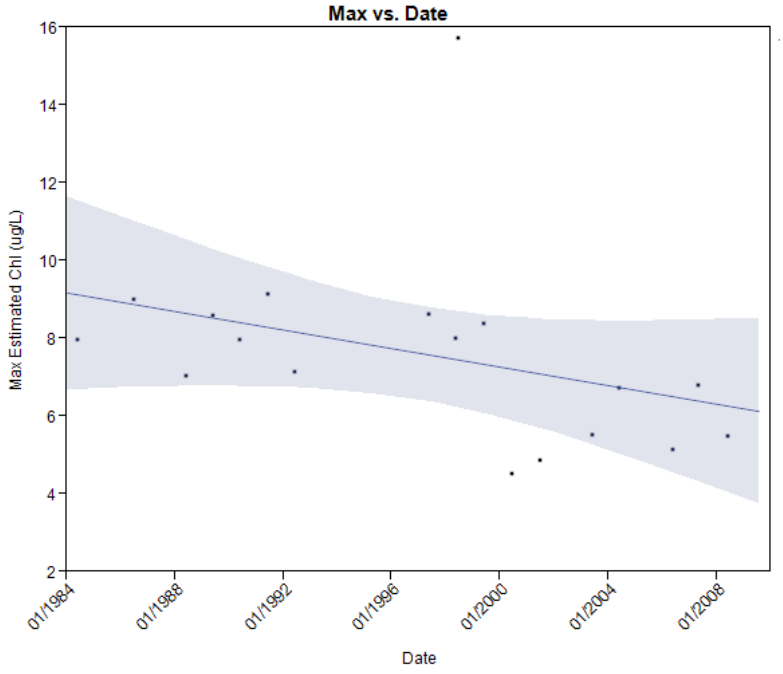

Figure D- 11 Max Estimated Chlorophyll from 1984-2012 for Early Summer Months in East Canyon Reservoir

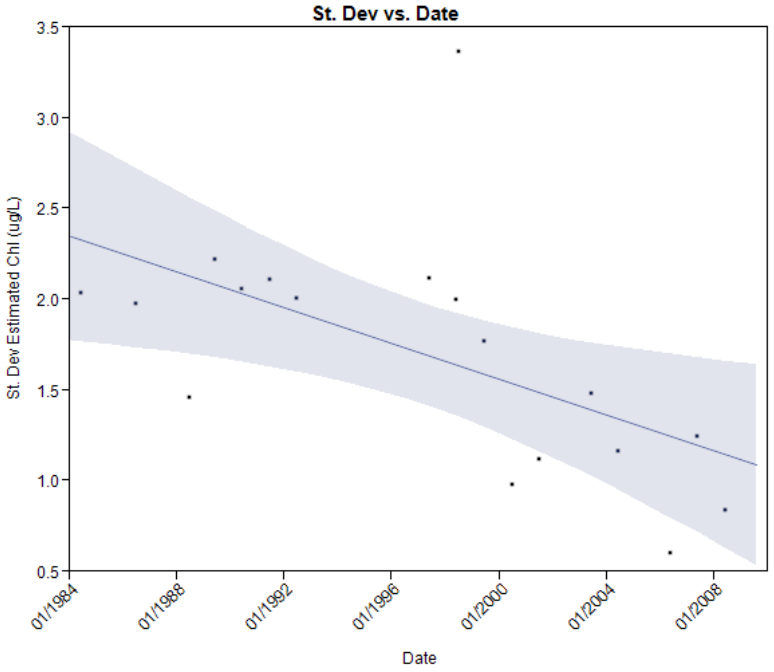

Figure D- 12 Standard Deviation of Estimated Chlorophyll from 1984-2012 for Early Summer Months in East Canyon Reservoir

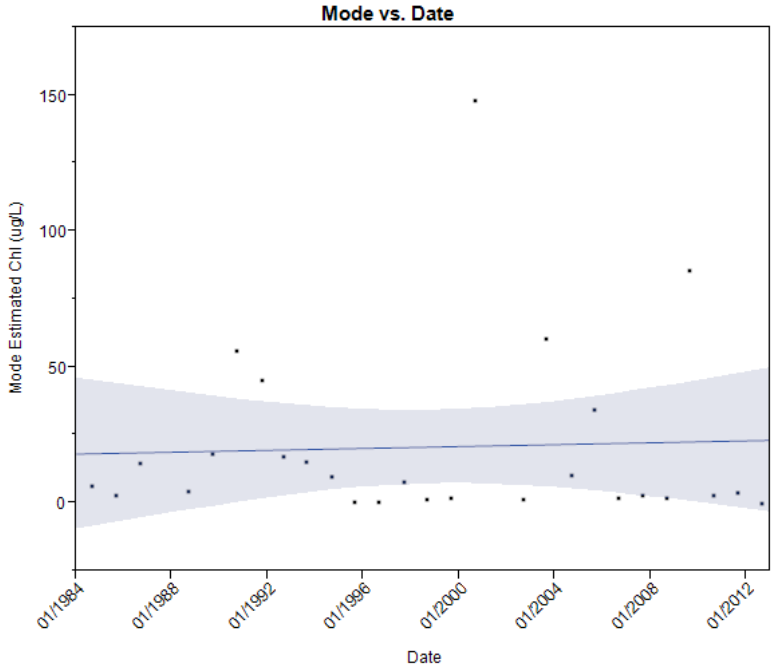

Figure D- 13 Average Estimated Max Estimated Chlorophyll from 1984-2012 for Mid-Summer Months in East Canyon Reservoir

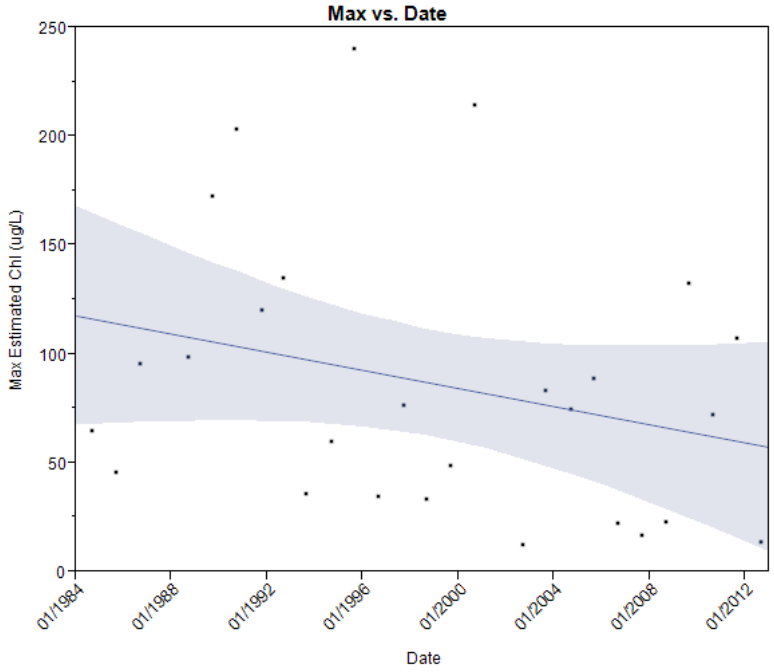

Figure D- 14 Max Estimated Chlorophyll from 1984-2012 for Mid-Summer Months in East Canyon Reservoir

