Note on Chlorophyll Analysis in Chassahowitzka Water Quality Appendix

Section 3.3 of this report describes an analysis of the effects of flow reductions on chlorophyll concentration in which measurements of chlorophyll concentration are compared to a threshold value and calculates risk of individual samples exceeding that value. To perform this analysis a threshold must be selected, and that threshold should be relevant to the system being studied. District staff identified a value of 3.9 µg/L as the most relevant threshold to use. Note, it is critical to distinguish between our use of this value as a threshold for analysis and its prescribed use as a criterion for determining impairment within the Estuary-Specific Numeric Interpretations of the Narrative Nutrient Criterion (NNC). This value, taken as the NNC, is applicable only within WBID 1361 and as an annual geometric mean value according to Rule 62-302.532, F.A.C. Contributing to our decision to use this 3.9 µg/L value, associated with the downstream 1361 WBID, is that the upstream WBID (1348D) does not have a chlorophyll NNC value. We used this same value of 3.9 µg/L, but for a different purpose than determination of impairment of the NNC. Thus, an instance of a single exceedance of this threshold, or an increased risk of this exceedance across several repeated samples both inside and outside the WBID boundary cannot and should not be interpreted in the context of impairment of the NNC.

Gabe Herrick, Ph.D.
Senior Environmental Scientist
Southwest Florida Water Management District
2379 Broad Street
Brooksville, FL 34604

EXPLORATORY EVALUATION OF WATER QUALITY AND FLOW RELATIONSHIPS FOR THE CHASSAHOWITZKA RIVER IN SUPPORT OF MINIMUM FLOWS REEVALUATION

TECHNICAL REPORT IN FULLFILLMENT OF WORK ASSIGNMENT 18TW0001116:

PREPARED FOR:

THE SOUTHWEST FLORIDA WATER MANAGEMENT DISTRICT, BROOKSVILLE, FL

PREPARED BY:

JANICKI ENVIRONMENTAL, INC. AND WSP, INC.

FINAL REPORT

December 7, 2018

Table of Contents

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

This report details efforts to quantify relationships between spring flows from the Chassahowitzka River head springs and water quality throughout the river system. Water quality is one of 10 “environmental values” defined in the State Water Resource Implementation Rule to be considered when establishing minimum flows. Salinity is a water quality constituent that represents a direct, physical driver for many estuarine processes. However, other water quality constituents can also affect biological resources in the river. This work effort focused on providing an exploratory examination of the relationships between flows and water quality constituents using the most up to date datasets available for the Chassahowitzka River System. The analysis focused on identifying water quality response endpoints that, under certain conditions, could result in adverse effects to a “resource of concern” within the river as a function of reduced flows. Resources of concern are those attributes of the system that relate to one or more of the 10 environmental values identified in the Water Resource Implementation Rule and have potential quantifiable responses to flow. This updated analysis on relationships between spring flows and water quality uses an expanded list of water quality constituents and additional data collected since the District’s original minimum flow report (Heyl et al. 2012) was prepared.

For this analysis, spatial attributes of the river including the headsprings (“Springs”), the “Mainstem” of the river, and the nearshore “Estuary” (outside the river mouth) were identified as potential resources of concern. The specific tasks associated with this work effort consisted of data gathering, exploratory data analysis, stochastic predictive modeling and synthesizing information to supplement existing knowledge on the effects of flows on water quality in the system. Initial tasks included the compilation of available water quality and water quantity data for the Chassahowitzka River and the creation of a Microsoft Access database and database inventory. Additionally, descriptive statistics and plots were generated for each metric of interest to describe both the univariate characteristics and the seasonal and inter-annual distributions. Screening methods were used to identify and qualify potential anomalous data evident in the datasets in the Access database and linear regression was used to explore bivariate relationships between the water quality constituent of interest and flow. Subsequent to initial data compilation and exploration, a statistical analysis plan was developed which outlined potential analytical methods used to approach each of the various data types that exist in the master database. Application of the statistical analysis plan led to the analytical results describing the effects of flows on water quality within the system. Previously developed acceptance criteria for using linear regression relationships in support of minimum flows were applied prior to reporting significant results for linear regression analysis.

For the Springs sites, several water quality constituents were significantly related to flows including alkalinity, calcium, chloride, magnesium, potassium, sodium, and sulfate. This was not a surprising result as it is well known that water that has been in contact with limestone for a relatively short length of time should have low concentrations of calcium and bicarbonate ions; water with a longer period of residency within the flow system should typically have higher concentrations. Total Dissolved Solids is a measure of chemical constituents dissolved in the

groundwater and, in west-central Florida, TDS is mostly influenced by the concentrations of the major ions: calcium, bicarbonate, magnesium, sodium, sulfate and chloride. TDS can be used to estimate the relative residence time of ground water in the aquifer and typically increases as the length of groundwater flow paths increase (SWFWMD 2001).

Nitrogen enrichment is an ongoing concern due to the presence of algal mats (filamentous and epiphytic algae) which were linked to excessive nutrient concentrations. We reevaluated relationships between flow and all forms of available nitrogen for completeness and found that while some statistically significant relationships with flow were established, the results were inconsistent and not directly useful for supporting reevaluation of minimum flows for the Chassahowitzka River System. No significant relationships were found with any organic or total forms of nitrogen in the Springs dataset. In some cases, regressions on inorganic forms of nitrogen (nitrate, nitrite, and ammonia) resulted in significant relationships to flows but were inconsistent. For example, total nitrite at Blue Run Spring suggested a small magnitude positive relationship; increasing concentrations with increasing flow. However, the results of the same analysis for dissolved nitrite in both Blue Run and Ruth Springs suggest an inverse relationship; lower concentrations with higher flows. The only other significant positive relationship observed for a form of nitrogen was for nitrate at Crab Creek Spring and the same analysis on the dissolved fraction of nitrate was negatively related to flow. These findings support those reported by Upchurch et al. (2008) in an analysis of the relationships of nitrate and flows for springs in the Suwannee River Water Management District where 50% of the springs demonstrate that nitrate concentrations increase as discharge from the spring increases; forty-five percent of the springs show no correlation between discharge and nitrate, and 5% have relationships where high discharge was related to lower nitrate concentrations. Upchurch et al. (2008) concluded that that minimum flows could not effectively be utilized to control nitrate discharging from the springs by promoting high discharge. The analysis in this evaluation of water quality in the Chassahowitzka River therefore supports the findings of Upchurch et al. (2008) and Heyl (2012) that the current evidence does not support the conclusion that there is a consistent relationship between these forms of nitrogen and flows.

A similar analysis was conducted for the Estuary data defined as those sample locations outside the mouth of the river in the nearshore estuarine environment. The hydrodynamic model developed to support the reevaluation of minimum flows for the Chassahowitzka River System is considered the best available tool for evaluating the effects of flows on salinity within the mainstem of the river and; therefore, while salinity relationships were evaluated as part of this study, they were only considered as potential criteria for that area outside the hydrodynamic model domain (i.e., the Estuary resource of concern). Analyses of the Estuary data led to conclusions similar to those for the Springs data. There were several significant relationships between spring flows and salinity for stations outside of the mouth of the river; however, given the distance from headsprings, it is more likely that salinity in the estuary is driven by a combination of spring discharge, coastal runoff, wetland storage, direct rainfall, and freshwater discharges from other nearby coastal areas that are all seasonally dependent and to some extent correlated with one another. Therefore, it is not recommended that these relationships in the Estuary be used to support the establishment of minimum flows for the Chassahowitzka though it may be worth future investigation to identify the relative impacts of these factors on the nearshore estuarine resources of the river.

Analysis of the Mainstem of the river did reveal evidence that chlorophyll a distributions, a proxy for phytoplankton abundance, were found to be significantly related to flows under certain conditions. While healthy phytoplankton populations are essential for a healthy estuary, an excess in phytoplankton abundance can have negative impacts on ecosystem health, and, while chlorophyll concentrations are generally low in the Chassahowitzka River System, there is evidence that the system can be susceptible to high phytoplankton biomass with several chlorophyll concentrations observed above 25 ug/l.

For regulatory purposes, the Florida Environmental Protection (FDEP) has adopted a Total Maximum Daily Load (TMDL) for nitrate in the springs and for total nitrogen (TN) in the upper river (Dodson, et al. 2014). The FDEP split the river into Waterbody Identifiers (WBID). A river kilometer (Rkm) system is also established for the river which begins at the mouth (Rkm 0) and ends at the headsprings (Rkm 9). WBID 1361 the downstream river segment, from Rkm 0 to Rkm 5.9. The TMDL applies to those WBIDs upstream of 1361 including the upper run of the river (WBID 1348D), the main spring vent complex (WBID 1348Z) and Betee Jay Springs (1361B). WBID 1361 is governed by state numeric nutrient criteria (NNC) which include nitrogen, phosphorus and chlorophyll a thresholds for compliance. The upper WBIDs are governed by the TMDL and include nitrogen thresholds but do not include a chlorophyll a threshold.

A mixed-effects logistic regression model was used to predict the probability of a chlorophyll a sample exceeding the NNC of 3.9 ug/l for WBID 1361. While the NNC is defined as an annual geometric mean, because the median and annual geometric mean are related statistics, increasing the exceedance of the standard value for individual samples increases the risk that the AGM will be exceeded. Given that the upstream WBID does not have a site-specific standard for chlorophyll and that the upstream segment is impaired and has a TMDL while the downstream WBID is currently meeting its designated use, the downstream criterion value for chlorophyll was applied to the entire system. The WBID boundary is in an unfortunate location for the evaluation of chlorophyll as it bisects the peak of the spatial distribution. That is, there are low concentrations in the most upstream section of the river, the chlorophyll concentrations tend to peak right at the WBID boundary, and then decrease towards the mouth. This complicates interpretation of the effects of flows on the regulatory criteria but the decision was made to model the effects of flows on the spatial distribution of chlorophyll irrespective of the WBID boundaries, though it limits direct inference to the application of the results to statutory rules regarding impairment. The model results suggested that chlorophyll a distributions were related to spring flows and that there were complex interactions between spring flows, season and location within the mainstem of the river that were predictive of chlorophyll distributions. The model results suggested that reduced flows increased the probability of a sample exceeding the state standard value. The effect of flows was most apparent in May when flows are typically at their seasonal minima.

A principal objective of this study was to perform analysis that could be used to support the reevaluation of the minimum flows for the river. To use the model described above to evaluate the effects of flow reductions on chlorophyll concentrations in the Chassahowitzka River System, a “Baseline” condition reflecting flows unimpacted by withdrawals and 1% to 15% flow reduction scenarios (in 1% flow-change increments) were developed. The period of record for evaluating the flow reduction scenarios was 1998 -2017 to correspond with the period of measured flows within the system. In addition, because the response of chlorophyll to flow was primarily constrained to the portion of the river above river kilometer 4.9, this area was used for the evaluation. Again, although the boundary for 3.9 ug/l chlorophyll standard bisects the peak of the spatial chlorophyll distribution and the WBID associated with the standard only extends upstream to river kilometer 5.9, the standard was used for assessment of chlorophyll concentrations for the entire portion of the river at or upstream of Rkm 4.9. The results of the flow reduction evaluation suggested that a 12% reduction in flows would increase the individual sample exceedance frequency over the Baseline condition by 15% for this section of the river. The 15% change threshold is a prescriptive standard commonly used to identify “significant harm” for criterion used to establish minimum flows.

While the chlorophyll-flow modeling effort utilized the site-specific chlorophyll threshold value to evaluate response to changes in flow, the results were not intended to be used as a direct assessment of whether or not changes in flow would result in compromises to the river’s “Designated Use” as defined in State statute. The chlorophyll concentrations tend to peak at the WBID boundary and then decrease both towards the river mouth and upstream of the boundary. This complicates interpretation of the effects of flows on the regulatory criteria. In addition, it would be beneficial to validate the model with additional data prior to application in a regulatory setting. Instead, the analysis illustrates the utility of this type of modeling to assess the sensitivity of chlorophyll a distribution (a proxy for phytoplankton abundance) in the upper 3 kilometers of the river to changes in flow, and suggests the need for more research in this upstream portion of the river that has displayed evidence of sensitivity to changes in flows. Transect data collected at sites throughout the river segment were valuable in this regard because they provided spatially-intensive water quality data. Because of the spatial distribution of the chlorophyll peak within the mainstem of the river, future data collection efforts should consider spatially-intensive sampling in this portion of the system to test hypotheses developed from this work that chlorophyll a distributions are sensitive to changes in flows in the upper reach of the river. Otherwise, this reevaluation has confirmed many of the findings of the District’s original minimum flows report.

1.0 INTRODUCTION

1.1 BACKGROUND

Florida law (Chapter 373.042 F.S.) requires Florida's Water Management Districts or the Department of Environmental Protection (DEP) to establish minimum flows for rivers, streams, estuaries and springs to identify the limit at which further withdrawals would cause significant harm to water resources or ecology of the area. A minimum flow rule for the Chassahowitzka River System was adopted in 2013 (Rule 40D-8.041, Florida Administrative Code or F.A.C.), with a directive to reevaluate the minimum flow within six years of its adoption. The Chassahowitzka River/Chassahowitzka Spring Group and Blind Spring are included on the Southwest Florida Water Management District's 2017 Minimum Flows and Levels Priority List for reevaluation (and its draft 2018 Priority List), with finalization due in 2019.

As part of the District's efforts to reevaluate the Chassahowitzka River System minimum flow, a work effort was contracted with Janicki Environmental, Inc. in July 2018 (Task Work Assignment (TWA) No: 18TW0001116) to conduct exploratory data analysis through investigation of relationships between springs flows and system water quality. The specific tasks within this TWA consisted of data compilation, exploratory data analysis, stochastic predictive modeling and synthesizing information in support of the minimum flow reevaluation.

This report addresses work under multiple tasks of the TWA. Initial tasks included the compilation of available water quality and water quantity data for the Chassahowitzka River and the creation of a Microsoft Access database and database inventory. Additionally, descriptive statistics and plots were generated for each metric of interest to describe both the univariate characteristics and the seasonal and inter-annual distributions. Screening methods were used to identify and qualify potential anomalous data evident in the datasets in the Access database. The Access database, summary statistics, and tabular/graphical output from the described analyses are provided as attachments to this report. Methods used to compile data, and complete the initial data exploration are described in Section 2.1. Subsequent to initial data compilation and exploration, a statistical analysis plan was developed which outlined the analytical methods used to approach each of the various data types that exist in the master database. The statistical analysis methodology is detailed in Section 2.2.

The results section (Section 3.0) of this report details the results of the outlined statistical analyses. The results section is organized by "Resources of Concern" which were identified as part of the exploratory data analysis process. Within each sub-section of the results section, the specific analytical approach is described as it was applicable to support the reevaluation. This effort included results of exploratory data analysis, identification of a conceptual model, identification of the statistical approach and results of application of the approach as they pertain to supporting the reevaluation for the water quality component. Finally, when applicable, the assumptions and limitations of the approach are also described.

1.2 REPORT OBJECTIVES

The specific objective of the report was to provide documentation, exploratory analysis, and statistical inference regarding the relationships between flows and water quality constituents in the Chassahowitzka River System to support the assessment of the water quality environmental value as part of the reevaluation of minimum flows for the Chassahowitzka River.

2.0 DATA AND METHODS

This section describes the sources of data utilized in this report, the preliminary methods for initial data exploration, as well as the statistical analyses utilized.

2.1 DATA COMPILATION

Figure 2-1 provides an organizational overview of the types of data compiled for this work assignment. Multiple datasets, including all water quality, groundwater discharge, and various ancillary datasets were provided by the District to the project team. Water quality data provided by the District include data from multiple fixed station sampling programs consisting of monthly/quarterly sampling, and from more recently deployed continuous water quality monitoring stations. These programs are described in the subsections below. Additionally, period of record river discharge, rainfall and tide data were downloaded as specified below. Individual datasets were input into the Statistical Analysis Systems (SAS) software package for summarization and analysis

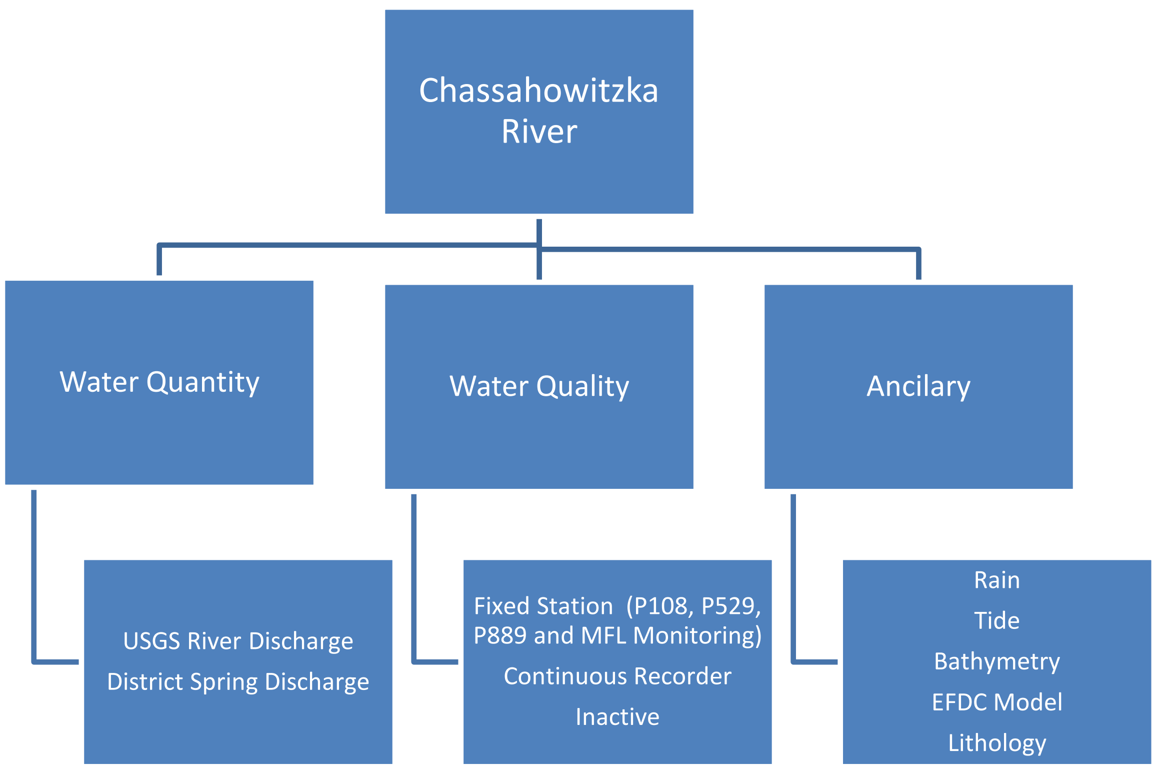


Figure 2-1. Organizational chart for data compiled for this report.