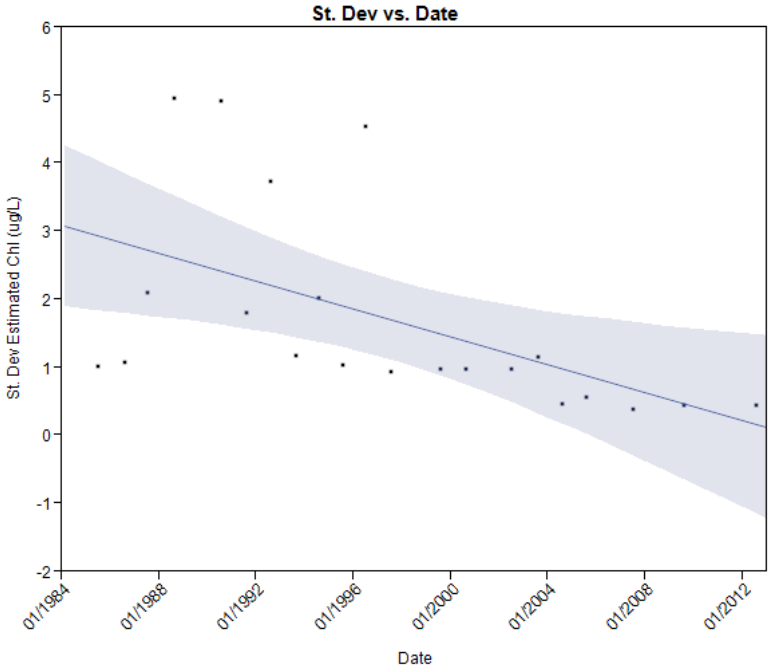

Figure D- 15 Standard Deviation of Estimated Chlorophyll from 1984-2012 for Mid-Summer Months in East Canyon Reservoir

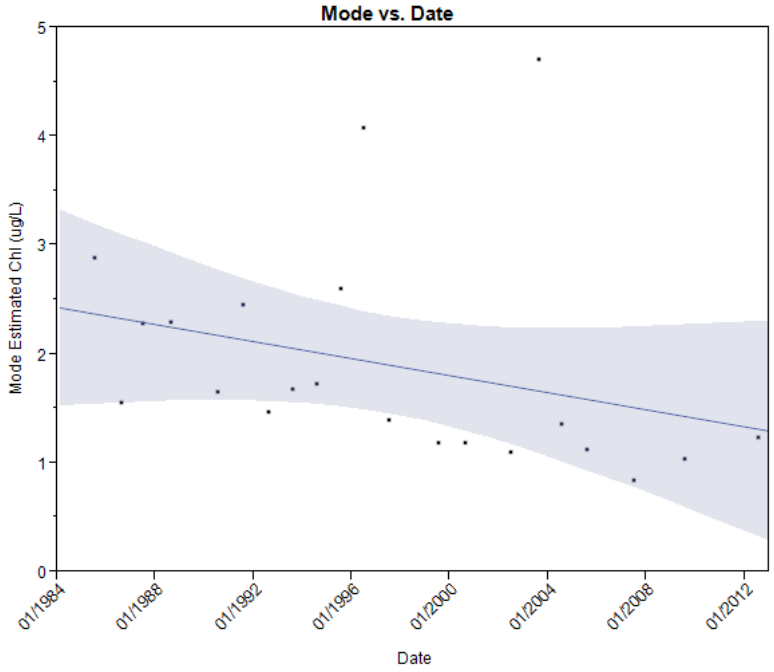

Figure D- 16 Average Estimated Chlorophyll from 1984-2012 for Late-Summer Months in East Canyon Reservoir

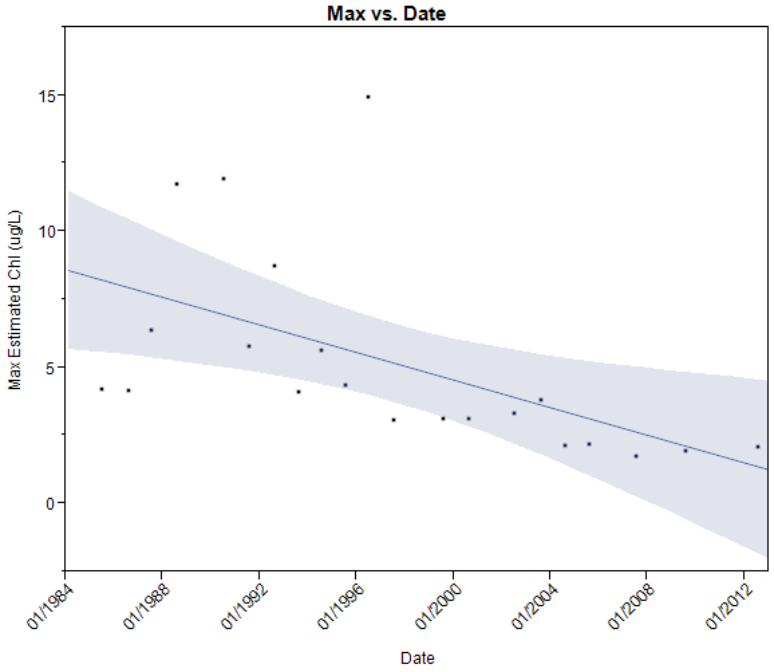

Figure D- 17 Max Estimated Chlorophyll from 1984-2012 for Late-Summer Months in East Canyon Reservoir

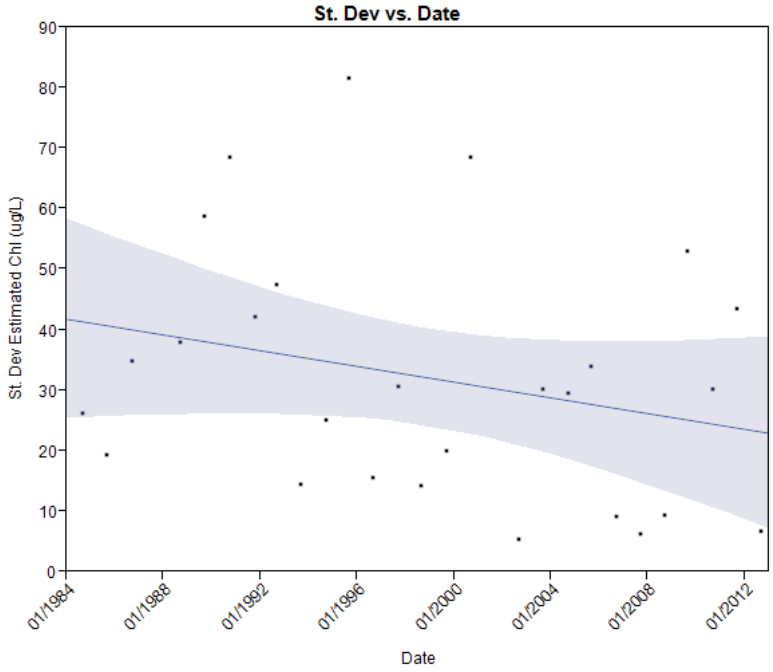

Figure D- 18 Standard Deviation of Estimated Chlorophyll from 1984-2012 for Late-Summer Months in East Canyon Reservoir

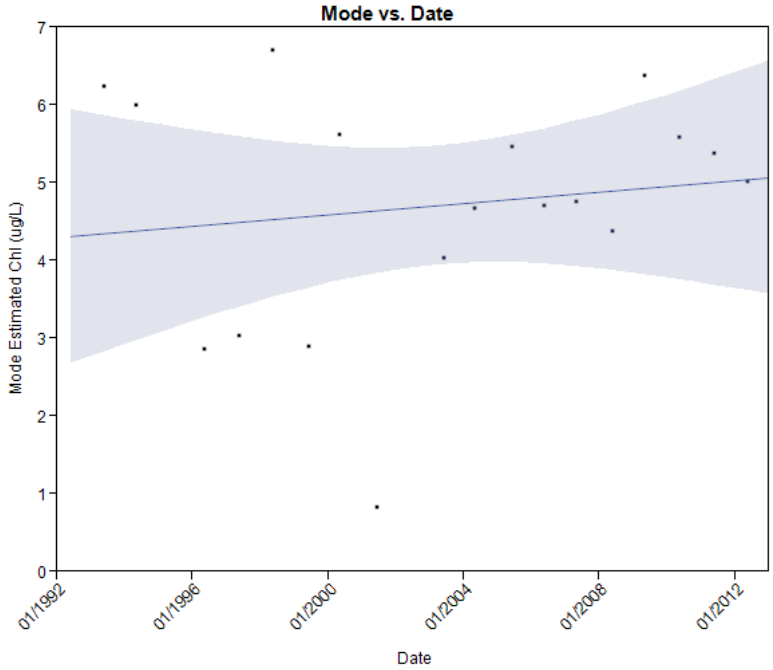

Figure D- 19 Average Estimated Chlorophyll from 1993-2012 for Early Summer Months in Jordanelle Reservoir

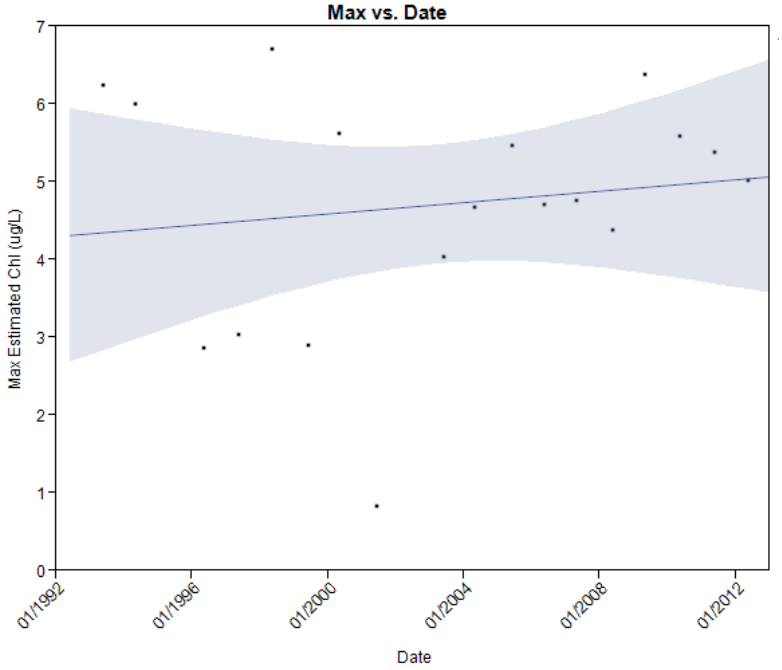

Figure D- 20 Max Estimated Chlorophyll from 1993-2012 for Early Summer Months in Jordanelle Reservoir

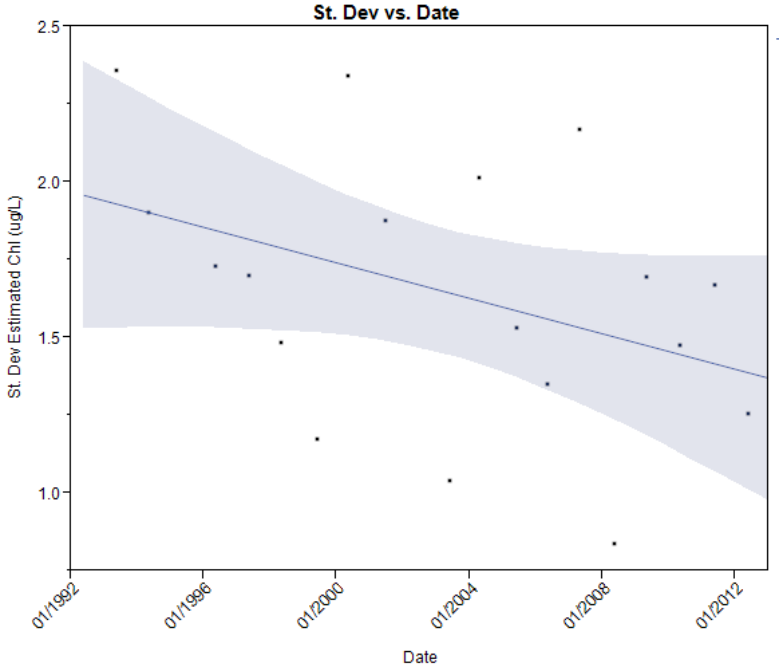

Figure D- 21 Standard Deviation Estimated Chlorophyll from 1993-2012 for Early Summer Months in Jordanelle Reservoir

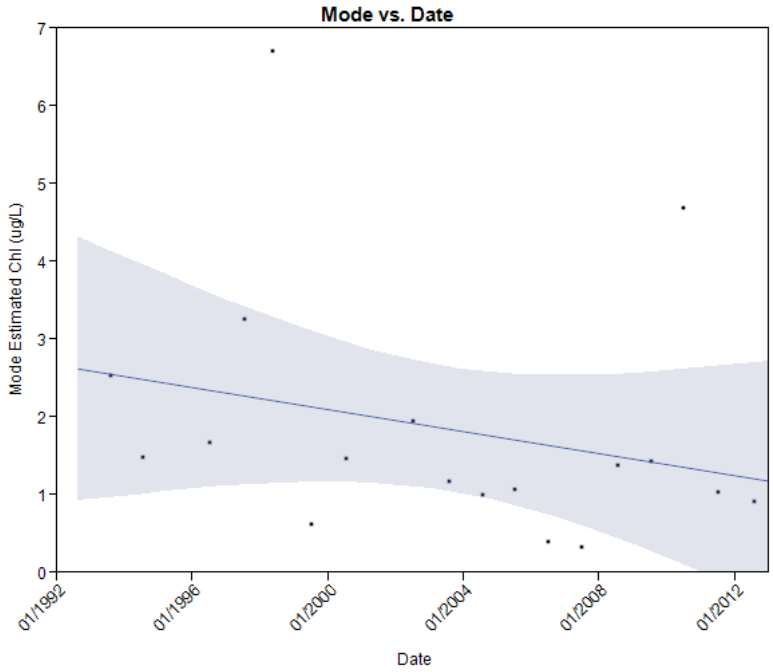

Figure D- 22 Average Estimated Chlorophyll from 1993-2012 for Mid-Summer Months in Jordanelle Reservoir

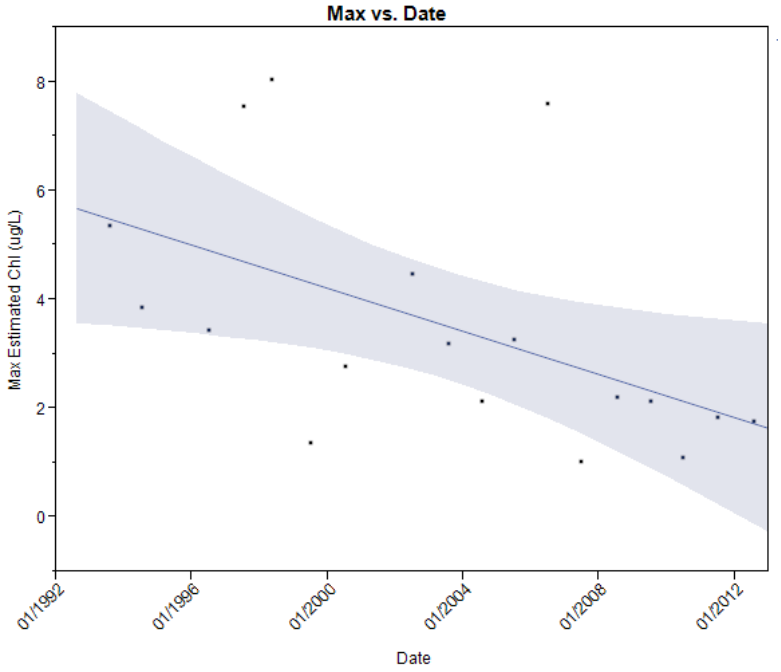

Figure D- 23 Max Estimated Chlorophyll from 1993-2012 for Mid-Summer Months in Jordanelle