2.1.1 Active Water Quality Monitoring

Active water quality monitoring networks include District Project P108 (Coastal Rivers, Figure 2-2), P529 (Project COAST, Figure 2-3), P889 (Springs, Figure 2-4), and continuous monitoring stations (Figure 2-5).

Fixed surface water stations sampled as part of the District's P108 (Coastal Rivers) monitoring network are displayed in Figure 2-2. Sampling began in late 2005. These stations are sampled quarterly for the District's standard surface water suite of field and laboratory water quality parameters (Table 2-1). Additionally, the continuous monitoring station "Chassahowitzka River Near Homosassa" is co-located at P108 station CV0.

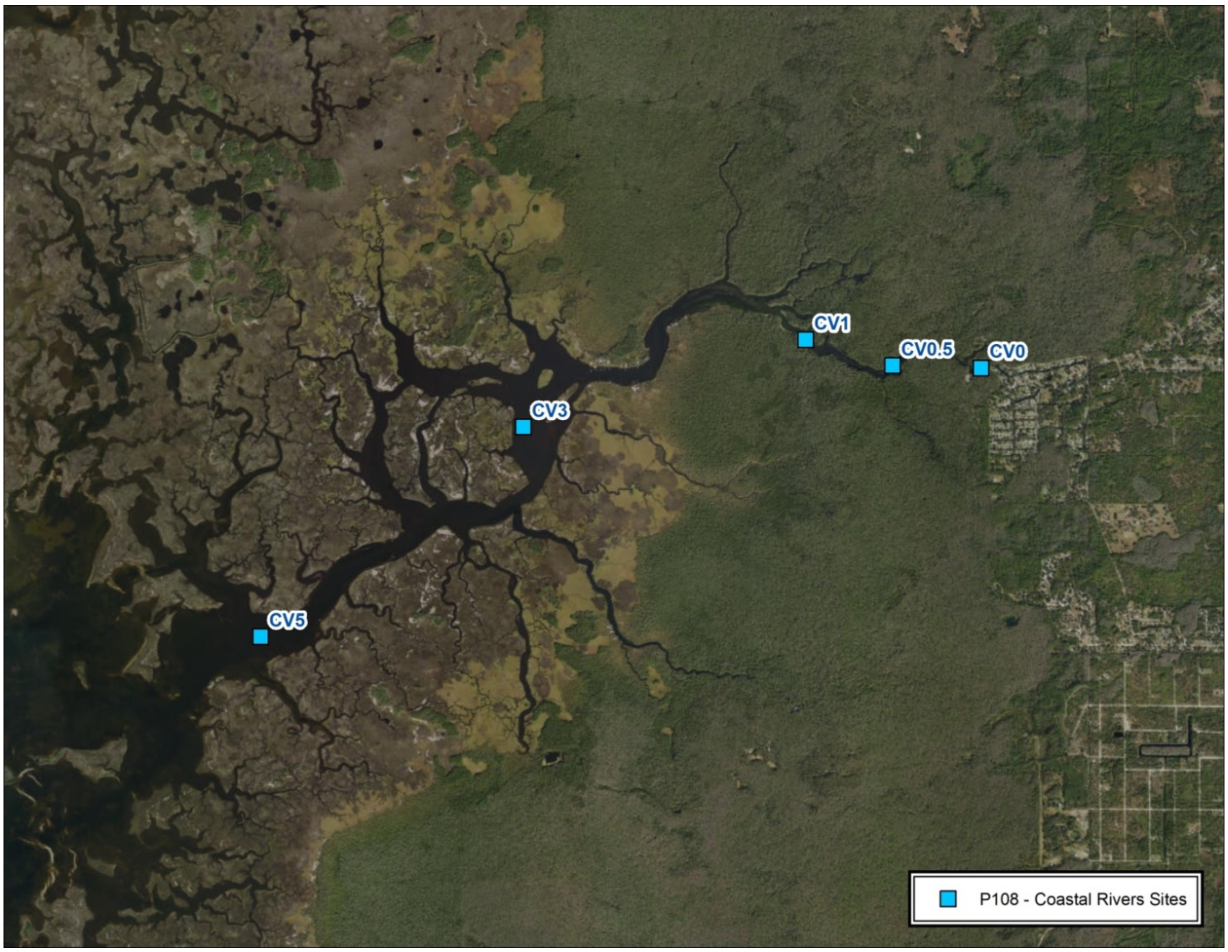


Figure 2-2. Active surface-water sampling conducted as part of the District's surface water quality monitoring network P108 (Coastal Rivers).

Table 2-1.	**Standard Suite of surface water quality parameters analyzed for the District's P108 monitoring network.**
Ammonia (N) (Total)	pH (Total)*
Calcium (Dissolved)	Phaeophytin (Total)
Chlorophyll a (Total)	Phosphorus- Total (Total)
Color (Dissolved)	Potassium (Dissolved)
Depth (Total)*	Residues- Nonfilterable (TSS) (Total)
Depth, bottom (Total)*	Residues- Volatile (Total)
Dissolved Oxygen (Total)*	Salinity (Total)*
Iron (Dissolved)	Secchi-horizontal (Total)*
Magnesium (Dissolved)	Secchi-vertical (Total)*
Nitrite+nitrate (N) (Total)	Sodium (Dissolved)
Nitrite (N) (Total)	Specific Conductance (Total)*
Nitrogen- Total (Total)	Temperature (Total)*
Orthophosphate (P) (Dissolved)	Turbidity (Total)

*indicates field parameters

Figure 2-3 displays the stations associated with Project P529 (Project COAST). Data collection at these stations began in 1997 and sampling occurred on a monthly basis for a limited suite of field and laboratory parameters by the University of Florida until 2010. Beginning in 2015, the District began collecting and analyzing the original 10 stations on a quarterly basis. The District also expanded the suite of parameters to match the parameter list for P108 (Table 2-1).

Figure 2-3. Active surface water stations associated with P529 (Project COAST).

The District established a dedicated monitoring program for the spring vents in the Chassahowitzka in 1992 and has conducted routine sampling on a quarterly basis since then. Standard District water quality parameters for spring sampling are provided in Table 2-2.

Table 2-2. Standard District groundwater (spring) parameters.

Alkalinity (Total)	Nitrogen- Total (Total)
Aluminum (Dissolved)	Orthophosphate (P) (Dissolved)
Ammonia (N) (Total)	pH (Total)*
Boron (Dissolved)	Phosphorus- Total (Total)
Calcium (Dissolved)	Potassium (Dissolved)
Carbon- Total Organic (Total)	Residues- Filterable (TDS) (Dissolved)
Chloride (Dissolved)	Silica – Dissolved (Dissolved)
Color (Dissolved)	Sodium (Dissolved)
Dissolved Oxygen (Total)*	Specific Conductance (Total)*
Fluoride (Dissolved)	Strontium (Dissolved)
Iron (Dissolved)	Sulfate (Dissolved)
Magnesium (Dissolved)	Temperature (Total)*
Manganese (Dissolved)	Turbidity (Total)
Nitrite (N) (Total)	

*indicates field parameters

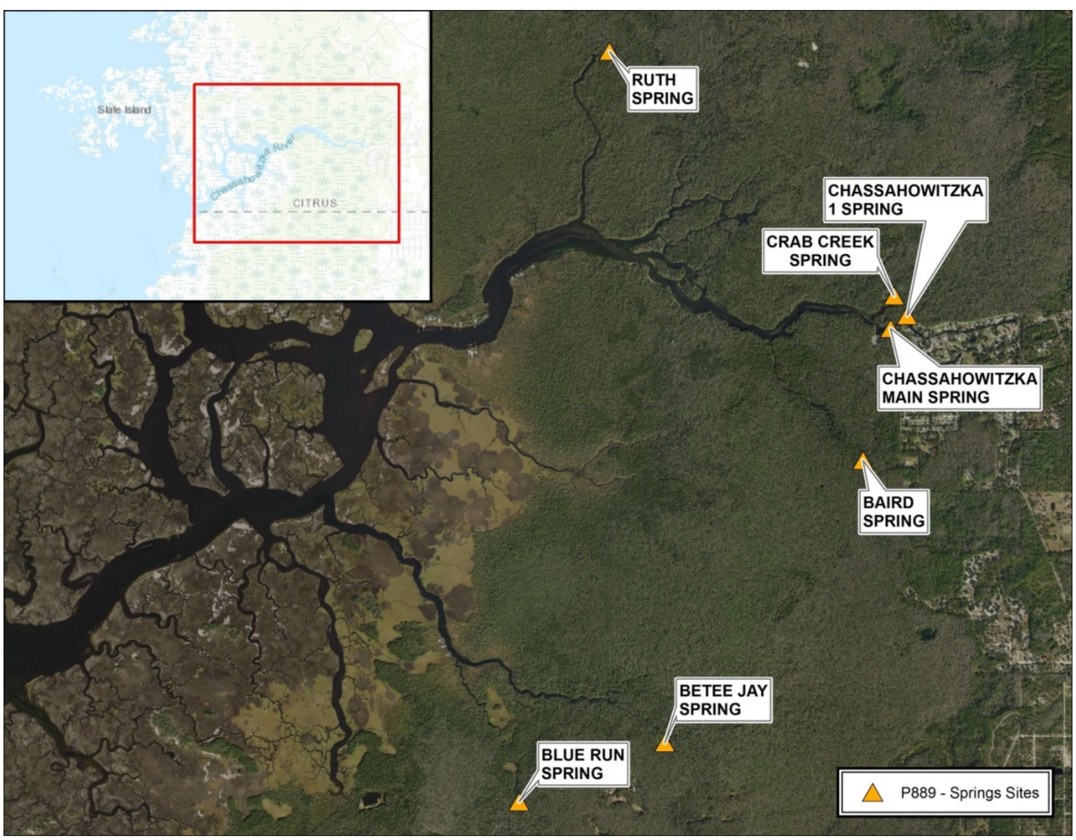


Figure 2-4. Active spring sampling locations in the Chassahowitzka watershed (project P889).

The District provided data for three experimental water quality continuous monitoring recorders on the Chassahowitzka River (Figure 2-5). These data have a short period of record, beginning in 2017. The Chassahowitzka River near Homosassa site is monitored continuously for temperature, pH, conductivity, salinity, dissolved oxygen and depth. The other two sites are monitored for a broader suite of parameters listed in Table 2-3. In addition to the continuously recorded parameters, grab samples are taken at a frequency of every three to four weeks for the full suite of parameters listed in Table 2-3.

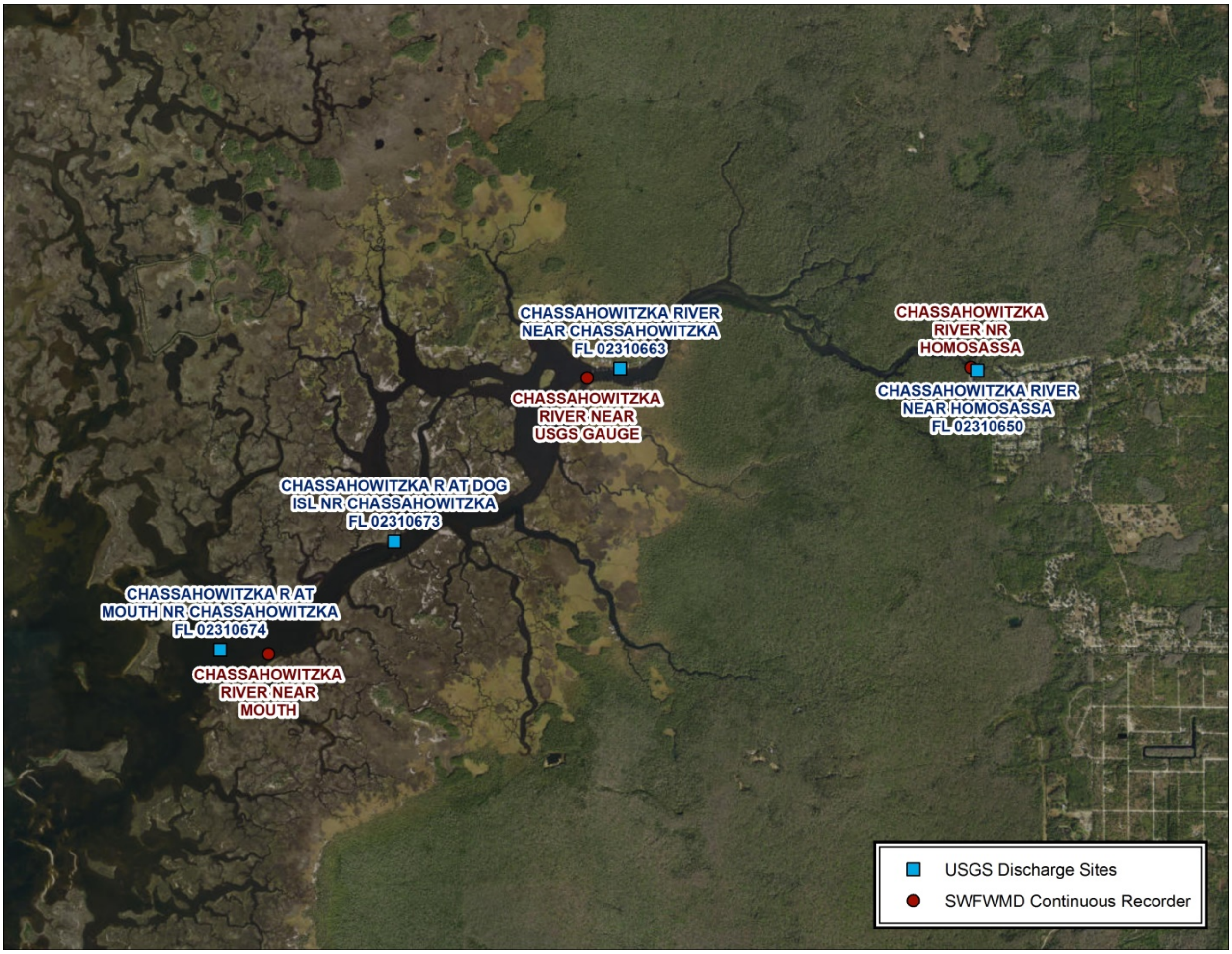


Figure 2-5. Water quality continuous recorders (red circles) on the Chassahowitzka River. Blue squares indicate the locations of USGS river discharge gages.

Table 2-3	Parameters measured at the continuous recorders at Chassahowitzka River Near Mouth and Chassahowitzka Near USGS Gauge.
Temperature	fDOM
Depth	Chlorophyll
Conductivity	Turbidity
pH	Salinity
Dissolved Oxygen (mg/L and %)	Nitrate
Light Spectrum	Dark Spectrum

2.1.2 Inactive Water Quality Sampling

In addition to the active sampling stations in the Chassahowitzka, the District provided datasets from previous studies in the Chassahowitzka that were initiated to better understand the relationship between water quality and ecology of the system. In particular, the District contracted with the University of Florida to conduct sampling and analysis to evaluate the relationship between water quality and ecology of the five spring fed river systems in west central Florida (Frazer et al. 2001; 2006). This study was conducted via quarterly sampling between 1998 and 2011 (with a data gap between 2001 and 2003). For this study, 20 locations (i.e. transects) were established along the length of the river, with three sampling points located laterally across the river at each of the 15 upstream transect locations and a single site at the lowest five sites in the system. The water quality constituents measured associated with the study included data on alkalinity (mg/l), chlorophyll a (ug/l), color (pcu), dissolved oxygen (mg/l), ammonium (ug/l),nitrate (ug/l), soluble reactive phosphorus (ug/l), salinity (psu), temperature , total nitrogen (ug/l), and total phosphorus (ug/l).The transect locations are provided in Figure 2-6. The other inactive stations had few sampled over short or intermittent duration and were not useful in support of reevaluation of water quality in support of minimum flows. Surface water quality constituents with more than 20 observations from inactive stations are listed in Table 2-4.