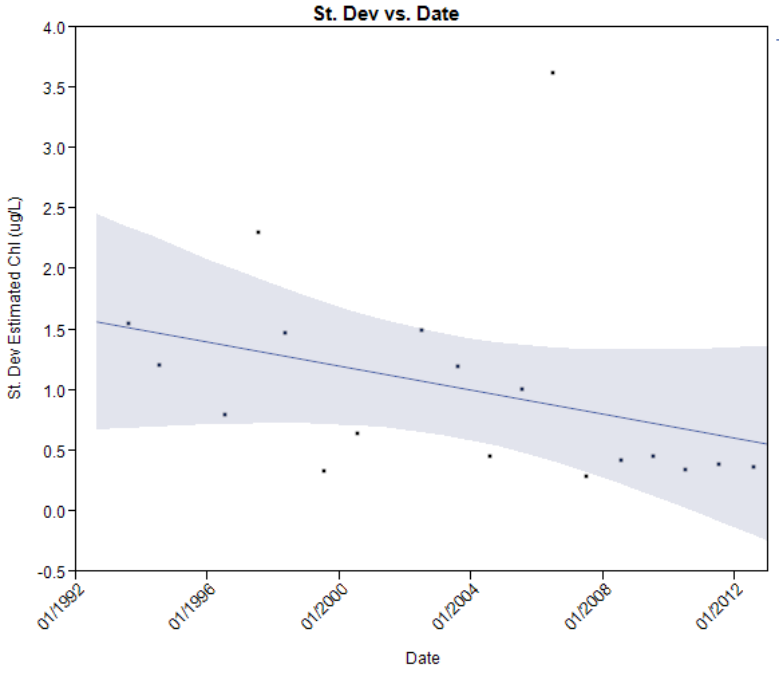

Figure D- 24 Standard Deviation Estimated Chlorophyll from 1993-2012 for Mid-Summer Months in Jordanelle Reservoir

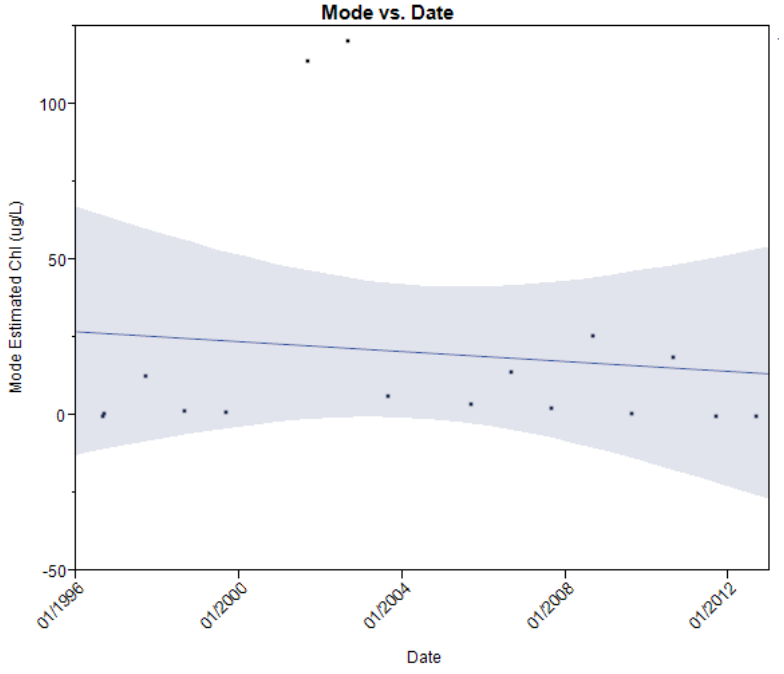

Figure D- 25 Average Estimated Chlorophyll from 1993-2012 for Late Summer Months in Jordanelle Reservoir

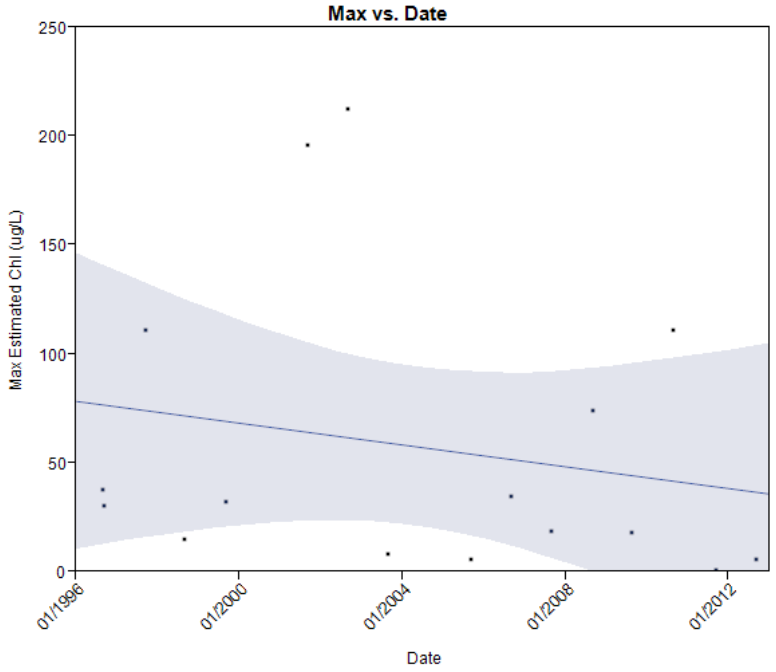

Figure D- 26 Max Estimated Chlorophyll from 1993-2012 for Late Summer Months in Jordanelle Reservoir

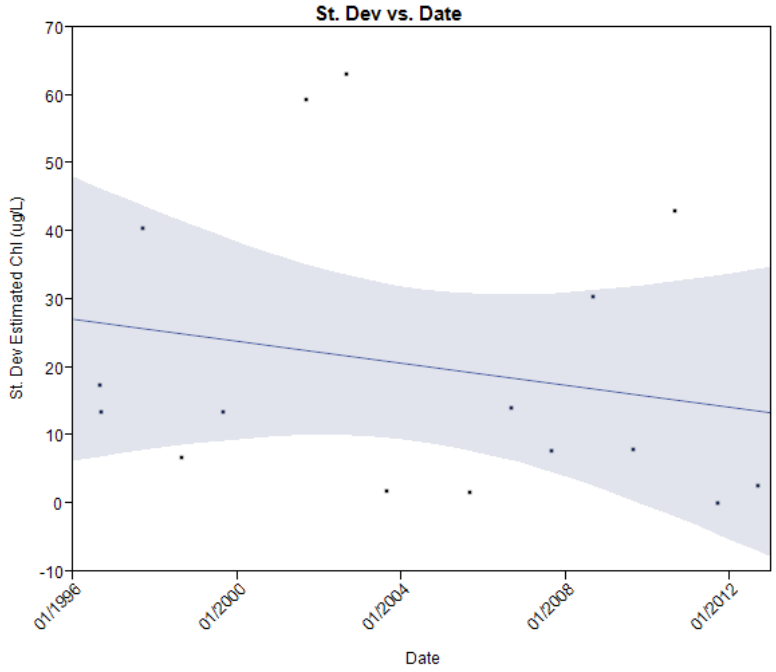

Figure D- 27 Standard Deviation Estimated Chlorophyll from 1993-2012 for Late Summer Months in Jordanelle Reservoir