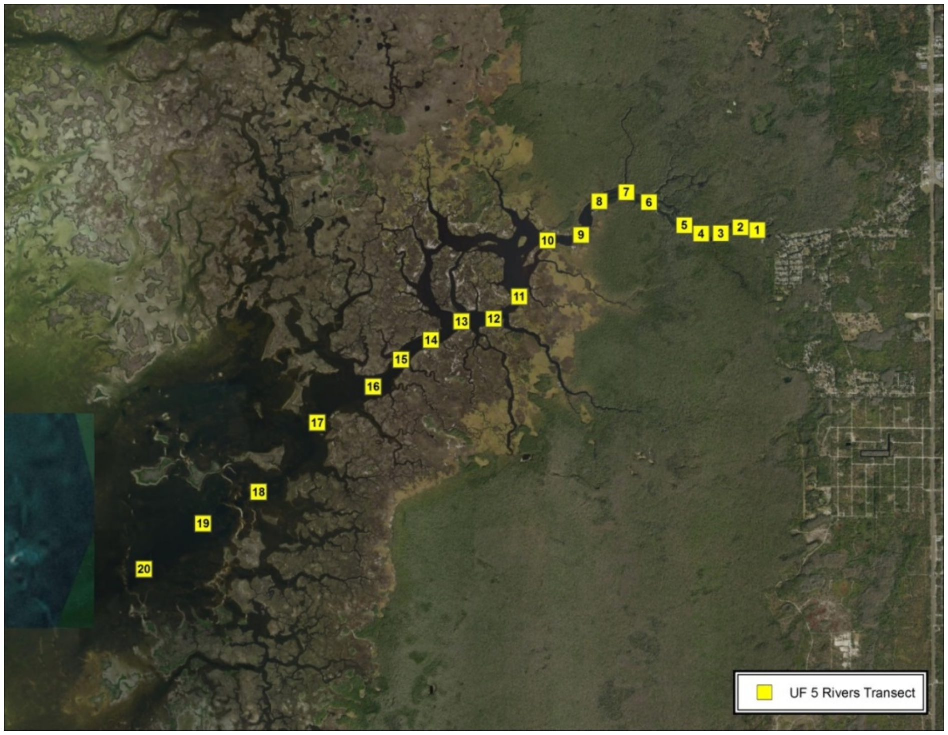


Figure 2-6. **University of Florida 5 Rivers Study transect locations on the Chassahowitzka River.**

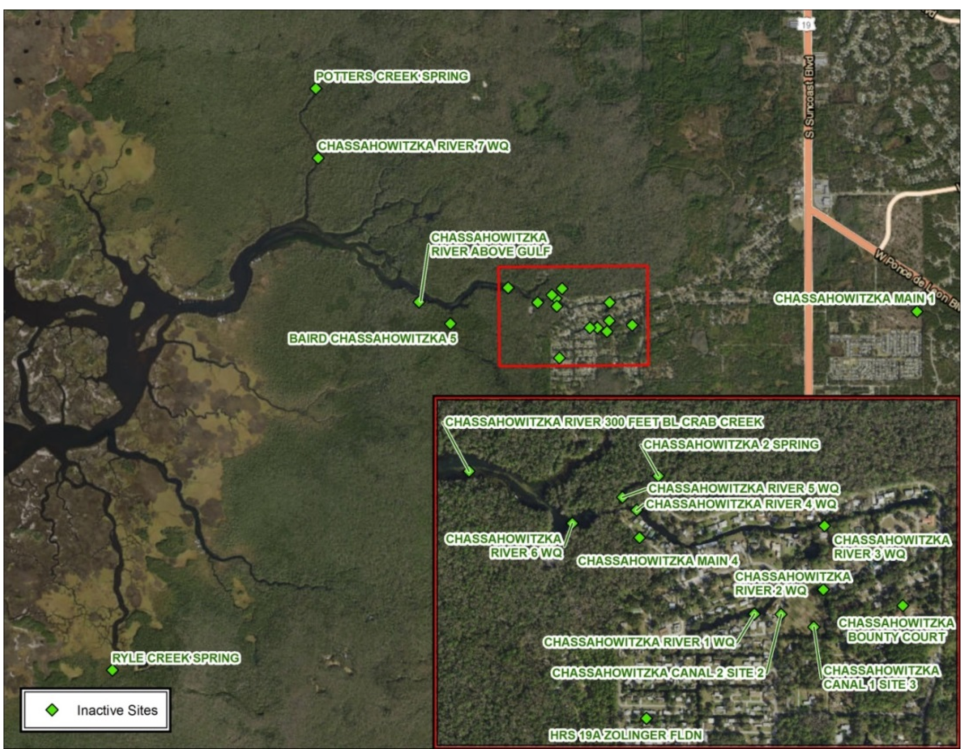


Figure 2-7. **Inactive water quality monitoring stations on the Chassahowitzka River for which the District provided data.**

Table 2-4. Surface water quality parameters sampled from Inactive water quality monitoring with n > 20.	
Ammonia (N) (Total)	pH (Total)*
Calcium (Dissolved)	Pheophytin (Total)
Chlorophyll a (Total)	Phosphorus (Total)
Color (Dissolved)	Potassium (Dissolved)
Depth (Total)*	Residues- Nonfilterable (TSS) (Total)
Depth, bottom (Total)*	Residues- Volatile (Total)
Dissolved Oxygen (Total)*	Salinity (Total)*
Iron (Dissolved)	Secchi-horizontal (Total)*
Magnesium (Dissolved)	Secchi-vertical (Total)*
Nitrite+nitrate (N) (Total)	Sodium (Dissolved)
Nitrite (N) (Total)	Specific Conductance (Total)*
Nitrogen- Total (Total)	Temperature (Total)*
Orthophosphate (P) (Dissolved)	Turbidity (Total)

*indicates field parameters

2.1.3 Hydrologic Data

USGS discharge and/or stage data for the Chassahowitzka River were downloaded for the available period of record from the National Water Information System (NWIS; https://waterdata.usgs.gov/nwis). These gages are shown below in Figure 2-8. USGS discharge data include a data qualifier indicating whether each data record has been "Accepted" (A) or remains "Provisional" (P). Much of the discharge data from late 2017 into 2018 are flagged as "Provisional" and such data should be used with discretion.

Figure 2-8. Location of Chassahowitzka River USGS flow/stage gages.

In addition, rainfall data (station name = Inverness3SE, station id = USC00084289) were downloaded for the period of record from the National Climate Data Center (NCDC; https://www.ncdc.noaa.gov/cdo-web/). Tide data (Station 872750 Cedar Key, FL; datum = MLLW) were downloaded for the period of record from the National Oceanic and Atmospheric Administration (NOAA) Center for Operational oceanographic Products and Services Tides and Currents webpage (https://tidesandcurrents.noaa.gov/). These sites are shown in Figure 2-9.

Figure 2-9. Locations of NCDC rainfall station and NOAA tidal gage relative to the Chassahowitzka and Homosassa Rivers.

2.1.4 Data Screening Methods

The Task 2 report was accompanied by a master database of all available water quantity, water quality, and available ancillary datasets compiled for the Chassahowitzka River and quality control procedures were used to identify potential anomalous values. Many of the raw datasets downloaded, or received from the District, contain at least one column indicating the quality of each data record. Water quality grab sample data include a designated qualifier column containing FDEP data qualifier codes as applicable. A list of these qualifier codes is provided in Appendix A. Certain qualifier codes indicate the data should not be utilized in analyses. The following qualifiers were not used in the analysis:

"Y = flag for improperly preserved sample;

"Q = flag for out of hold time;

"T"= not to be used for analysis; and

"?" = data rejected and should not be used.

It is important to note that a data qualifier including “U” indicates that a compound was analyzed for but not detected. The value associated with the qualifier is the laboratory method detection limit (MDL) though the actual MDL values were not always reported. These values were retained as reported for analytical purposes.

Two data screening methods were used to identify potential anomalous values in each examined dataset. The purpose of the screening methods was simply to identify data points for further investigation, no data were eliminated from the database based on this analysis. Two screening methods were used; an extreme value screening method (e.g. +- 3 standard deviations from the mean) and a functional screening method that evaluates deviations from an expected value based on a timeseries of data using robust regression (SAS Institute, Inc. 2016). Columns were added to the database to identify whether or not each value met the criteria to qualify as a specific data point worthy of further investigation as an anomalous value and these columns were later used in the statistical analysis. A complete list of descriptive statistics and descriptive plots for all constituents evaluated and delivered in the master database was delivered in fulfillment of Task 2 and is provided in this document as Appendix B.

2.2 Statistical Analysis Methods

A statistical analysis plan was developed as Task 3 of this work effort. The Task 3 document outlined a conceptual analytical pathway to guide the analysis. The analytical pathway is outlined in Figure 2-10.

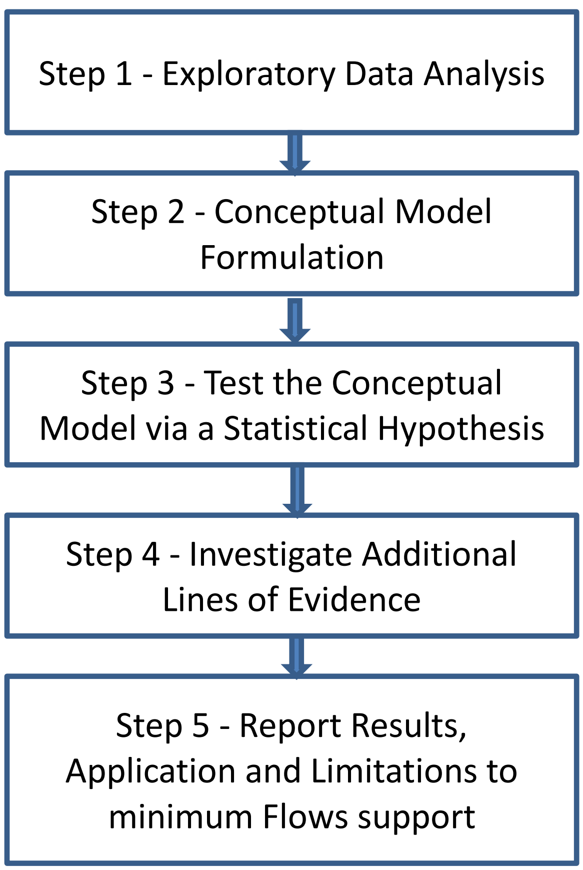


Figure 2-10. Analytical flow path in support of reevaluation of the Chassahowitzka River MFL.

The statistical tools applied to this project include ordinary least squares linear regression, general and generalized linear models with the potential inclusion of random effects, timeseries modeling, and semi- and non-parametric techniques including regression and classification trees. Each of these tools is described in a sub-section below.

2.2.1 Ordinary Least Squares Regression

Ordinary least squares (OLS) regression was used primarily as an exploratory data analysis tool to describe bivariate relationships between flow and a particular analyte of interest. OLS regression maps a response variable such as salinity to a potential explanatory (predictor) variable (e.g. spring discharge). This is accomplished by defining an intercept and a slope for the predictor that defines a straight line minimizing the sum of squared deviations from the line (Zar 1984). Application of OLS regression provides a reference line through the bivariate distribution that can aid in the selection and elimination of those analytes that are not linearly responsive to flow. A linear regression equation is expressed as:

$$Y_i = \alpha + \beta X_i + e_i$$

Where :

Y is the response, X is the predictor, i is an index to the indididual observation,

e is the error, α is the intercept, and β is the slope of the regression line.

The regression coefficient of determination (R^2) is one measure of the variance in the dependent variable that is explained by the model. In linear regression, it is assumed that the data are independent samples from the population. Another important assumption of linear regression is that the error term of the model is normally distributed, with constant variance. However, at times, water quality data do not conform to these assumptions. Data transformations such as natural log transforms can help to satisfy the assumption of normality and heteroscedasticity but will not correct for dependencies in the data structure. Given the number of analytes evaluated in the exploratory data analyses, linear regression was used as a screening method to refine hypotheses related to the overarching project goal and more sophisticated regression modeling techniques were used where there was a potential to generate inferences that could inform the reevaluation of minimum flows for the Chassahowitzka River. These more sophisticated techniques are described in the next sub-section.

2.2.2 General and Generalized Linear Models

General and generalized linear models are extensions (generalizations) of OLS regression models that allow for more flexibility in accounting for artifacts of the data that may affect the underlying assumptions of OLS regression. General linear models are applied when the response variable is continuous and generalized linear models are applied when the response variable is binary or count data. Both classification and continuous predictor variables are allowed and can be expressed as either fixed or random effects representing the deterministic component or the variance component of the model, respectively (Littell et al. 1996).

An example of using a general mixed effects model is provided by the equation below that regresses chlorophyll concentrations on spring discharge. The deterministic component produces a parameter estimate of the intercept and slope and tests that they are different from zero while the random component of the model allows for each station to have a separate intercept and for the correlation among samples collected at different stations to be accounted for in the error variance. The likelihood ratio test can be used to compare the mixed effects parameterization relative to the null model and Akaike Information Criteria (AIC) can be used to evaluate the model improvement for nested models of the same family. We used a modification of AIC that includes a penalty for including additional model parameters (AICC) in the model evaluation (SAS Institute, Inc. 2016). Residual diagnostics are also helpful to assess model fit and assumptions associated with the regression.

$$Y_{ij} = \beta_0 + \beta_{0j} + \beta_1 * X_{ij} + e_{ij} + e_j$$

Where:

Y_{ij} = chlorophyll concentraion for each sample (i), and station (j)

X_{ij} = spring discharge for each date (i), and station (j)

β_0 = overall intercept

β_{0j} = random intercept for station

β_1 = deterministic effect of spring discharge on chlorophyll

e_{ij} = residual ($N_{(0)\ iid}$)

ej = residual covariance among samples taken at the same station

Generalized linear models are similar to general linear models except the response variable can be from alternative distributions. Logistic regression is an example of a generalized linear model. The formulation is nearly identical to those models above in that they are all linear (i.e. additive) models but generalized linear models use a link function to map the response variable to a known distribution, generally of the exponential family. Logistic regression in particular is useful if there are important critical threshold values for an analyte of interest, above which results in some adverse effect; for example, if it were known that a chlorophyll concentration above some threshold value resulted in an adverse effect to the ecology of the system. The general equation for a mixed effect logistic regression model is provided below. Notice that there is no error term as the variance is modeled as a function of the mean value; the formulation for the random effects is therefore included within the link function.

$$g(E[Y \mid \gamma]) = \log(\frac{p_{(y=1)}}{1-p_{(y=1)}}) = X'\beta + Z'\gamma$$

Where :

$X'\beta$= Fixed effects

$Z'\gamma$ = Random effects

$\log(\frac{p_{(y=1)}}{1-p_{(y=1)}})$ = link function mapping presence absence to independent terms

$g(E[Y \mid \gamma])$ = response conditional on random effects

γ=Random effect $N(0,\sigma^2\gamma)$

We used this model formulation to evaluate the effects of changes in flows on the water column phytoplankton distribution throughout the mainstem of the river.

2.2.3 Robust Regression

Robust regression is a method to account for highly influential data points that may affect statistical inference (Chen 2002). The method relies on iteratively reweighted least squares which successively down weights observations with large residuals until reaching convergence criteria established to evaluate the change in the parameters' estimates. By iteratively reweighting the values, the resulting parameter estimates are "robust" to the influence of extreme values in the dataset and therefore the procedure is robust to deviations from assumptions about the data distribution, heterogeneity of variance, and other assumptions of traditional OLS regression approaches. Robust residuals are calculated along with robust standard errors and these estimates can be used to evaluate the robust regression fit to the data. Robust regression was performed as part of the quality control checks where outliers were identified in the dataset for further examination. For example, Figure 2-11 provides an example of a timeseries of turbidity measurements for a water quality station in the Chassahowitzka River. The robust regression identified several outliers in the timeseries (denoted by red open triangles) and adjusted the intercept and slope of the regression line to down-weight the influence of those outliers. The principal use of robust regression for this project is to detect outliers in an objective way.

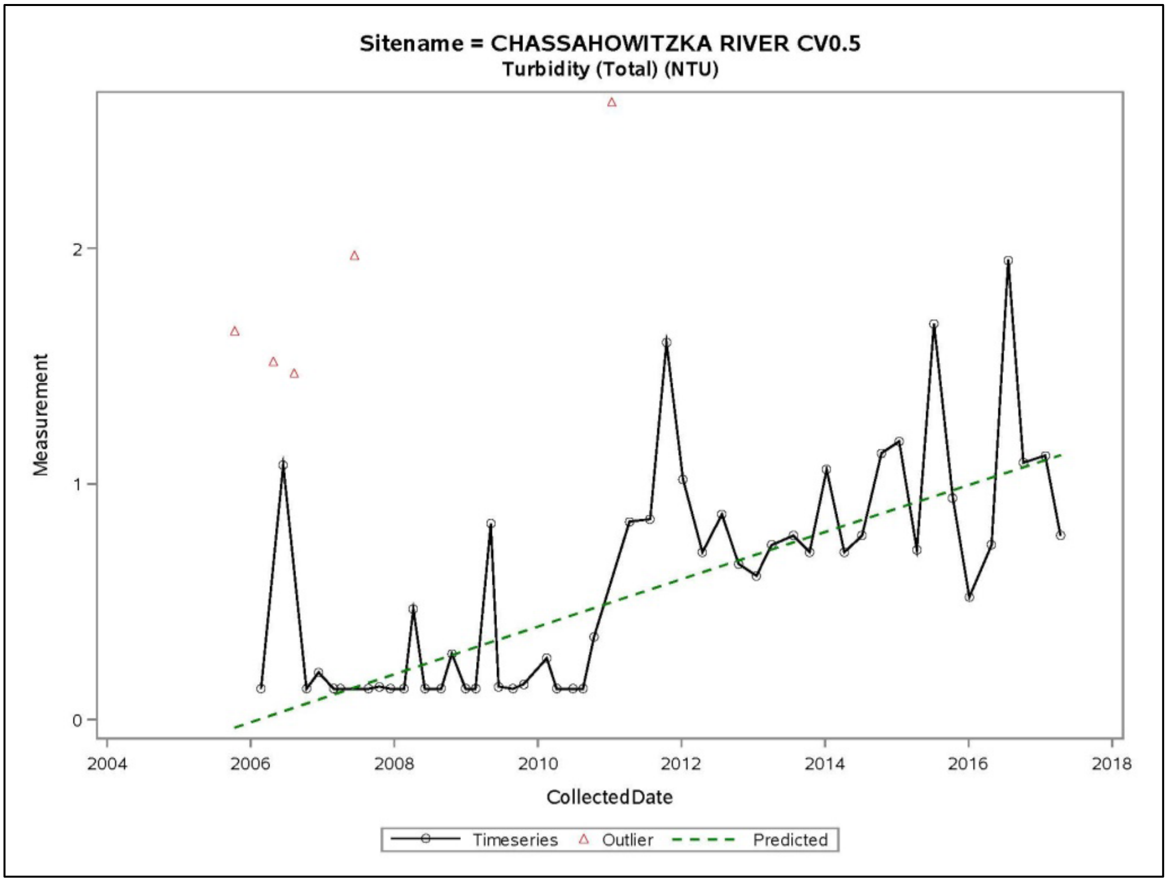


Figure 2-11. Example of application of robust regression to a total turbidity timeseries for a station in the Chassahowitzka River.

2.2.4 Time Series Trend Tests

Evaluation of long-term trends was performed using the seasonal Mann-Kendall (SMK) test for trend (Hirsch et al., 1982; Hirsch and Slack, 1984) which was developed by the USGS in the 1980s to analyze trends in surface-water quality throughout the United States. The SMK test was modified from the Mann-Kendall trend test (MK, a measure of rank correlation to measure the association between measured quantities), in that the MK test is first performed for individual seasons (months or quarters), and the individual results are combined into an overall test for whether the dependent variable changes in a consistent direction (monotonic trend) over time (Helsel et al., 2006). Time series trend tests were conducted for the long-term spring discharge records as well as specific water quality constituents that may be affected by anthropogenic influences over time as well as changes in spring discharge.

2.2.5 Conditional Inference Trees

Conditional inference trees are a class of permutation-based methods also known as “Decision Trees” or “Regression and Classification Trees”. This class of methods is applicable to all kinds of regression problems, including nominal, ordinal, and continuous data. A conditional inference tree methodology (Hothorn et al., 2006) was used as one line of evidence for evaluating water quality stressor-response relationships. The approach is based on recursive partitioning. The partitioning process iteratively searches for a point in the stressors variable which maximizes the difference in the response values between two groups of response data. No *a priori* threshold is specified. The regression tree approach defines the breakpoint as that which maximizes the difference between groups by minimizing the p value associated with some statistical test. The point in the stressor variable at which the p value is minimized, after adjustment for multiple comparisons, is assigned as the breakpoint defining the split of the response variable into two groups. Once the first split is made the process continues to test for subsequent splits that are conditional on the first split; hence, the term "conditional inference" or "conditional probability analysis". Multiple explanatory covariates can be included in the analysis to identify multiple drivers of response dependent on the range of values and can indicate the presence of synergistic relationships among potential explanatory analytes including discharge.