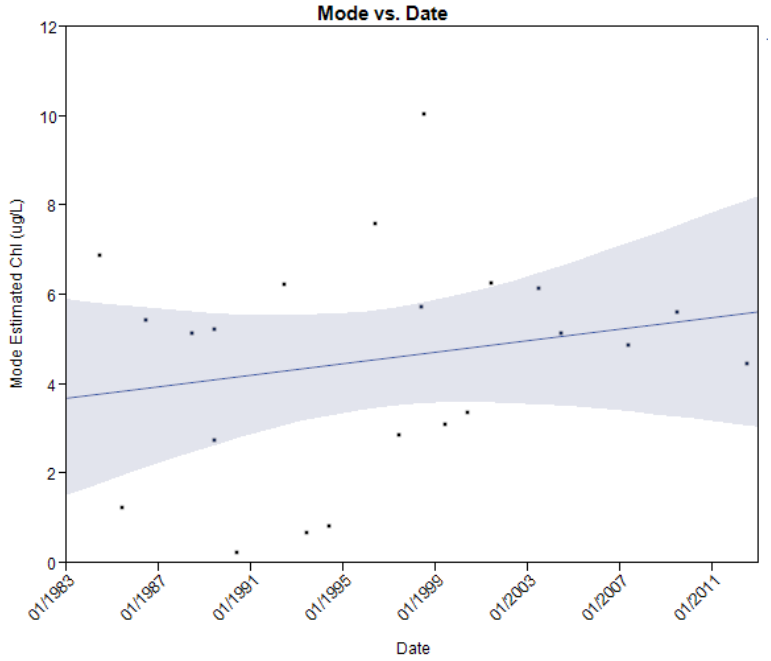

Figure D- 28 Average Estimated Chlorophyll from 1984-2012 for Early Months in Echo Reservoir

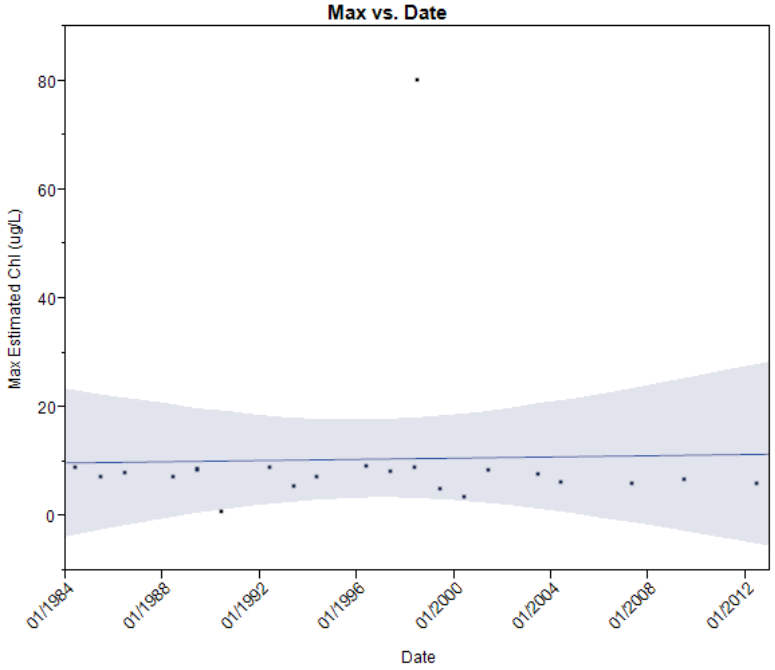

Figure D- 29 Max Estimated Chlorophyll from 1984-2012 for Early Months in Echo Reservoir

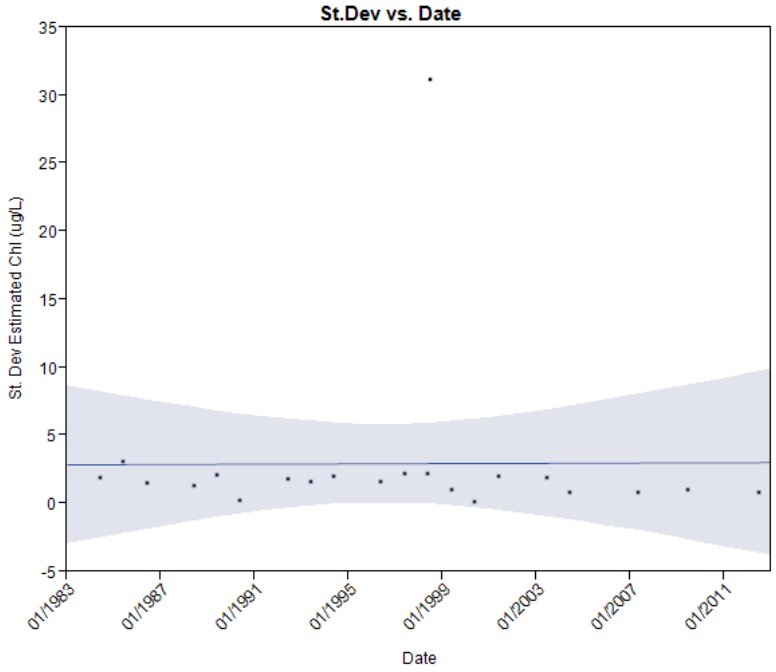

Figure D- 30 Standard Deviation of Estimated Chlorophyll from 1984-2012 for Early Months in Echo Reservoir

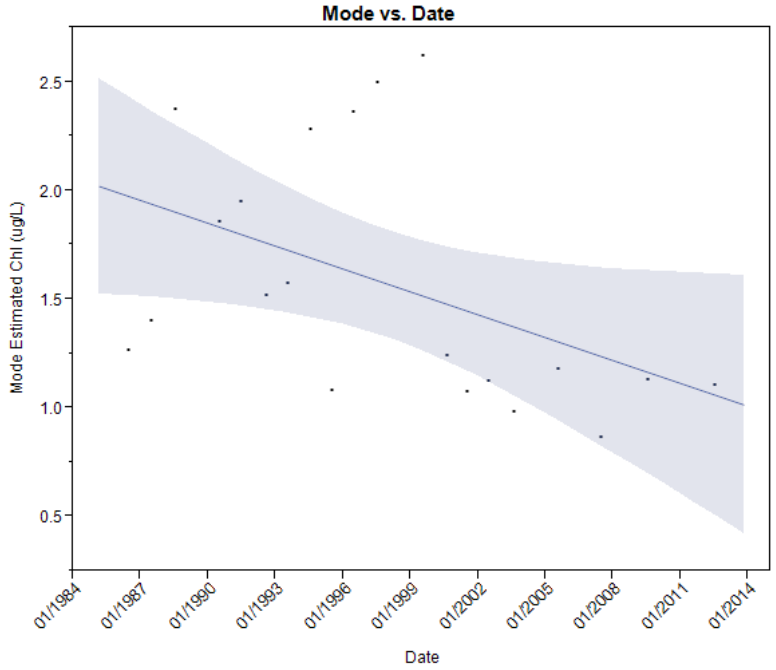

Figure D- 31 Average Estimated Chlorophyll from 1984-2012 for Mid-Summer Months in Echo Reservoir

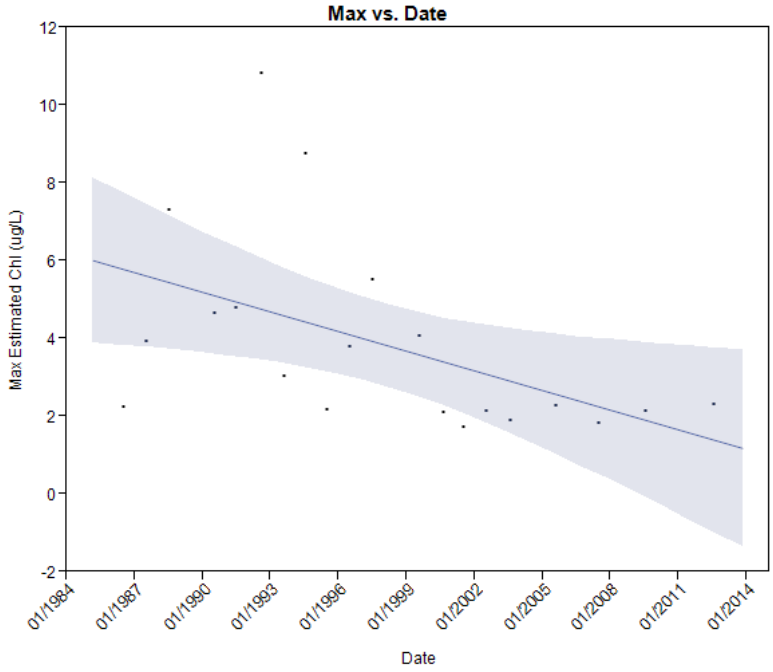

Figure D- 32 Max Estimated Chlorophyll from 1984-2012 for Mid-Summer Months in Echo Reservoir