2.2.6 Statistical Analysis of High Frequency Water Quality Data

High frequency (aka continuous) water quality data collection can be a useful tool to identify the different periods of time over which cyclical variability is observed, and to relate these period scales to physical (for example, tides, both diurnal and lunar) and biological (for example, primary production, typically daily) drivers (Downing et al. 2017). One goal of these data collection networks is often to identify important patterns of variation (frequencies) in the continuous monitoring data. For example, the relative variability of within- and between-day variation is important to put grab sample data in the context of within-day variability. In some instances, frequency analysis such as wavelet or spectral analysis can be used to assess different temporal signatures in the underlying data if those periodicities can predict return intervals. Descriptive plots were constructed to put the timeseries data into context of within- and between-date variability. Base functions in the R computing language (R Core Development Team 2008), and the timeseries regression and seasonal decomposition functions in the forecast package (Hyndman 2018) were used as necessary to decompose the timeseries of continuous water quality data, and to identify frequencies relevant for future evaluation. Given the limited period of record for the high frequency data collected in the Chassahowitzka River it is unlikely that analysis of these data will directly inform criteria useful to the reevaluation of the minimum flows. However, analysis provided information on the dominant forms of variability on these data relative to high-frequency periodicities, like fluctuations in tidal amplitudes associated with moon phase, and low-frequency periodicities that represent more long-term seasonal, or possibly flow related, signals.

3.0 RESULTS

The water quality monitoring networks described in Section 2 were established to evaluate different aspects of water quality in the Chassahowitzka River. For example, the "Springs" stations were established to characterize the water quality within the spring vents at, or near, low tide. On the other hand, the UF transects were established to characterize the spatial distribution in water quality for the mainstem of the river using a spatially intensive water quality monitoring design. In this way, each monitoring network represents not only the data sources, but also critical resources of concern within the system. Therefore, the following sections describe the results of data analysis for each of these resources of concern including: the Spring Vents, the Mainstem of the River, and the Estuary. For each of these resources of concern we evaluated the relationship between water quality constituents of interest and flows as directed by the scope of work, and investigated the potential for the constituent to result in an adverse effect that could, under sufficient magnitude and duration, result in significant harm to the integrity of the resource. Therefore, a summary of the flows used for this analysis are also provided as a section in this report.

The organization of the results section follows the description and application of the analytical pathway discussed in Section 2 by first describing the results of exploratory data analysis, then defining a conceptual model that related flows to the resource of concern and then evaluating whether or not a relationship exists that can be useful to support the development of minimum flows for the system. The continuous water quality recorder data represent a special case to evaluate fine scale temporal changes in water quality both within a day and between dates within weeks, months, and seasons. These data have been collected for approximately one year at specific locations in the main stem of the river and therefore are more representative of high frequency variability at a particular location in the river rather than a long-term representation of the expected condition for the system as a whole. The continuous recorder data were evaluated in this context and results provided as a separate section in the report.

3.1 FLOW GAGE OF RECORD

The gage of record for the Chassahowitzka is the USGS Chassahowitzka Near Homosassa (USGS 02310650:Figure 3-1 in red). The District provided a long-term flow record for the reevaluation to be consistent with the methods used in the 2012 MFL report (Heyl et al. 2012) but updated with new data. The continuous daily discharge record for USGS 02310650 began in 1997 and stage measurements began in 1999. The latest reported gage data from 02310650 were downloaded from USGS NWIS. Where index velocity data were available, those data were used; otherwise, data reported by USGS based on regression methods were used. Where data were missing, a regression relationship developed by the District (Heyl 2010) between Weeki Wachee well and flows at the gage of record was used to predict missing values and extend the flow record back in time. It is important to note that the Chassahowitzka River system receives discharge from smaller springs as well as receiving diffuse groundwater discharge but that these are not part of the flow record used to establish minimum flows. While other flow gages are available, their period of record is quite short and more tidally influenced.

Figure 3-1. **Chassahowitzka River USGS flow gages with the long term flow gage of record for the established minimum flow and minimum flow reevaluation identified (in red).**

The timeseries of the final long-term flow record is presented in Figure 3-2. The top plot is the daily flows and the bottom plot is the daily flows with a 21-day lag average overlay. The change in estimation method is evident with reduced variability in the daily record for the period where the estimates are based strictly on well data. The 21-day average timeseries reduced the intra-daily variability considerably but maintains the same underlying trend in the data.

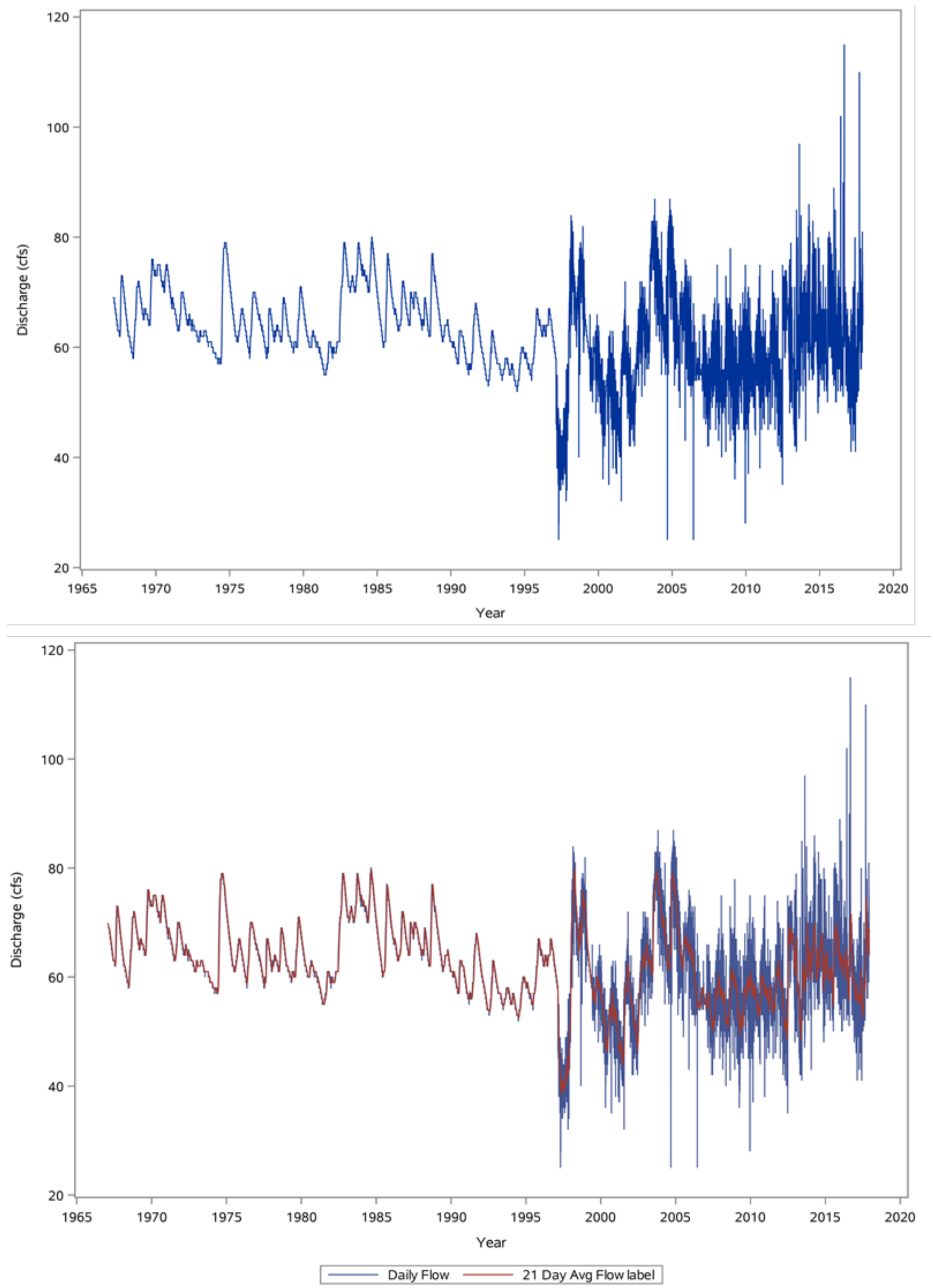


Figure 3-2. **Long-term flow record for the Chassahowitzka River minimum flow analyses, including the daily flows (Top) and the 21 day average flows red overlaid on the daily flow in blue (Bottom).**

Summary statistics and a histogram for the long-term Chassahowitzka River flows are provided in Figure 3-3. The mean (62 cfs) and median (62.5 cfs) values are very similar, within 1 cfs of each other. The range between the 5th (P5) and 95th (P95) percentiles is around 24 cfs, which is less than 50% of the median flow. Quantile plots against the normal distribution (Figure 3-4) suggest slight deviations from the normal distribution.

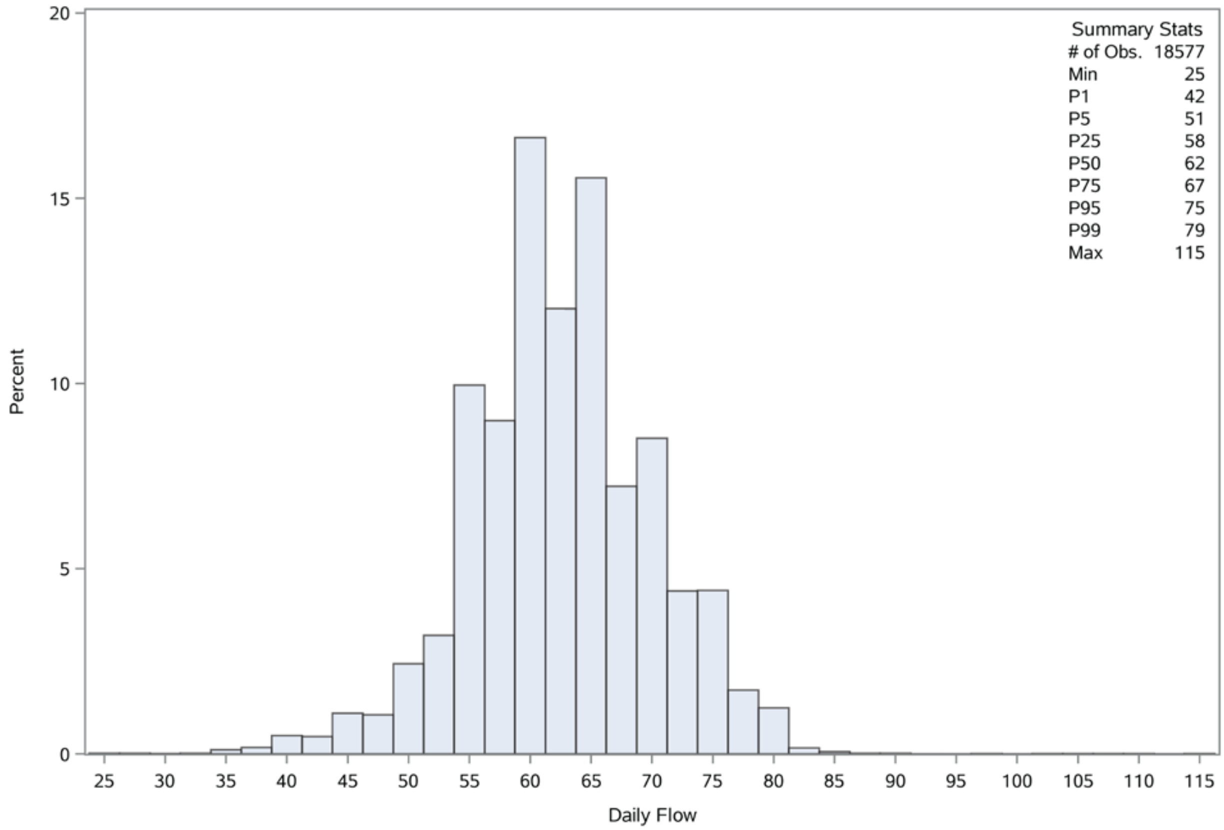


Figure 3-3. Summary statistics and histogram for the long-term flow record for the Chassahowitzka River minimum flow reevaluation.

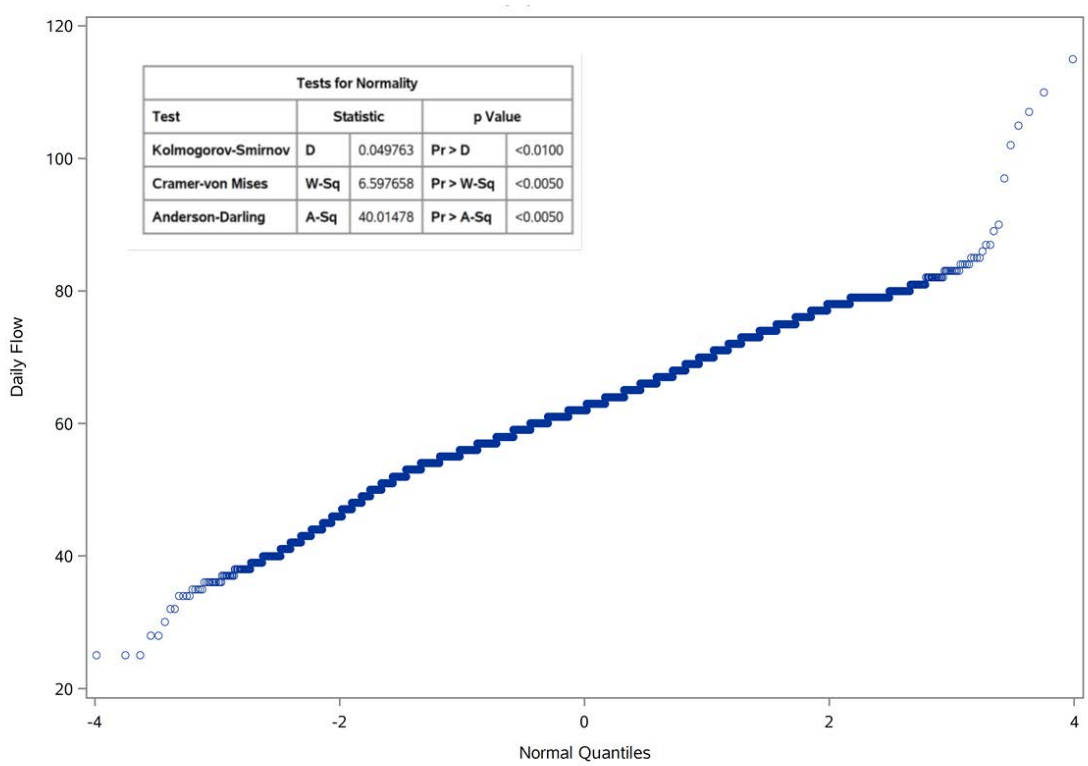


Figure 3-4. Quantile plot of long-term Chassahowitzka flow record against the normal distribution.

Lag averages of the flows were highly correlated out to a 30-day average (Table 3-1) indicating that any of these estimates are nearly equally likely to serve as deterministic components to assess the effects of flows on changes in resources of concern.

Table 3-1. Spearman's rank correlation matrix for lag average flows to 30 days using the long term flow record for the Chassahowitzka River.

Flow Statistic	Daily Flow	2d mean	3d mean	5d mean	7d mean	8d mean	14d mean	21d mean	30d mean
Daily Flow	1.00	0.98	0.96	0.95	0.95	0.94	0.94	0.93	0.92
2d mean	0.98	1.00	0.99	0.98	0.97	0.97	0.96	0.95	0.95
3d mean	0.96	0.99	1.00	0.99	0.98	0.98	0.97	0.97	0.96
4d mean	0.95	0.98	0.99	1.00	0.99	0.99	0.98	0.97	0.96
5d mean	0.95	0.98	0.99	1.00	1.00	0.99	0.99	0.98	0.97
6d mean	0.95	0.97	0.99	1.00	1.00	1.00	0.99	0.98	0.97
7d mean	0.95	0.97	0.98	1.00	1.00	1.00	0.99	0.99	0.98
8d mean	0.94	0.97	0.98	0.99	1.00	1.00	0.99	0.99	0.98
14d mean	0.94	0.96	0.97	0.99	0.99	0.99	1.00	1.00	0.99
21d mean	0.93	0.95	0.97	0.98	0.99	0.99	1.00	1.00	1.00
30d mean	0.92	0.95	0.96	0.97	0.98	0.98	0.99	1.00	1.00

The seasonality in the monthly median flows is portrayed in the box and whisker plots of Figure 3-5. Seasonality is evident in the plot with lower flows in May – July and highest flows in September and October; however, as noted in Heyl et al. 2012, the variation in these flows is small relative to a typical surface water dominated system. Monthly median flows were calculated for the period of record and the Seasonal Mann-Kendal (SMK) test for trend was used to evaluate the trend over time. The SMK tests the slope of the time series trend for each month and combines the results to report a statistic representing the significance of the combined results. The results of the SMK test suggest a significant declining trend in discharge values over time ($p<0.001$:Figure 3-6). These flow trends are very similar to the trends observed in Weeki Wachee Wells that were used to develop estimates of historical flows for the Chassahowitzka (Heyl 2010:Figure 3-7).

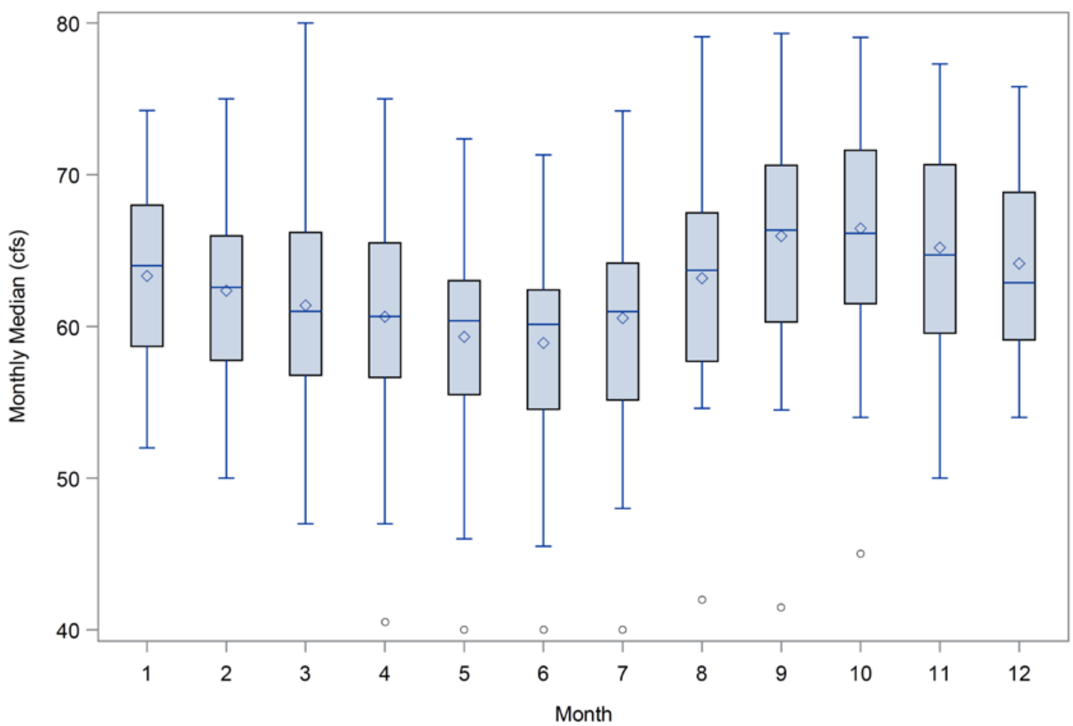


Figure 3-5. Seasonal (monthly) distribution of flows based on the long term record.

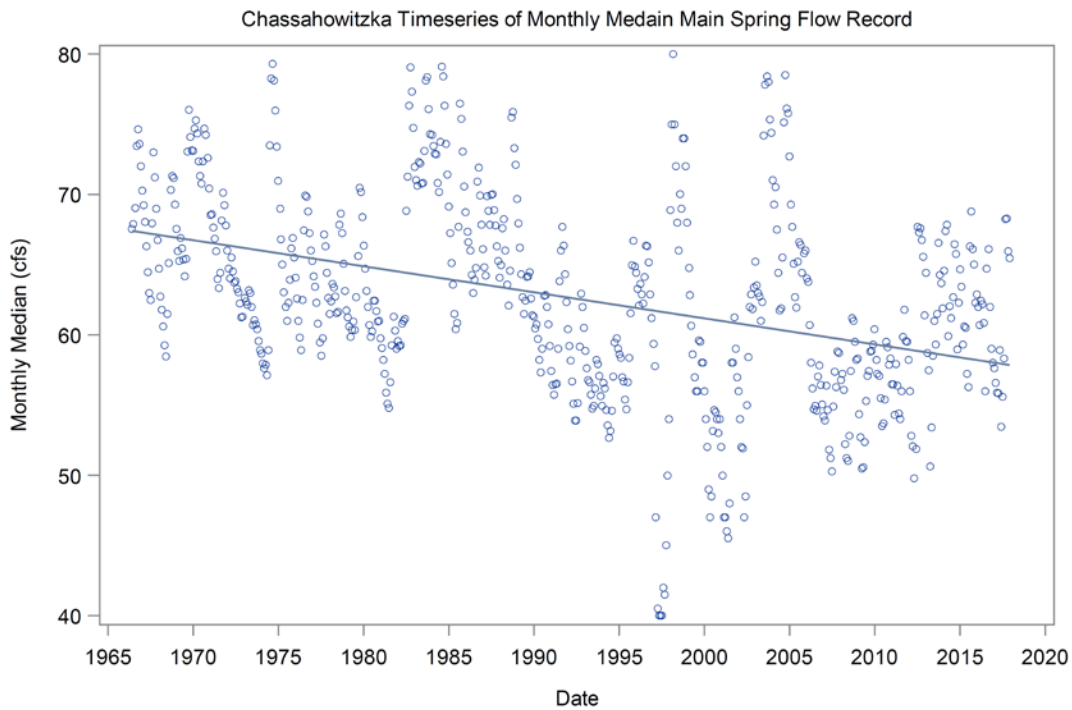


Figure 3-6. **Timeseries of monthly median long-term Chassahowitzka flow record with trend line depicting the decreasing trend in flows over time.**

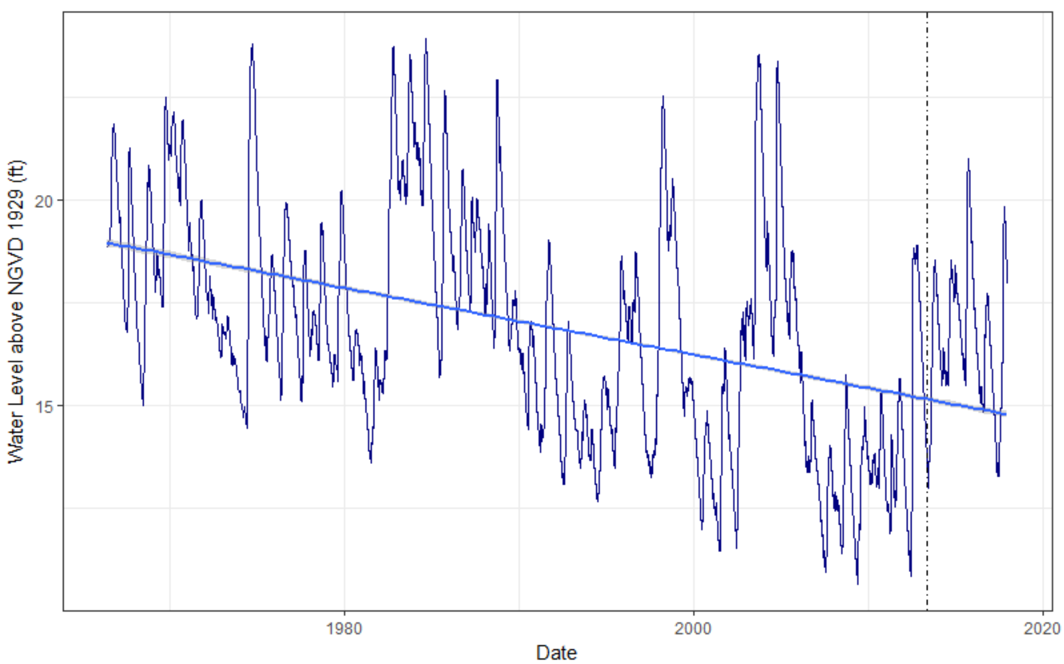


Figure 3-7. **Water levels in Weeki Wachee Well from 16,268 daily values. Dashed vertical line is at start of new well location on 2013-04-30, which has been adjusted by adding 0.3 feet to match with old well location following regression adjustment by the USGS (Kevin Grimsley, personal communication, 2018).**

3.2 Spring Vents

The Springs data included quarterly sampling events beginning in the early- to mid-1990's collected at or near low tide to minimize tidal mixing with seawater. Data come from seven spring vents shown in Figure 3-8

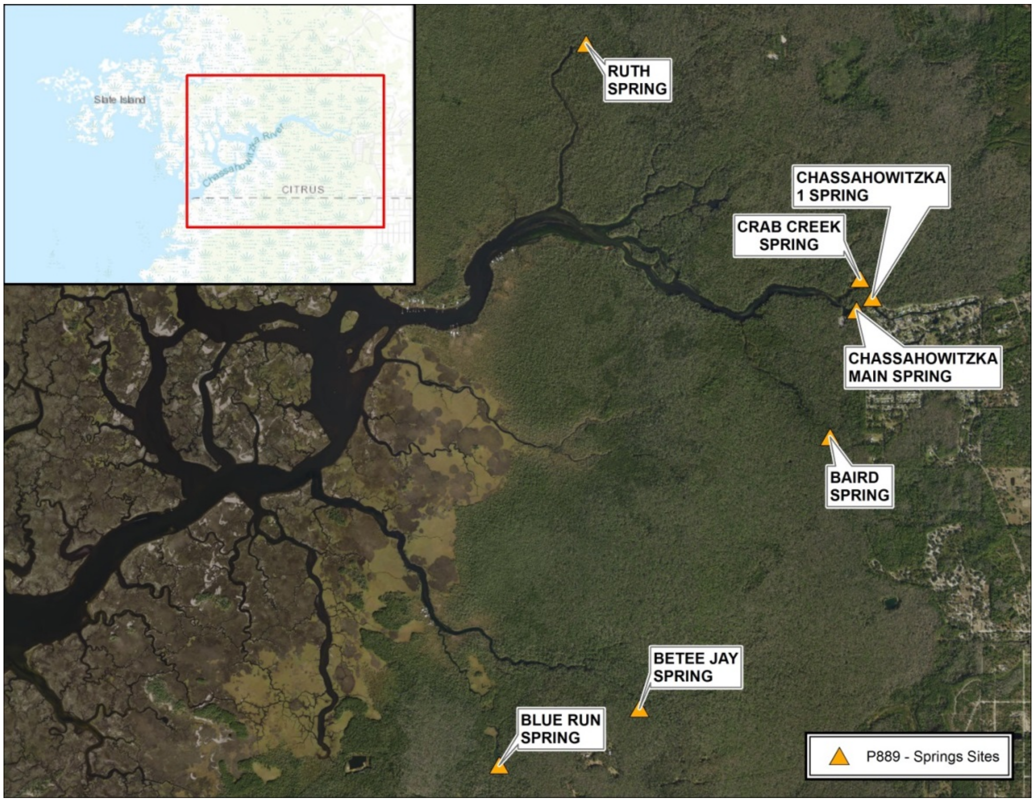


Figure 3-8. Location of sampling sites for the Springs Vent sampling program (District project number P889).

3.2.1 Exploratory Data Analysis

To evaluate the utility of these data to support reevaluation of the Chassahowitzka River minimum flow, linear regression analysis was conducted to test the hypothesis that concentrations of water quality constituents emanating from the spring vents were related to spring discharge. Because spring discharge is estimated as a function of Weeki Wachee well for the majority of the time series, and because, these spring vents are located in several areas without long term discharge records, the long-term flow record developed for reevaluation of the Chassahowitzka River minimum flows was used as the estimate of spring discharge for all spring sites. The water quality constituents evaluated are listed in Table 3-2 and the same-day flow and the 3-day lagged average flow were considered as predictor variables. Outliers indicated by robust regression and data points with qualifiers indicating unreliable data were removed from the analyses. The District developed acceptance criteria for using regression analysis in support of minimum flows evaluations for the Chassahowitzka River (Heyl et al. 2012). The acceptance criteria state that regressions must include a) a minimum 10 observations per variable, b) a plausible trend in the response as a function of flow, c) no significant serial correlation and d) an adjusted coefficient of determination (R^2) of at least 0.3.

Table 3-2. List of water quality constituents evaluated for linear relationships with flow.			
Alkalinity (Dissolved)	Depth (Total)	Nitrogen- Total Kjeldahl (Dissolved)	Stage (Total)
Alkalinity (Total)	Depth, bottom (Total)	Nitrogen- Total Kjeldahl (Total)	Strontium (Dissolved)
Ammonia (N) (Dissolved)	Dissolved Oxygen (Total)	Nitrogen15/Nitrogen14 Isotope Ratio	Strontium (Total)
Ammonia (N) (Total)	Eh, Field (hydrogen electrode)	Orthophosphate (P) (Dissolved)	Sulfate (Dissolved)
Bicarbonate (Total)	Fluoride (Dissolved)	Orthophosphate (P) (Total)	Sulfate (Total)
Biological Oxygen Demand (Total)	Fluoride (Total)	Phaeophytin (Total)	Temperature (Total)
Boron (Dissolved)	Hardness (Total)	Phosphorus (Dissolved)	Total depth at monitored location
Boron (Total)	Iron (Dissolved)	Phosphorus- (Total)	Transparency (Total)
Cadmium (Total)	Iron (Total)	Phosphorus – Soluble Reactive	Turbidity (Total)
Calcium (Dissolved)	Lead (Total)	Potassium (Dissolved)	Zinc (Dissolved)
Calcium (Total)	Light, Attenuation Coefficient	Potassium (Total)	Zinc (Total)
Carbon- Total Organic (Total)	Magnesium (Dissolved)	Purge Volume (Total)	pH (Total)
Chloride (Dissolved)	Magnesium (Total)	Residues- Filterable (TDS) (Dissolved)	
Chloride (Total)	Manganese (Dissolved)	Residues- Nonfilterable (TSS) (Total)	
Chlorophyll (Total)	Manganese (Total)	Residues- Volatile (Total)	
Chlorophyll a (Total)	Molybdenum (Dissolved)	Salinity (Total)	
Chlorophyll b (Total)	Nitrate (N) (Dissolved)	Secchi-horizontal (Total)	
Chlorophyll c (Total)	Nitrate (N) (Total)	Secchi-vertical (Total)	
Cobalt (Dissolved)	Nitrite+Nitrate (N) (Dissolved)	Selenium (Dissolved)	
Coliform Fecal (Total)	Nitrite+Nitrate (N) (Total)	Selenium (Total)	
Coliform Total (Total)	Nitrite (N) (Dissolved)	Silica- Dissolved (Dissolved)	
Color (Dissolved)	Nitrite (N) (Total)	Silica- Dissolved (Total)	
Color (Total)	Nitrogen- Organic (Dissolved)	Sodium (Dissolved)	
Copper (Dissolved)	Nitrogen- Total (Dissolved)	Sodium (Total)	
Copper (Total)	Nitrogen- Total (Total)	Specific Conductance	

Those regressions that met the acceptance criteria are described in the paragraphs below. An example of the results for the Chassahowitzka 1 Spring site is provided in Figure 3-9, all displaying inverse relationships with flow (i.e. constituent concentrations decrease with increasing flows). The greatest number of significant results was observed in the Chassahowitzka 1 Spring (Table 3-3) and the Chassahowitzka Main Spring (Table 3-4) where constituents were regressed against either the daily or 3-day lag average flows.

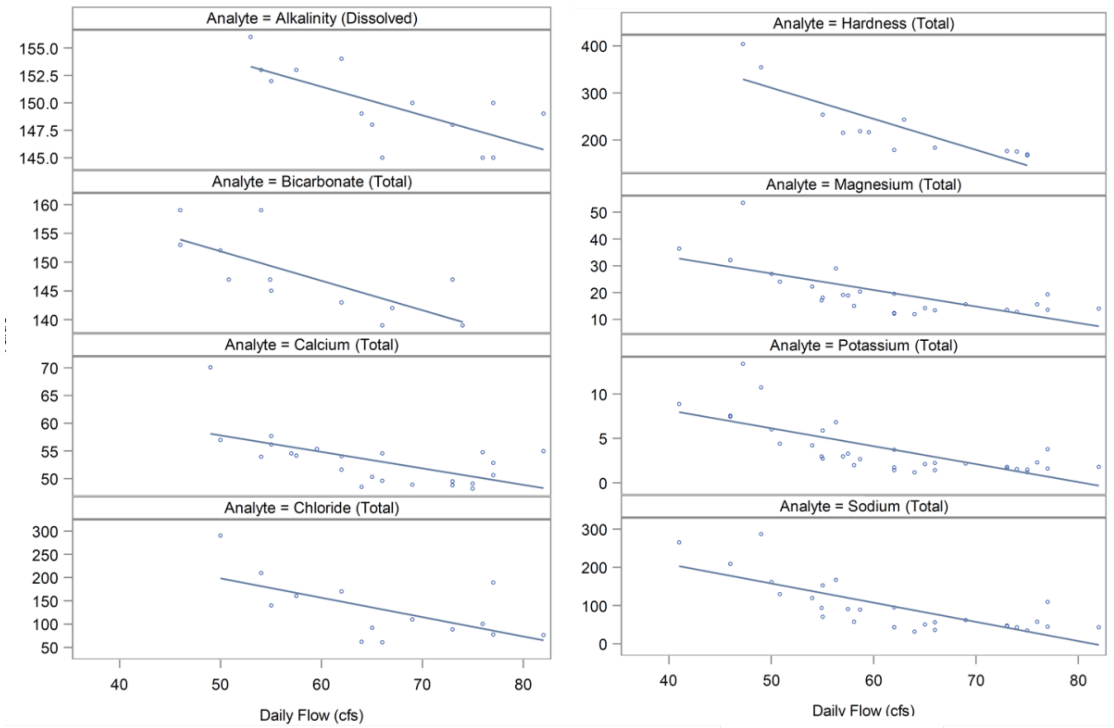


Figure 3-9. Regression relationships between a select group of water quality constituents of interest (all units are in milligrams per liter) and daily flows for Chassahowitzka 1 Spring.

While many of the components of TDS (i.e. bicarbonate, magnesium, and sodium) illustrated a decrease in concentration with increasing flows, there was no evidence that these trends would result in significant harm to the system. The fact that there were statistically significant relationships does not imply that there was an ecologically meaningful interpretation of this result that could aid in reevaluation of the minimum flows. TDS concentrations vary greatly across the spring group and chloride concentrations also widely range, indicating that water quality at the spring group is strongly influenced by the coastal transition zone, even at low tide (SWFWMD 2001).