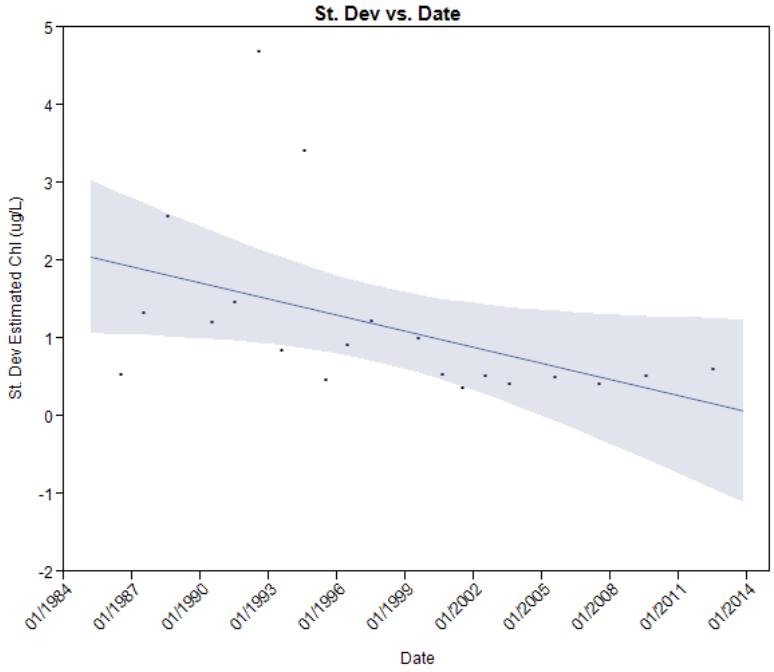

Figure D- 33 Standard Deviation of Estimated Chlorophyll from 1984-2012 for Mid-Summer Months in Echo Reservoir

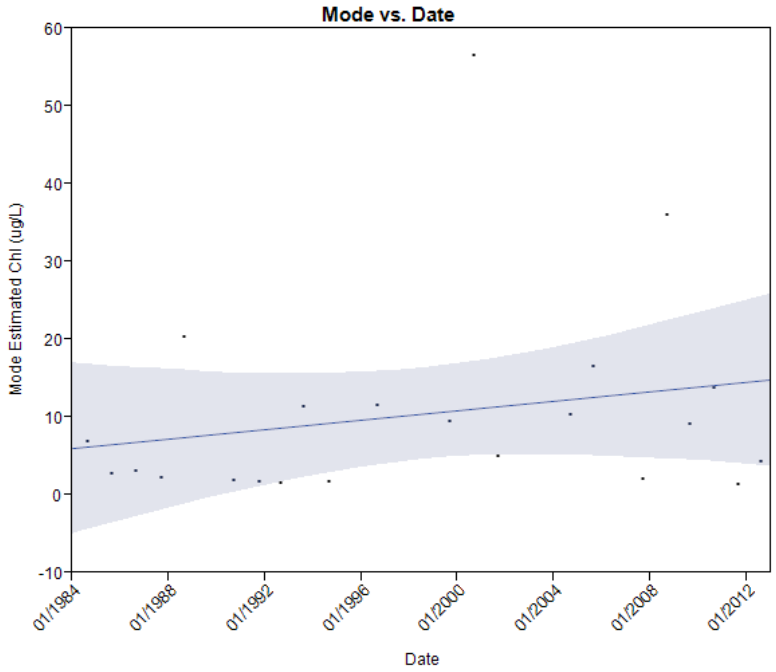

Figure D- 34 Average Estimated Chlorophyll from 1984-2012 for Late Summer Months in Echo Reservoir

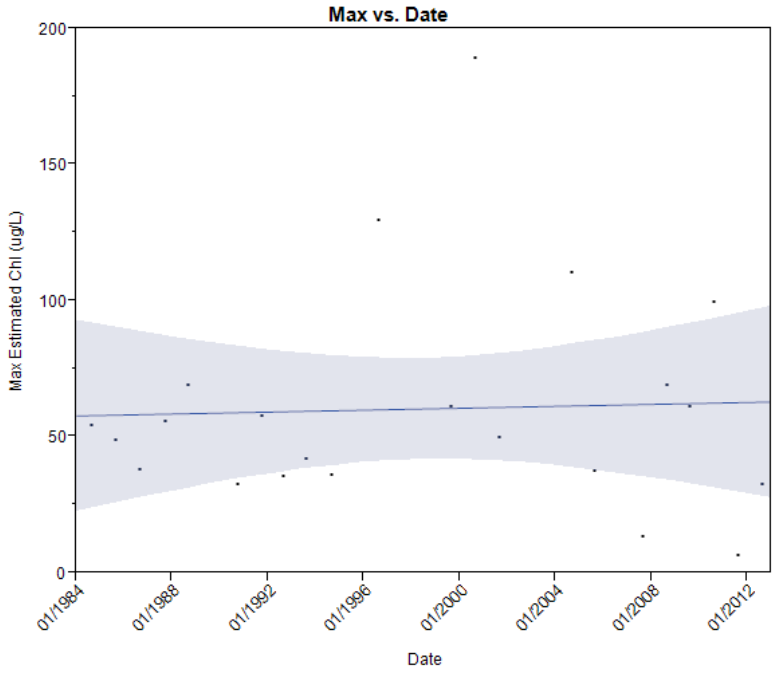

Figure D- 35 Max Estimated Chlorophyll from 1984-2012 for Late Summer Months in Echo Reservoir

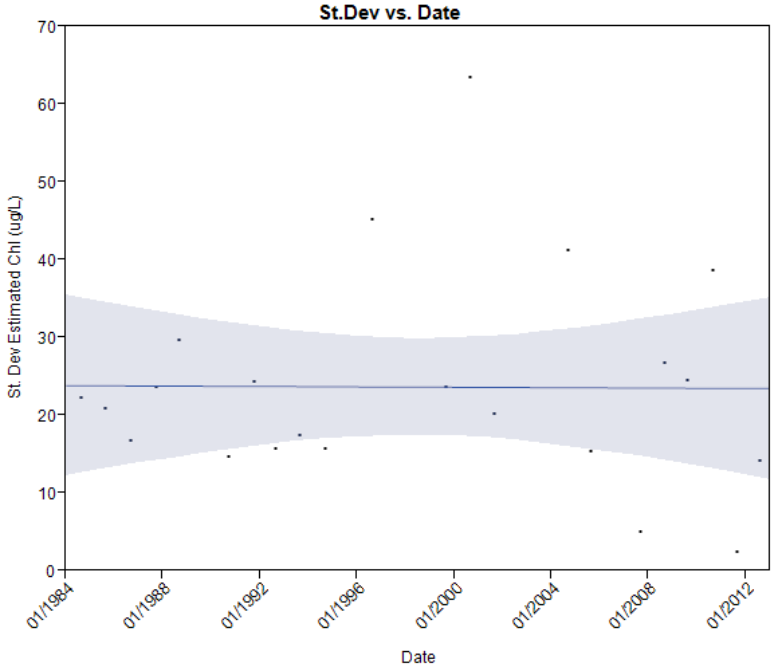

Figure D- 36 Standard Deviation of Estimated Chlorophyll from 1984-2012 for Late Summer Months in Echo Reservoir