Table 3-3.	**Significant regression results for Chassahowitzka 1 Spring data for regressions based on same-day flow (no asterisk) or 3-day lagged flow (asterisk).**					
Parameter	Units	Intercept	Slope	DF	R Square	P Value
Alkalinity (Dissolved)	mg/L	165.80	-0.2383	11	0.45	0.0120
Bicarbonate (Total)	mg/L	177.49	-0.5121	10	0.55	0.0059
Calcium (Total)	mg/L	72.65	-0.2971	21	0.36	0.0024
*Chloride (Dissolved)	mg/L	920.63	-11.3559	113	0.36	0.0000
Chloride (Total)	mg/L	406.17	-4.1590	12	0.38	0.0185
Hardness (Total)	mg/L	640.88	-6.5992	11	0.73	0.0002
*Magnesium (Dissolved)	mg/L	69.78	-0.7545	102	0.34	0.0000
Magnesium (Total)	mg/L	57.93	-0.6160	25	0.49	0.0000
*Nitrite (N) (Dissolved)	mg/L	0.01	-0.0001	31	0.33	0.0005
*Potassium (Dissolved)	mg/L	19.31	-0.2372	102	0.34	0.0000
Potassium (Total)	mg/L	16.22	-0.2016	31	0.54	0.0000
*Residues- Filterable (TDS) (Dissolved)	mg/L	1,830.93	-20.4523	110	0.36	0.0000
*Sodium (Dissolved)	mg/L	518.46	-6.3935	102	0.36	0.0000
Sodium (Total)	mg/L	408.95	-5.0233	29	0.61	0.0000
*Specific Conductance (Total)	uS/cm	3,766.01	-42.9613	186	0.37	0.0000
*Sulfate (Dissolved)	mg/L	133.70	-1.5659	113	0.36	0.0000
Sulfate (Total)	mg/L	63.68	-0.5947	12	0.40	0.0149

Table 3-4.	**Significant regression results for Chassahowitzka Main Spring data for regressions based on same-day flow (no asterisk) or 3-day lagged flow (asterisk).**					
Parameter	Units	Intercept	Slope	DF	R Square	P Value
Alkalinity (Dissolved)	mg/L	166.45	-0.2679	12	0.38	0.0183
Calcium (Total)	mg/L	123.52	-0.9669	22	0.56	0.0000
*Chloride (Dissolved)	mg/L	3,047.87	-39.3967	130	0.35	0.0000
Chloride (Total)	mg/L	1,778.16	-21.5393	12	0.68	0.0003
Hardness (Total)	mg/L	1,246.03	-15.3002	20	0.48	0.0003
Magnesium (Dissolved)	mg/L	199.33	-2.3823	103	0.31	0.0000
Magnesium (Total)	mg/L	138.73	-1.6791	39	0.56	0.0000
*Nitrite (N) (Dissolved)	mg/L	0.01	-0.0001	38	0.31	0.0002
*Phosphorus (Dissolved)	mg/L	-0.01	0.0004	11	0.32	0.0431
*Potassium (Dissolved)	mg/L	67.92	-0.8697	103	0.37	0.0000
Potassium (Total)	mg/L	48.48	-0.6377	44	0.57	0.0000
Residues- Filterable (TDS) (Dissolved)	mg/L	5,350.67	-65.2569	118	0.30	0.0000
*Sodium (Dissolved)	mg/L	1,822.54	-23.4989	104	0.39	0.0000
Sodium (Total)	mg/L	1,132.51	-14.6894	41	0.66	0.0000
*Specific Conductance (Total)	uS/cm	11,895.44	-149.2152	186	0.40	0.0000
Strontium (Total)	ug/L	1,376.70	-16.9961	19	0.36	0.0039
*Sulfate (Dissolved)	mg/L	1,317.95	-16.2433	19	0.34	0.0051
Sulfate (Total)	mg/L	253.80	-3.0051	12	0.68	0.0003

In addition to significant relationships at Chassahowitzka 1 Spring and Chassahowitzka Main Spring, some of the same ions also had significant inverse relationships with flow at Betee Jay Spring (Table 3-5) including total dissolved solids (TDS). Total dissolved solids is a measure of chemical constituents dissolved in the groundwater and in west-central Florida, TDS is mostly influenced by the concentrations of the major ions: calcium, bicarbonate, magnesium, sodium, sulfate and chloride and can be used to estimate the relative residence time of ground water in the aquifer which typically increases as the length of groundwater flow paths increase. In coastal areas, TDS is often used to determine the influence of salt water on water quality (SWFWMD 2001).

Table 3-5. Significant regression results for Betee Jay Spring data

Parameter	Units	Intercept	Slope	DF	R Square	P Value
Chloride (Dissolved)	mg/L	603.15	-7.5417	110	0.30	0.0000
Hardness (Total)	mg/L	501.73	-4.3677	11	0.57	0.0030
Magnesium (Total)	mg/L	44.59	-0.4752	20	0.51	0.0002
Nitrate (N) (Total)	mg/L	0.49	-0.0036	14	0.47	0.0034
Potassium (Total)	mg/L	12.54	-0.1744	23	0.54	0.0000
Residues- Filterable (TDS) (Dissolved)	mg/L	1,284.83	-13.4910	108	0.32	0.0000
Sodium (Dissolved)	mg/L	340.00	-4.2019	92	0.31	0.0000
Sodium (Total)	mg/L	303.22	-4.1284	22	0.56	0.0000
Sulfate (Dissolved)	mg/L	95.62	-1.1265	111	0.33	0.0000

Nitrogen enrichment in the Chassahowitzka springs group is an ongoing concern due to the presence of algal mats (filamentous and epiphytic algae) which were linked to increases in nitrogen concentrations. The DEP has adopted a Total Maximum Daily Load (TMDL) for nitrate in the springs and for total nitrogen (TN) in the upper river (Dodson et al. 2014). We reevaluated these relationships and found no significant relationships with any organic or total forms of nitrogen in the Springs dataset. In some cases, regressions with inorganic forms of nitrogen (nitrate, nitrite, and ammonia) resulted in significant relationships to flows but the results were tenuous with low numbers of observations and less than 50% of the total variability explained by the model. The nitrite results are considered especially tenuous since the concentrations tend to be very small and near the detection limits. In addition, the results of the nitrogen regressions were conflicting with respect to the direction of the relationship with flow. For example, the strongest nutrient relationship observed in the Chassahowitzka Spring group in this study was for nitrite (total) at Blue Run Spring with an R^2 value of 66% and $p<0.001$(Table 3-6) but the results suggest a small magnitude positive relationship; increasing concentrations with increasing flow. However, the results of the same analysis for Nitrite dissolved in Blue Run and Ruth Springs suggest an inverse relationship. The only other significant positive relationship observed for a form of nitrogen was for nitrate at Crab Creek Spring (Table 3-6). The nitrite results should be taken with caution since nitrites typically are very close to their detection limits.

Table 3-6. Significant regression results for Baird, Blue Run and Ruth Springs data							
Spring Name	Parameter	Units	Intercept	Slope	DF	R Square	P Value
Baird Spring	Boron (Dissolved)	ug/L	1,659.67	-11.8581	17	0.36	0.0062
Blue Run Spring	Magnesium (Total)	mg/L	347.55	-2.8716	14	0.33	0.0188
Blue Run Spring	Nitrite (N) (Dissolved)	mg/L	0.02	-0.0002	19	0.36	0.0041
Blue Run Spring	Nitrite (N) (Total)	mg/L	-0.01	0.0003	13	0.66	0.0002
*Ruth Spring	Nitrite (N) (Dissolved)	mg/L	0.01	-0.0001	23	0.31	0.0041
Crab Creek Spring	Nitrate (N) (Dissolved)	mg/L	0.24	0.0043	14	0.41	0.0072
Crab Creek Spring	Nitrite (N) (Dissolved)	mg/L	0.01	-0.0001	24	0.32	0.0026

In a technical memorandum by Heyl (2012), included as an appendix to the District's original minimum flow report for the Chassahowitzka River System (Heyl et al. 2012), the relationships between nitrite+nitrate (NO23) nitrogen and flows in spring systems of the Homosassa and Chassahowitzka rivers were examined. The memorandum indicated that flows in the Chassahowitzka have been declining since the 1960s, and that since NO23 monitoring began in 1993, concentrations have been cyclic but with a slight overall positive trend. Since NO23 concentrations have increased over time, the memorandum evaluated whether changes in NO23 were the result of change in flow or time. For the Chassahowitzka data, Heyl noted that once the time effect was accounted for, the relationship with flow was not significant. The trend over time was attributed to inland management practices that increased NO23 loads to the springshed.

In an analysis of the relationships of nitrate to flows in springs in the Suwannee River Water Management District (Upchurch et al. 2008), the objective was to address the question "can management of spring flows be utilized to mitigate nitrate discharging from the springs?". The analytes reported upon included spring discharge, and nitrite+nitrate using data obtained from all of the first, and most of the second, magnitude springs within the Suwannee River Water Management District (n=52). The report concluded that minimum flows cannot be utilized to control nitrate discharging from the springs by promoting high discharge. Data from 50% of the springs showed that nitrate concentrations increased as discharge from the springs increased. Forty-five percent of the remaining springs showed no correlation between discharge and nitrate, and only 5% (2 springs with poor data) had relationships where high discharge was related to lower nitrate concentrations.

3.2.2 Conceptual Model

Despite the existence of many significant water quality relationships with flow, there was no evidence that a conceptual model could be developed that provided a plausible connection between these relationships and the establishment of a minimum flow for the Chassahowitzka River. The relationship between major ions and flow would only be problematic if they were considered contaminants. Instead, many of these constituents are trace nutrients that are valuable for biological growth. In addition, even if the concentrations decrease with flow, the total mass of the constituent may be increasing and that total mass may be a more important driver of response of biota in the receiving water bodies. In summary, there was no evidence that the relationship of any of these constituents with flow would result in significant harm to the receiving waters of the Chassahowitzka River. Therefore, the investigation of relationship between these water quality constituents and flow for the Springs dataset was not pursued further. Plots for all relationships examined for the Springs data are provided in Appendix C.

3.3 RIVER MAINSTEM

Data from the river mainstem includes water quality samples collected from various monitoring networks from just below the headsprings to the mouth of the river. The monitoring programs in the mainstem of the Chassahowitzka River were described in Section 2.1. The following sections describe the exploratory analysis and implementation of the statistical analysis plan for those data.

3.3.1 Exploratory Data Analysis

Linear regression analysis was conducted on data collected in the mainstem of the river in a similar manner to that described above for the Springs resource. The Project Coast monitoring network data (P-529) is the most data rich of the sampling programs. After application of the District acceptance criteria, the only significant constituent related to flow was salinity which was significantly inversely related to flow for 9 of the 10 stations in the network (no results for station 1 met acceptance criteria). This is confirmatory evidence that flows affect water quality in the mainstem of the river. The hydrodynamic model developed for the Chassahowitzka River is considered the best available tool for evaluating the effects of flows on salinity in the mainstem of the river and therefore these relationships were not further pursued. The Coastal Rivers network (P-108), has a reduced sampling frequency relative to P-529 (most constituents only have 25 observations). Major ions including calcium, magnesium, sodium and potassium were significantly related to flows along with total organic carbon. All major ions were inversely related to flows, consistent with the findings for the Springs data, while total organic carbon was positively related to flows. For the UF transect data, only salinity, specific conductivity and alkalinity were significantly linearly related to flows. The hydrodynamic model developed for the Chassahowitzka as part of a separate work effort is considered the best available tool for evaluating the effects of flows on salinity and therefore salinity was not considered further for this analysis within the mainstem of the river. In general, the same inference described for these results in the Springs data can be applied to the majority of the results for the mainstem of the river. That is, there is no evidence that these constituents (other than salinity) would result in significant harm to the system. However, nonlinear patterns in the relationship between flows and chlorophyll were observed for several of the UF transect sites in the upper portion of the mainstem of the river. Given the ecological significance of chlorophyll, these relationships had the greatest potential to serve as valuable water quality indicators to support the reevaluation of minimum flows as described in detail in the following paragraphs. Appendix D details the results of this exploratory analysis with plots against flow for all water quality constituents and details of the statistical output.

Chlorophyll is a green pigment found in all plants that is responsible for the absorption of light, the energy for photosynthesis. In general, the greater the measured chlorophyll concentration, the greater the phytoplankton abundance. Phytoplankton are microalgae found in the water column and form the basis of the estuarine food web. While healthy phytoplankton populations are essential for a healthy estuary, an excess in phytoplankton can have negative impacts. For example, extremely high phytoplankton abundance can lead to an algae bloom, turning water green and preventing sunlight from reaching submerged aquatic vegetation. Factors that control the abundance and distribution of phytoplankton in estuaries are very complex. Several factors can impact phytoplankton populations, including nutrient and carbon availability, solar radiation, predation, and discharge velocity and residence time.

The findings with respect to chlorophyll were of particular interest since it appeared that the effects were somewhat nonlinear at the lowest flows for some of the transects in the upper portion of the river. The UF transect data collected between 1998 and 2011 were deemed the best available information from which to evaluate the effects of flows on chlorophyll distributions in the mainstem of the river. The sampling design is spatially intensive with 20 transect locations within 9 kilometers of the river (Figure 3-10) with all transects sampled on the same date for each event. At each of the first 15 transect locations, three samples were collected along a lateral cross-section of the river while at the five most downstream locations, a single water quality sample was collected (Frazer et al. 2001).

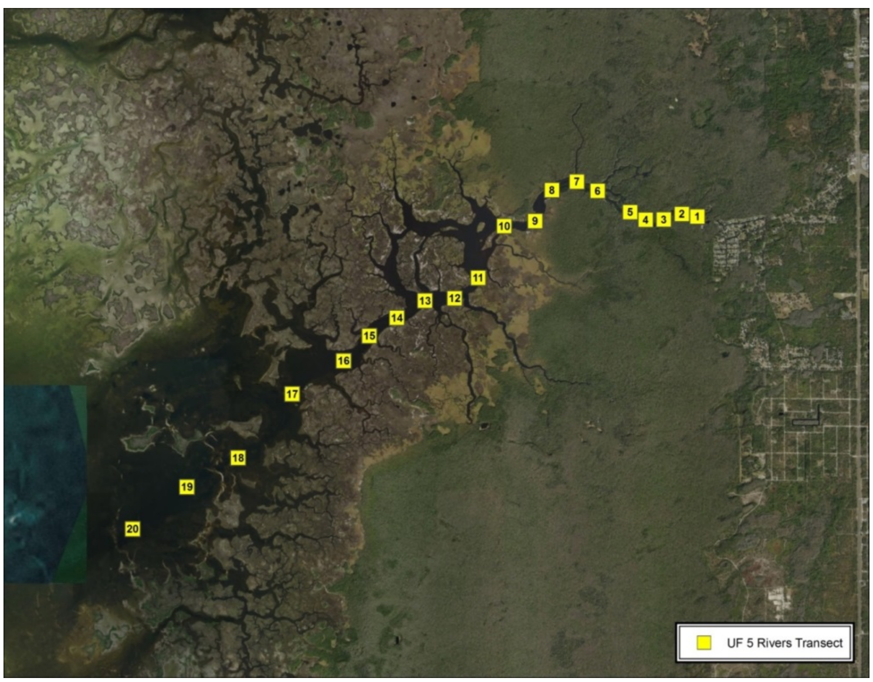


Figure 3-10. River kilometer and transect numbering system for the Chassahowitzka River.

The three samples at each site are not exactly replicates in that they were collected at different lateral positions across the river at the site location to correspond with measurements of macrophyte and other biological measurements across the river. Samples were collected quarterly between 1998 and 2011 with a data gap between 2001 and 2003. Typically, samples were collected in February (Q1), May (Q2), August (Q3), and November (Q4).

There are three major types of chlorophyll pigments, chlorophyll a, b, and c. Chlorophyll a is the pigment in greatest concentrations. Chlorophyll b and c are known as accessory pigments and while they are important to the light harvesting apparatus of plant cells, they are typically found in much lower concentrations in the water column, often at or below laboratory detection limits. Pheophytin is brownish or grey-green compound formed by the degradation of phytoplankton that is also found in the water column. When reporting chlorophyll, most researchers report chlorophyll a, either corrected or uncorrected for pheophytin. The UF data reported uncorrected chlorophyll a concentration for the entire time series of data but only reported chlorophyll corrected after 2005. While pheophytin corrected chlorophyll a is the preferred analyte, using uncorrected chlorophyll a should not significantly change the outcome of chlorophyll-flow relationships. For the purpose of this report, the term chlorophyll implies chlorophyll a uncorrected for pheophytin.

Chlorophyll concentrations were mostly under 10 ug/l in the mainstem of the river with a median value of 3.5 ug/l and a tendency for the highest concentrations between river kilometer 3.8 and 7.4 (Figure 3-11). However, the plots do suggest that the mainstem of the river was susceptible to chlorophyll a concentrations that were in many cases more than three times the median value as seen in the broken y axis plot of Figure 3-11 to highlight the spatial distribution of the majority of the data. Data from the active water quality sampling programs (Figure 3-12) confirm the general spatial distribution of chlorophyll concentrations though the sampling is less spatially intensive. The temporal distribution (Figure 3-13) suggests generally higher concentrations in Quarter 2 (May) and Quarter 3 (August) though concentrations higher than 10 ug/l were found in all quarters. Extreme chlorophyll values (>50ug/L) were also observed though it is possible that these samples were contaminated by other plant material not associated with phytoplankton.

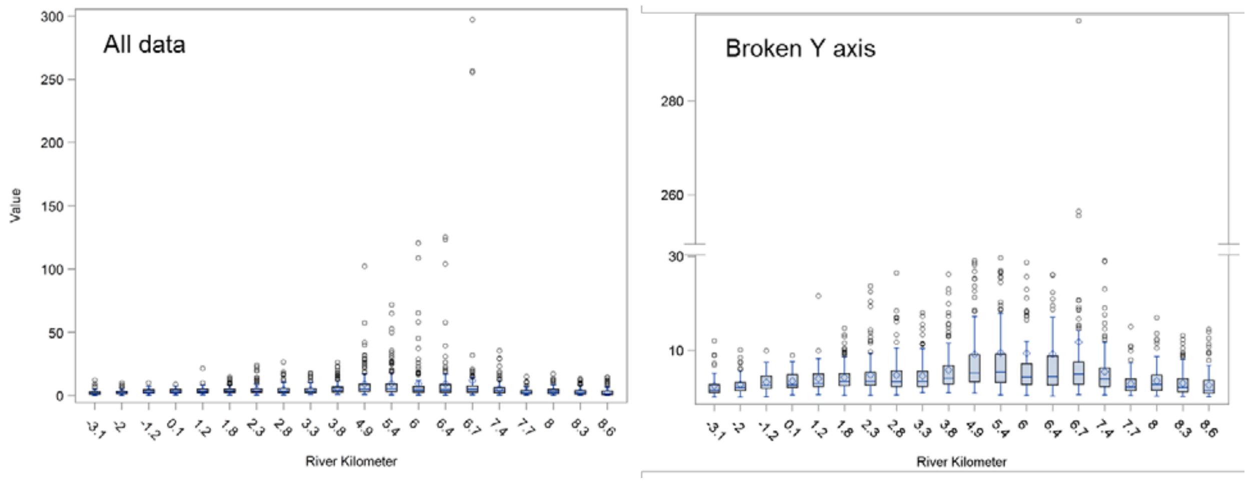


Figure 3-11. **Distribtuon of chlorophyll concentrations from the University of Florida transect data collection effort between 1998 and 2011. All data (left) and broken Y axis (right) to show distribution of data by river kilometer. River kilometer (x axis) is displayed as a discrete value to distribute the box plots more evenly.**

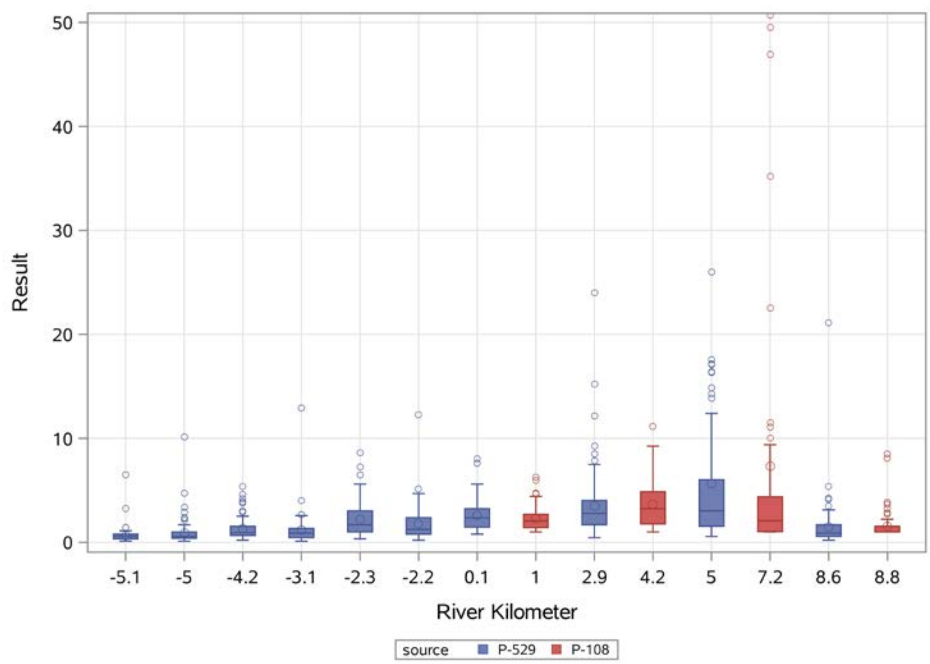


Figure 3-12. **Chlorophyll distribution at fixed locations in the Chassahowitzka River from the active sampling programs in the Chassahowitzka River. River kilometer (x axis) is displayed as a discrete value to space out the box plots.**

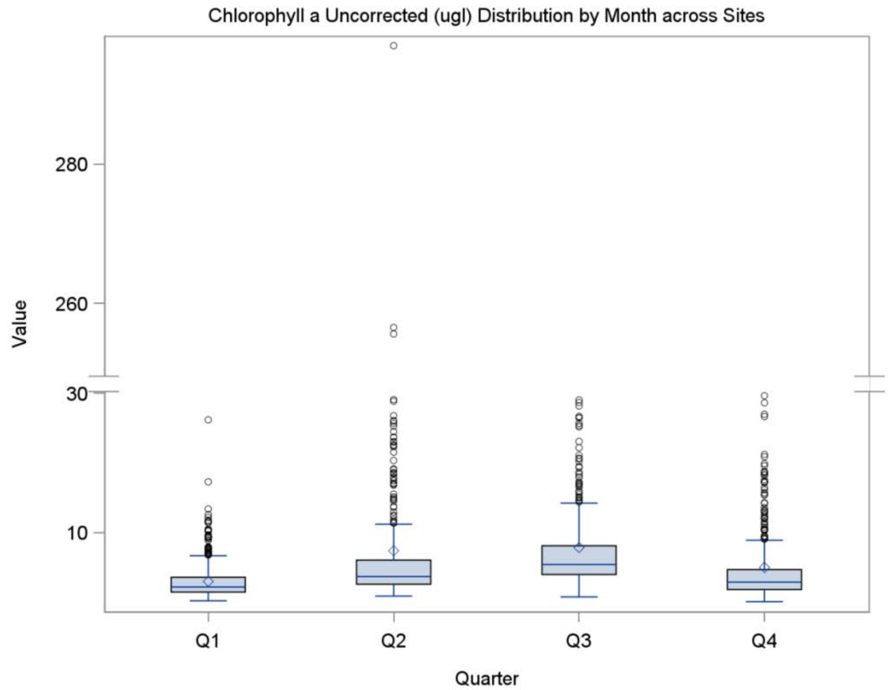


Figure 3-13. Temporal distribution of chlorophyll concentrations based on quarterly sampling from University of Florida transect study in the Chassahowitzka River.

3.3.2 Conceptual Model

Evaluating the effects of flows on water quality in the mainstem of the Chassahowitzka River system necessitated development of a conceptual model for characterizing relationships between flows and water quality. Following development of the conceptual model, the process was enhanced by formulation and use of a statistical model that addressed the conceptual model by testing the hypothesis that the distribution of chlorophyll concentrations in the mainstem of the river was related to flows from the Springs. In particular, a chlorophyll a criterion established by the State for a lower portion of the river (WBID1361:62.303.530 F.A.C.) was chosen to be indicative of a threshold value that, if exceeded, could be associated with the probability of an increased adverse impact. The state water quality standard for chlorophyll in the Chassahowitzka River estuary (WBID 1361:Figure 3-14) is an annual geometric mean (AGM) of 3.9 ug/l (FDEP 2013). The upstream WBID (1348D) does not have a chlorophyll standard. Instead, WBID 1348D has a total nitrogen TMDL, and the springs have a nitrate TMDL. The Chassahowitzka TMDL was based on evidence of excessive filamentous algae growth on the spring and river bottom and not related to phytoplankton abundance. Ironically, the WBID boundary between 1361 and the upstream 1348D WBID bisects the peak spatial distribution of chlorophyll in the river. For this reason, the 3.9 ug/l value was used for the entire river as an indicator of the potential of an adverse condition. However, it should be noted that exceedances in this regard are not directly applicable to inference regarding the declaration of the river as “Impaired” according to state laws.

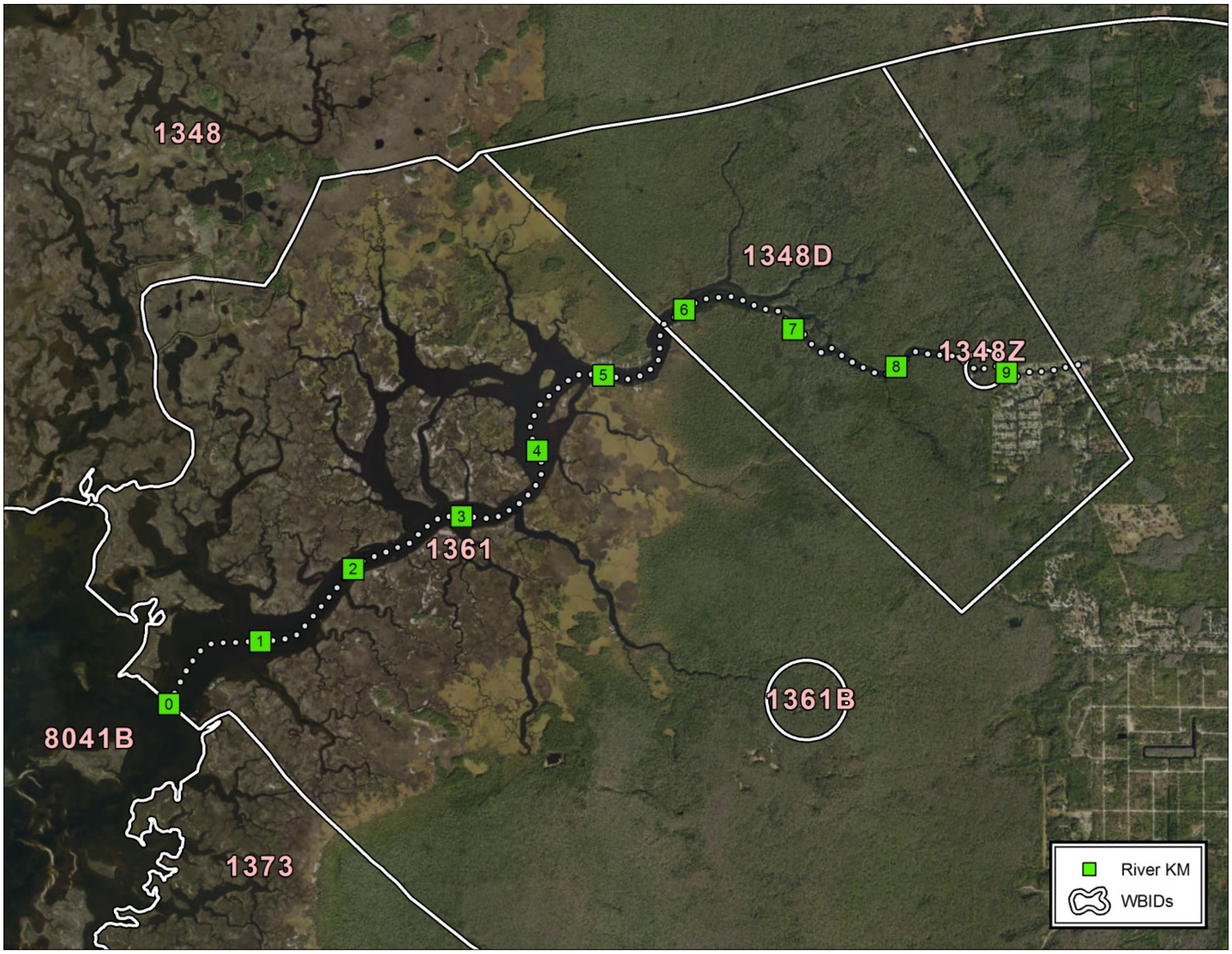


Figure 3-14. Map of Chassahowitzka identifying waterbody identifiers (WBIDS) of relevance within the systems.

While the chlorophyll water quality standard does not apply in a regulatory sense to the uppermost portion of the mainstem of the river (WBID 1348D), we used that value for the entire portion of the river above Rkm 4.9 to represent an indicator that has relevance to an adverse effect for the following reasons. Excessive phytoplankton concentrations, as measured by chlorophyll a, are known to: a) reduce water clarity and limit sunlight available to submerged aquatic vegetation such as the native macrophytes that are considered an indicator of good water quality conditions in the mainstem of the river; b) increase the production of organic material that, upon deposition, can reduce dissolved oxygen concentrations in the river bottom; and c) change the ratio of water column to benthic primary production that is thought to be an important characteristic of historically oligotrophic, spring-fed tidal river systems in Florida (Burghart et al. 2013).

The percent of (3.9 ug/l) threshold exceedances for each sampling location (denoted by its river kilometer) are presented by quarter in Figure 3-15. The plots suggest that the exceedance frequencies change as a function of season and location with increased exceedances in Q2 (May) and an increasing exceedance rate in downstream waters during Q3 (August) relative to Q2. In the winter season, phytoplankton (and therefore chlorophyll concentrations) become limited by photoperiod and temperature and are lower overall.

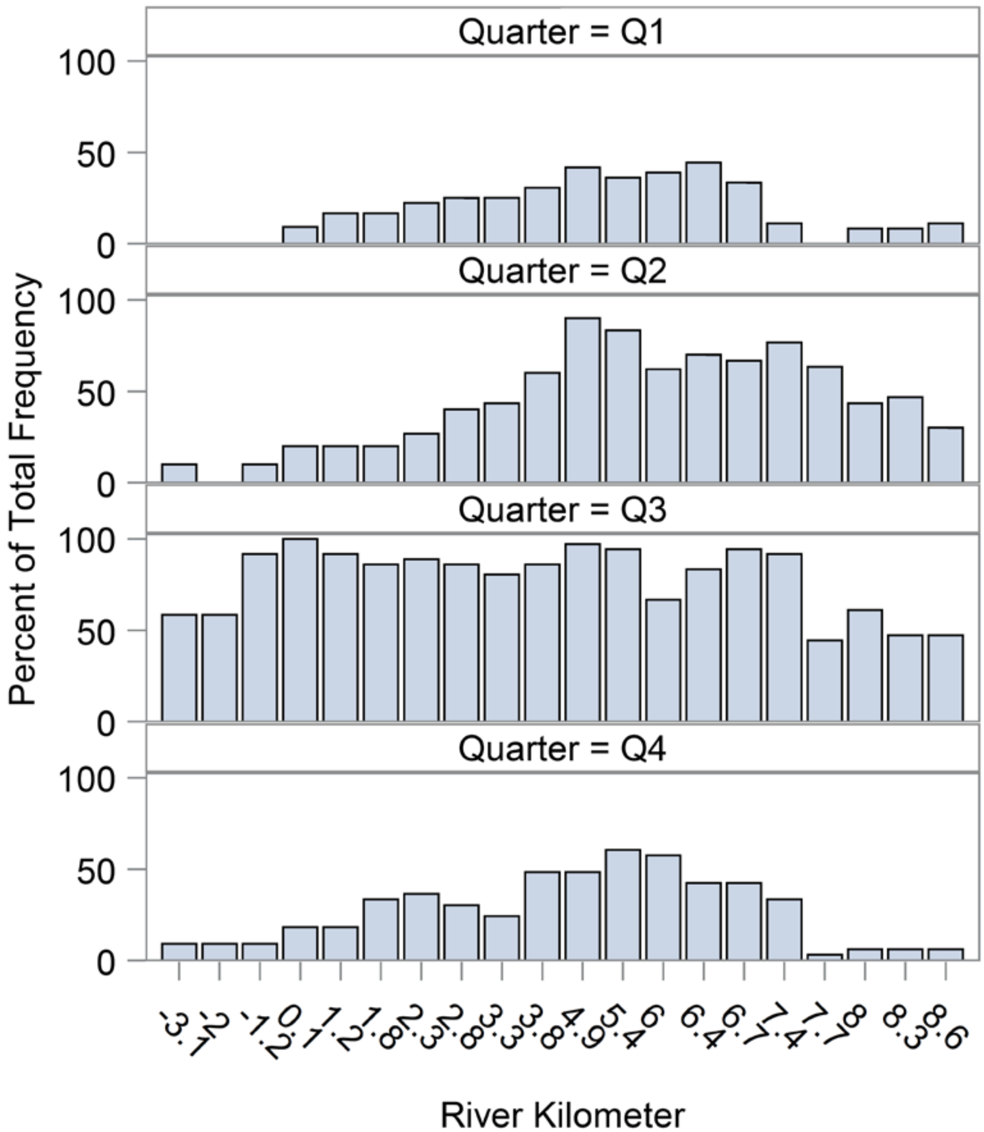


Figure 3-15. Empirical exceedance frequency by station location and by quarter. River kilometer increases towards the headsprings.

Based on the information described above, an analytical pathway was specified to evaluate the effects of flow on the probability of exceeding the state standard for chlorophyll a of 3.9 ug/l. As described in Section 2, the pathway includes:

- the development of a conceptual model;
- development of a hypothesis;
- an analytical approach;
- application of the analytical approach; and
- application of the results to evaluate the effects of flow reductions on the response of interest.

The final bullet point is vital with respect to the reevaluation as it requires that the modeling approach be amenable to conducting a series of simulated flow reduction scenarios. Thereby, the data used in the model must be available for a long term daily timeseries to conduct the flow reduction evaluations.

A conceptual model is presented in Figure 3-16 that illustrates the proposed relationship between spring flows, season and the distribution of chlorophyll in Chassahowitzka River. The model considers the effects of spring flow along with season as principal effects on the chlorophyll distribution, and that the effect of flow and season is location dependent. That is, the effect of flow differs depending on the location in the river and season. The plot of the 3-dimensional locally weighted average (uncorrected) chlorophyll concentrations as estimated using locally weighted scatter plot smoothing (LOESS) is presented in (Figure 3-17). The shape of the plot confirms the complex relationship between location and flow with peak average chlorophyll concentrations below the median of 62 cfs in the most upstream locations of the system. In particular, at the most upstream locations, the lines of increasing flows indicate that chlorophyll peaks around 47 cfs and slowly declines at higher flows.

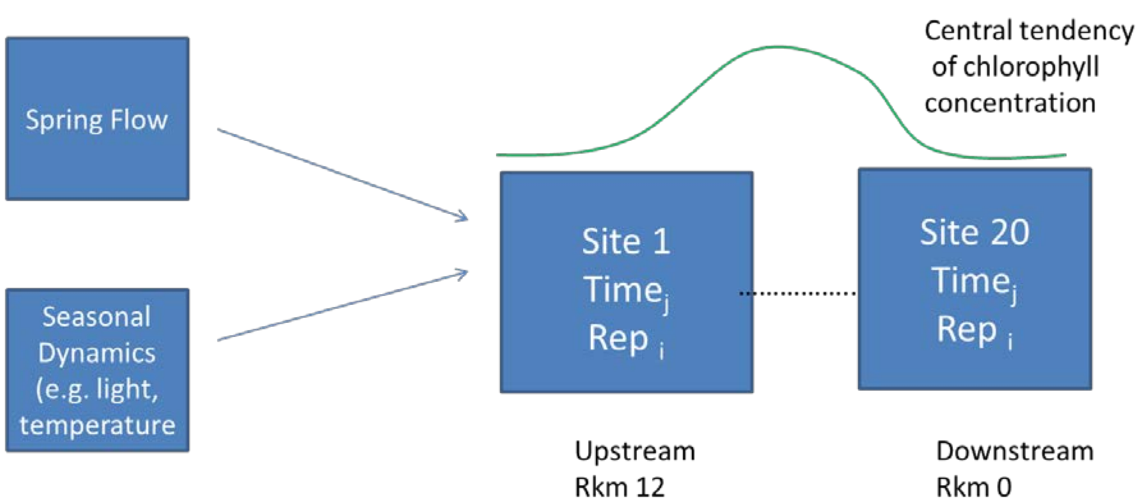


Figure 3-16. Conceptual model of the effects of spring flow and seasonal dynamics on chlorophyll concentrations in the Chassahowitzka River.