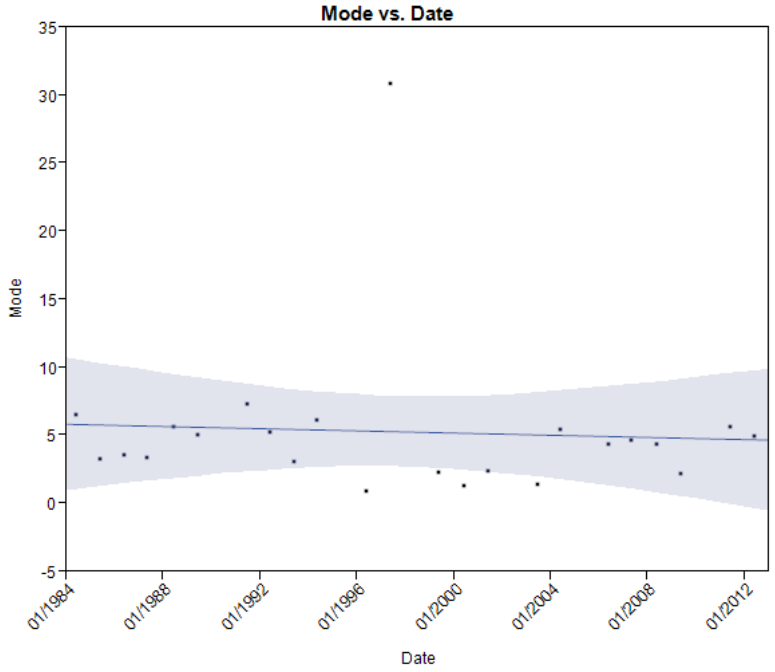

Figure D- 37 Average Estimated Chlorophyll from 1984-2012 for Early Summer Months in Rockport Reservoir

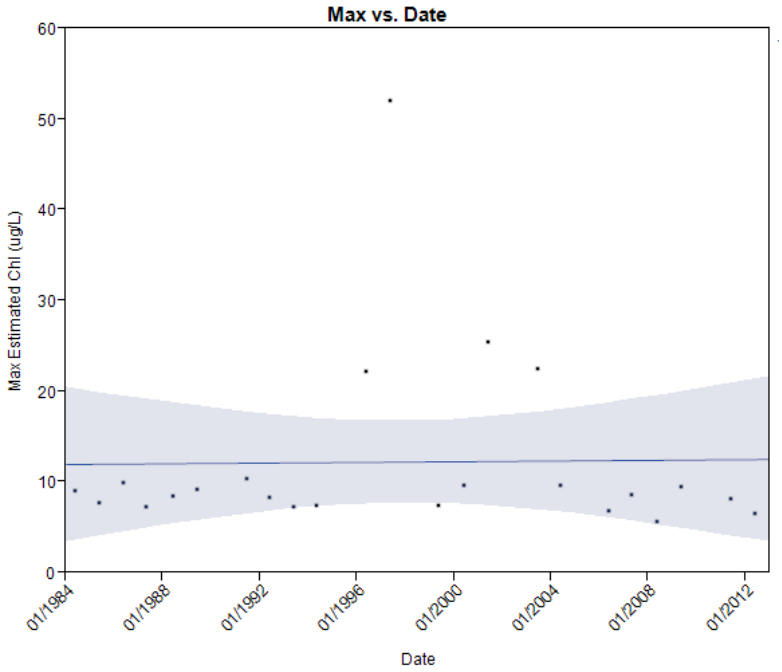

Figure D- 38 Max Estimated Chlorophyll from 1984-2012 for Early Summer Months in Rockport Reservoir

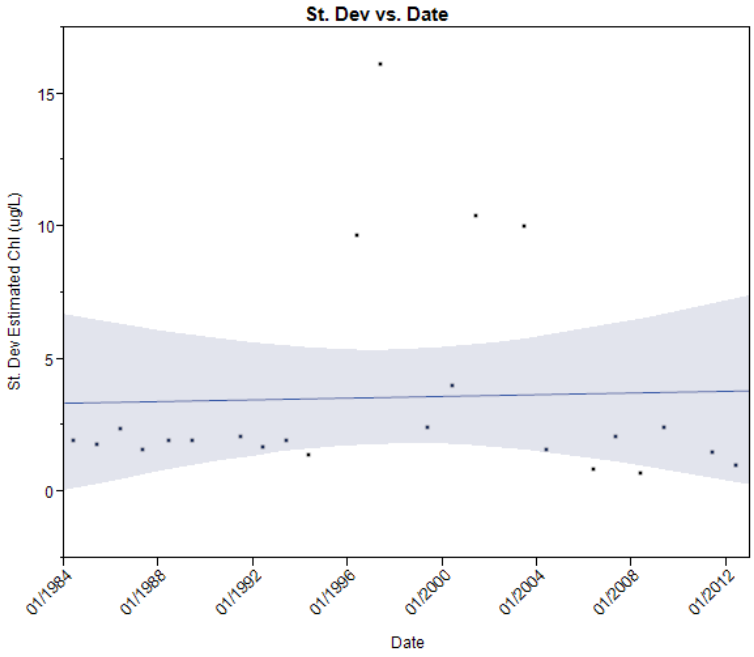

Figure D- 39 Standard Deviation of Estimated Chlorophyll from 1984-2012 for Early Summer Months in Rockport Reservoir

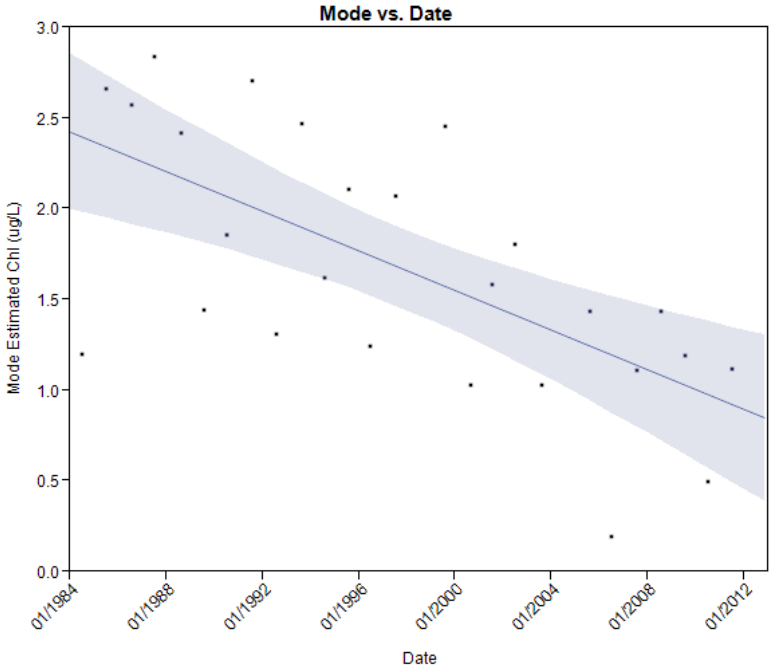

Figure D- 40 Average Estimated Chlorophyll from 1984-2012 for Mid-Summer Months in Rockport Reservoir

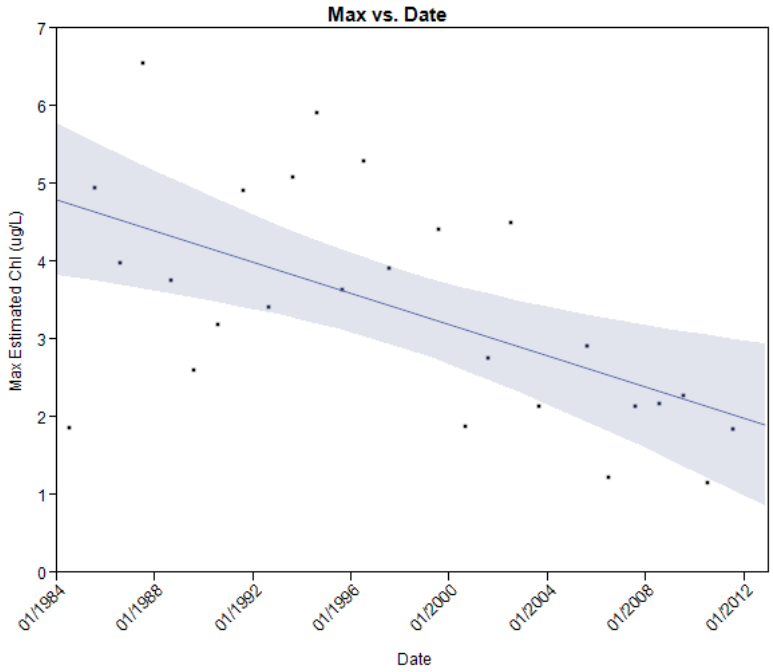

Figure D- 41 Average Estimated Chlorophyll from 1984-2012 for Mid-Summer Months in Rockport Reservoir

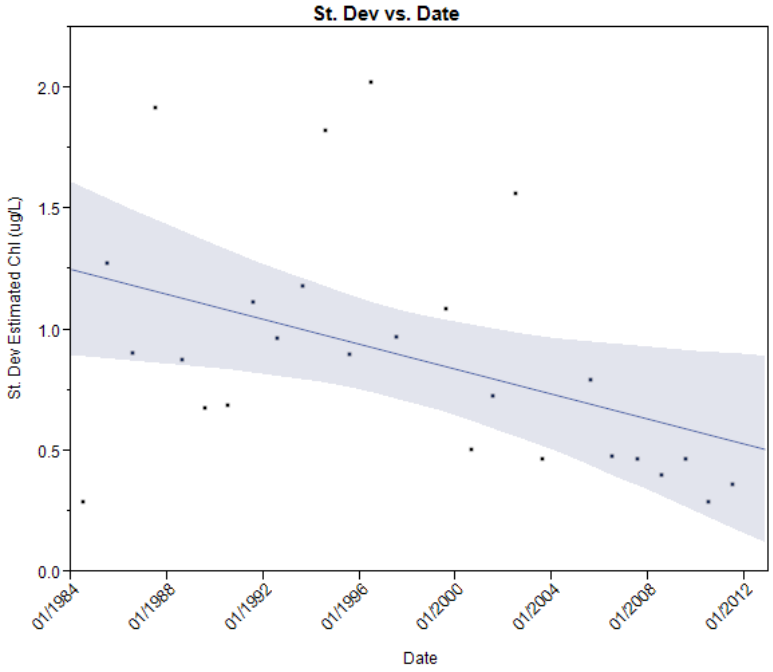

Figure D- 42 Estimated Chlorophyll from 1984-2012 for Mid-Summer Months in Rockport Reservoir

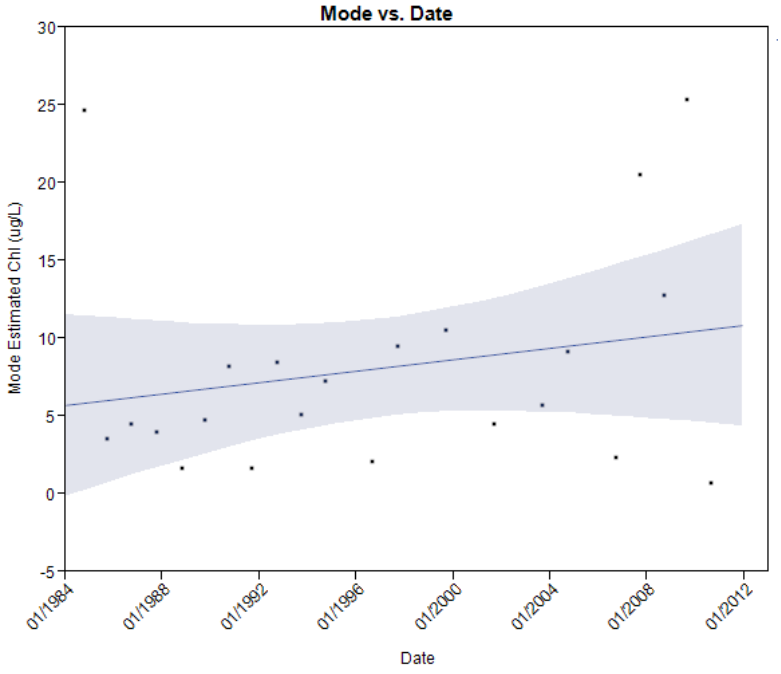

Figure D- 43 Average Estimated Chlorophyll from 1984-2012 for Late Summer Months in Rockport Reservoir

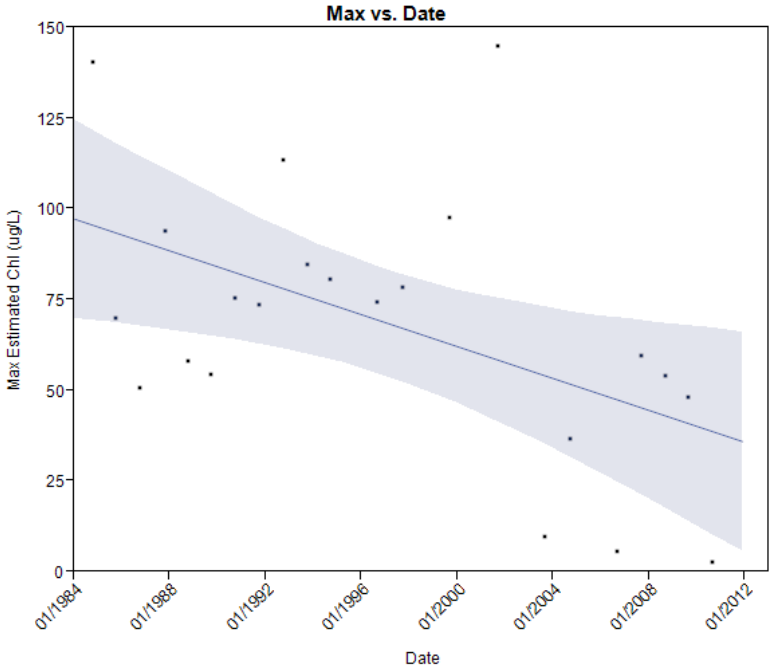

Figure D- 44 Max Estimated Chlorophyll from 1984-2012 for Late Summer Months in Rockport Reservoir

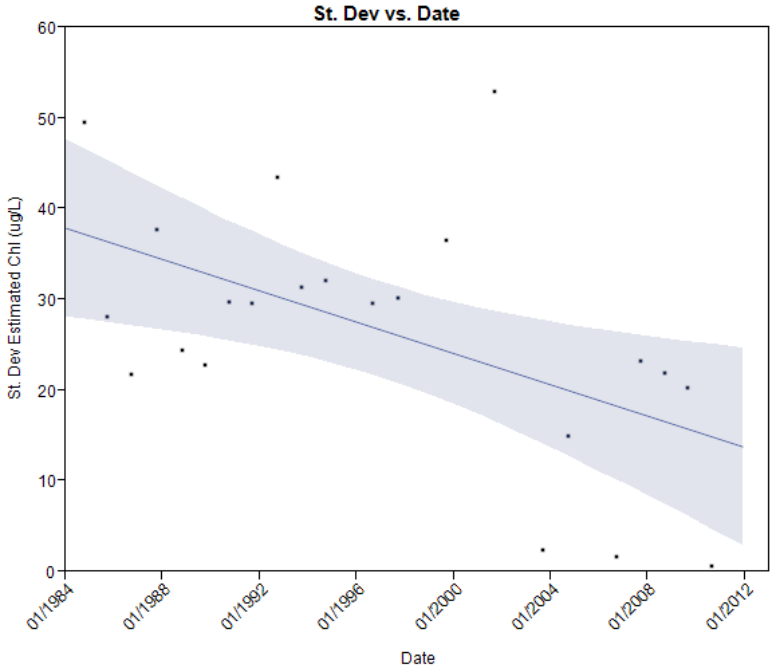

Figure D- 45 Standard Deviation of Estimated Chlorophyll from 1984-2012 for Late Summer Months in Rockport Reservoir