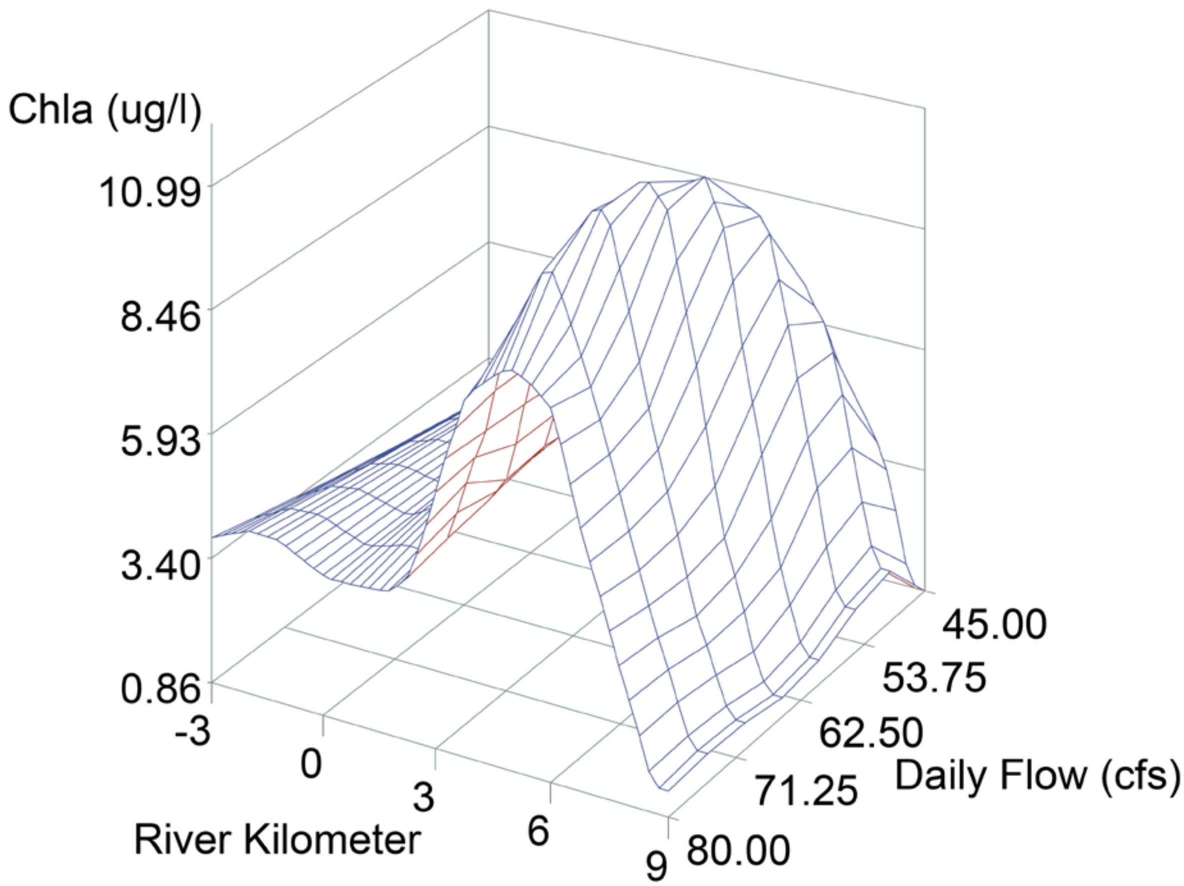


Figure 3-17. **LOESS 3 dimensional smoothed curve of chlorophyll concentrations in the Chassahowitzka River as a function of location and spring flow (using the existing condition flow record) from the UF transect data. River kilometer values increase with increasing distance from the mouth.**

The curve generally corresponds to the results of an analysis of “water age” by the District using the revised hydrodynamic model. Water age is defined as the estimated time it takes for a particle to move downstream of a particular location in the river. There is a rather dramatic increase in water age in the upper portion of the system as flows decrease. An example using the 3-day average flows (to smooth out the influence of tides) for three flow values (30, 50, and 80 cfs) is provided in Figure 3-18. The increase in water age in the upper portion of the system is especially apparent when flows drop below 50 cfs.

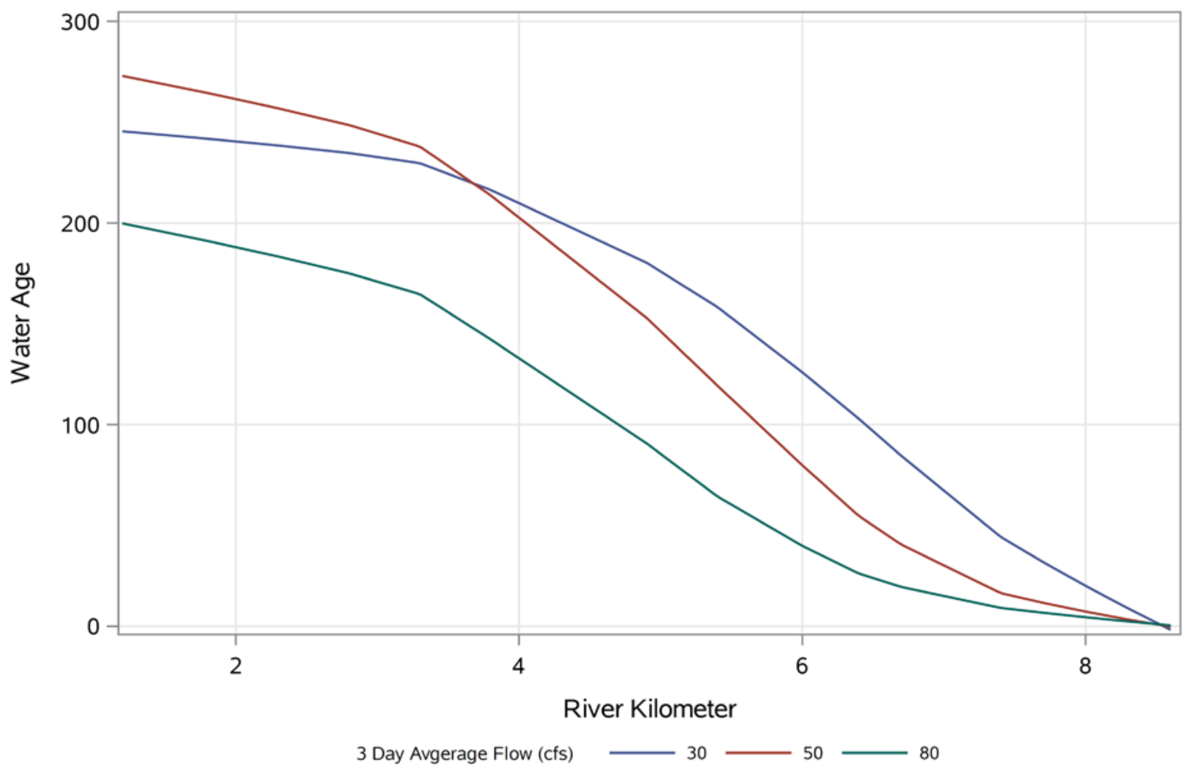


Figure 3-18. **Water age (in hours) curves for three different 3-day average flows.**

3.3.3 Analytical Approach

The conceptual model was then formulated as a statistical model to test the hypothesis that exceedances of the site-specific chlorophyll threshold established for the lower river were related to spring flow for the entire mainstem of the river. The general form of the model is expressed by the equation below as a generalized linear mixed effects model predicting the probability of an exceedance of the chlorophyll standard (a binomial response) as a function of flow and season (i.e. quarter) with interaction terms to allow for the effects of flows on chlorophyll to be location and seasonally dependent. The model is similar to a standard logistic regression model in that it is linear (additive) on the logit (log odds) scale but includes random effects components. Flow and river kilometer were treated as continuous variables in the model while quarter is treated as a categorical variable with Quarter 1 (i.e. Winter) being considered as the reference level. Because the UF data were sampled quarterly (February, May, August, and November), quarters were defined as Q1=Jan-Mar, Q2=Apr-Jun, Q3=Jul-Sep, and Q4=Oct-Dec. A quadratic term for the river kilometer effect was also initially included in an attempt to capture the parabolic curve observed in the empirical data as a function of location in the river. These are the “fixed” effects defining the deterministic component of the model (i.e. the predictive equation).

The “random effects” component of the model allows for specific properties of the sampling design to be incorporated into the analysis in order to appropriately estimate the standard errors associated with the statistical tests used to evaluate significance of the model. This results in what is called “design-based inference” and is important in this analysis to account for the site-specific properties of the sampling locations within the river. Three parameterizations of the random effects component of the statistical model were considered:

Parameterization 1: Random Site Intercepts

Parameterization 2: Random Site and nested Rep within (Site) effect

Parameterization 3: Rep averaged with random site intercepts

The “rep” term refers to the fact that three samples are taken in close proximity to one another at a particular longitudinal location along the river. The equation for the model using Parameterization 1 is given below:

$$E(y) = \log(\frac{p_{(y=1)}}{1-p_{(y=1)}}) = \beta_0 + \beta_{0s} + \beta_1 * \text{flow} + \beta_2 * \text{rkm}_j + \beta_3 * \text{quarter}_k + \beta_4 \ * \text{rkm} * \text{flow} + \beta_5 * \text{quarter}_k * \text{flow} + \beta_6 * \text{Rkm}^2$$

Where:

$\text{logit}(p_{ijk}) =$ probability of exceedence for each sample

S = Site specific properties of location at rkm $_i$

β_0 = Intercept

$\beta_{0s} =$ random intercepts for site $N(\ 0, \sigma^2{}_s)$

$\beta_{1-6} =$ regression coefficients

The random intercepts for the site term is a “variance component” to allow for the fact that each sampled site in the river has a random but quantifiable difference from the overall effect. The benefit of adding the random effects is that it allows the model to capture a variance component associated with variability in sites when estimating the statistical significance of the fixed effects (Zuur et al 2009) and allows for inference at any location within the modeled portion of the system. The difference between Parameterization 1 and Parameterization 2 is an additional term to describe the correlation that exists between the three replicate samples that were taken at the same longitudinal point in the river (although at a different location laterally) on the same date. Parameterization 3 is similar to Parameterization 1 except that the three replicates were first averaged and then the average was used to determine if the value exceeded the site-specific chlorophyll threshold.

The statistical model was implemented using the GLIMMIX procedure in SAS (V9.4: SAS Institute, 2016) using the general principles for model fitting outlined by Zuur et al. (2009) and described as follows. The full fixed effects model was implemented first and the benefit of including the random effects was evaluated using Restricted Maximum Likelihood (REML) and the residual pseudo-likelihood as described in the SAS Stats User’s Guide for the GLIMMIX procedure (SAS Institute: STAT User’s Guide v14.1: 2016). Once the random effects were established, Maximum Likelihood (ML) methods were used to evaluate the benefit of the fixed effects model terms using the goodness of fit evaluated by changes in likelihood ratio and the

Akaike Information Criteria (AICC) statistics. Individual terms were dropped from the model if they did not contribute improvement to the model fit as evaluated based on a reduction in either the log likelihood or AICC. Once the fixed effects were established for the final model, the final model was run and parameter estimates reported using REML.

3.3.4 Analytical Results

The three candidate model parameterizations for the random effects were considered using the full model fixed effects. Because the random effects parameterization can change the fixed effects estimates, the fixed effects parameter estimates were evaluated for all models as well. Since interaction terms are present in the models this can alter the results of the significance tests of the main effects if the interaction results in a cross over effect. For example, in the models below, the main effect for flow is not statistically significant as a Type 3 test because the location effect is set to the mean and at the mean the effect crosses over (Table 3-7). However, the solutions table (Table 3-8) displays the statistical test results evaluated at the reference level for the categorical effects and therefore provides estimates of the significance at specific levels of the other variables and demonstrate that under specific conditions, the effects of flow are significant. Since the interaction terms were highly significant, the main effects were retained in the model. The inference from the results suggests that the effects of flow are dependent on location in the river and season.

The random intercepts term for site significantly improved the model fit over the fixed effect model based on the likelihood ratio test results (LRT under Fit Statistics: Table 3-7). The dispersion statistic is also used as a test of model goodness of fit. An over-dispersed model can lead to improper estimates of the standard errors and inflated Type I error while a dispersion parameter below 1 is generally conservative with respect to the statistical significance of the parameter estimates (Burnham, K. P. and D. R. Anderson. 2002).

Table 3-7. Results of Type 3 tests for fixed effects for the three mixed effects model parameterizations evaluated to predict chlorophyll exceedances in the Chassahowitzka River.

	Parameterization 1 (P Value)	Parameterization 2 (P Value)	Parameterization 3 (P value)
Fixed Effects			
Flow	0.2161	0.3875	0.2892
Rkm	<.0001	<.0001	<.0001
Flow*Rkm	0.0007	0.0112	0.0055
Quarter	<.0001	<.0001	<.0001
Flow*Quarter	<.0001	<.0001	<.0001
Rkm^2	<.0001	<.0001	<.0001
Fit Statistics			
Random Effects : LRT	<0.001	<0.001	<0.001
Dispersion (Chisq / DF)	0.97	0.54	0.98

Table 3-8.	**Solutions table for fixed effects for the three mixed effects model parameterizations evaluated to predict chlorophyll exceedances in the Chassahowitzka River.**					
	Parameterization 1		Parameterization 2		Parameterization 3	
	Estimate	P Value	Estimate	P Value	Estimate	P Value
Fixed Effects						
Intercept	-8.9634	<.0001	-10.1523	<.0001	-8.6462	0.0004
Flow	0.1179	<.0001	0.1329	<.0001	0.1102	0.0005
River Kilometer	0.958	<.0001	1.0693	<.0001	1.051	<.0001
Flow*River Kilometer	-0.00793	0.0007	-0.00866	0.0112	-0.0088	0.0055
Quarter 2	14.0614	<.0001	16.3239	<.0001	12.9666	<.0001
Quarter 3	9.2264	<.0001	10.5851	<.0001	9.6258	<.0001
Quarter 4	5.5209	<.0001	6.1723	0.0018	4.5266	0.0273
Quarter 1	0		0		0	
Flow*Quarter 2	-0.2159	<.0001	-0.2512	<.0001	-0.1958	<.0001
Flow*Quarter 3	-0.1043	<.0001	-0.1181	0.0003	-0.1035	0.0029
Flow*Quarter 4	-0.08302	<.0001	-0.09256	0.0036	-0.06457	0.0494
Flow*Quarter 1	0		0		0	
Rkm^2	-0.06124	<.0001	-0.07042	<.0001	-0.06435	<.0001

The area under the receiver operator curve (ROC: Hosmer and Lemeshow 2000) is another metric used to evaluate the model fit. The ROC curves plot the sensitivity (defined as correctly predicting an exceedance when one is observed in the empirical data) against 1-specificity (defined as correctly predicting a non- exceedance when a non-exceedance is observed in the empirical data). A ROC curve that is high into the upper left-hand corner of the plot is most preferred because it has both high sensitivity and high specificity. Parameterization 3 had the largest area under the ROC curve (0.83: Figure 3-19) which is considered excellent discrimination according to Hosmer and Lemeshow (2000). Based on the fact that the Parameterization 3 dispersion statistic is close to 1 and that parameterization had the highest ROC value, in addition to results of internal discussions with the project team, Parameterization 3, the “rep averaged” model, was considered the most appropriate representation of the system under study for evaluating the effects of flows on the probability of a chlorophyll exceedance.

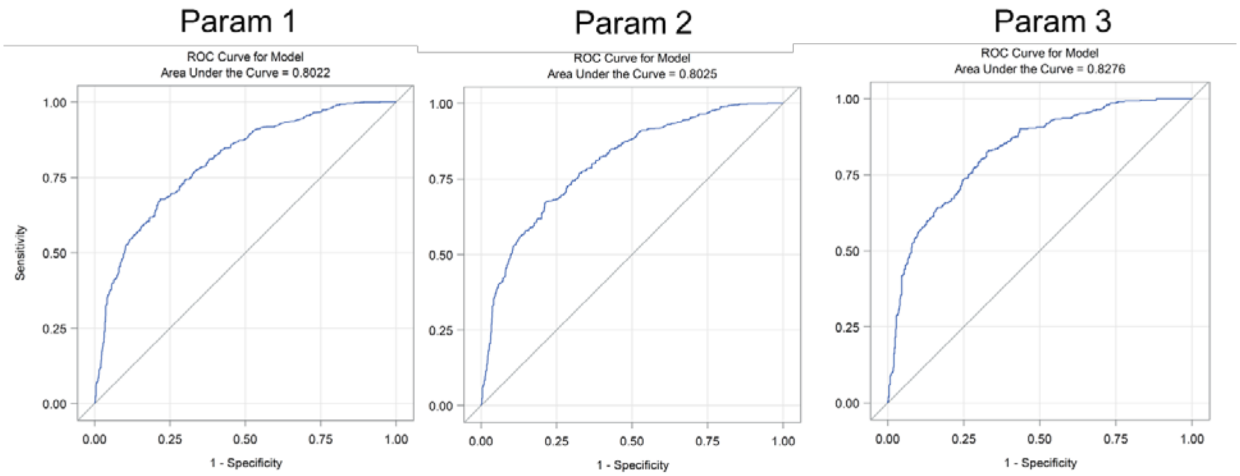


Figure 3-19. Receiver operator curves based on the fixed effects for the three generalized mixed effects models considered (note Param = Parameterization).

Once the parameterization for the random effects was established, the final fixed effects were evaluated by sequentially eliminating effects from the full model, beginning with the interaction terms, and evaluating the effects on AICC. The full model, which included all main effects and interactions, was the best of the candidate models for describing the fixed effects (Table 3-9).

Table 3-9. Comparison of Akaike Information Criteria for nested models with various fixed effects. For AICC smaller numbers represent improved model fit. Models were fit using maximum likelihood.

Model Parameterization	AICC
Flow rkm quarter	2440.62
Flow rkm Flow*rkm quarter	2431.23
Flow rkm Flow*rkm quarter Flow*quarter	2368.47
Flow rkm Flow*rkm quarter Flow*quarter rkm_sq	2352.94

Once the final model was selected, diagnostic plots and prediction curves were generated to evaluate the model fit across a range of conditions. A summary plot of the predicted probability of occurrence as an effect of season and river kilometer is provided in Figure 3-20a. The results suggest an increasing probability of occurrence with movement upstream in the river and higher overall probability of exceedance in Q3. These predictions are based on the linear component of river kilometer as the quadratic effect is held constant. Quarters 1 and 4 had lower overall probability of exceedance and multiple comparisons of the quarters term suggested that Q2 and Q4 were not statistically different in terms of their general relationship with river kilometer as evidenced in Figure 3-20b, which provides multiple comparison test-adjusted differences for the least squares mean estimates for the Quarter main effect (SAS Institute, Inc. 2016). However, the interaction terms again complicate this result. For instance, when evaluating the effects by quarter for four different flow scenarios (Figure 3-21), one can see that at the lowest flows, Q2 had the highest probability curve which quickly diminished as flows increased and had the lowest probability curves at the highest flows, implicating residence times as a potential factor limiting chlorophyll exceedances in the spring season. Quarter 3 (August) tended to be rather resistant to changes in flows.

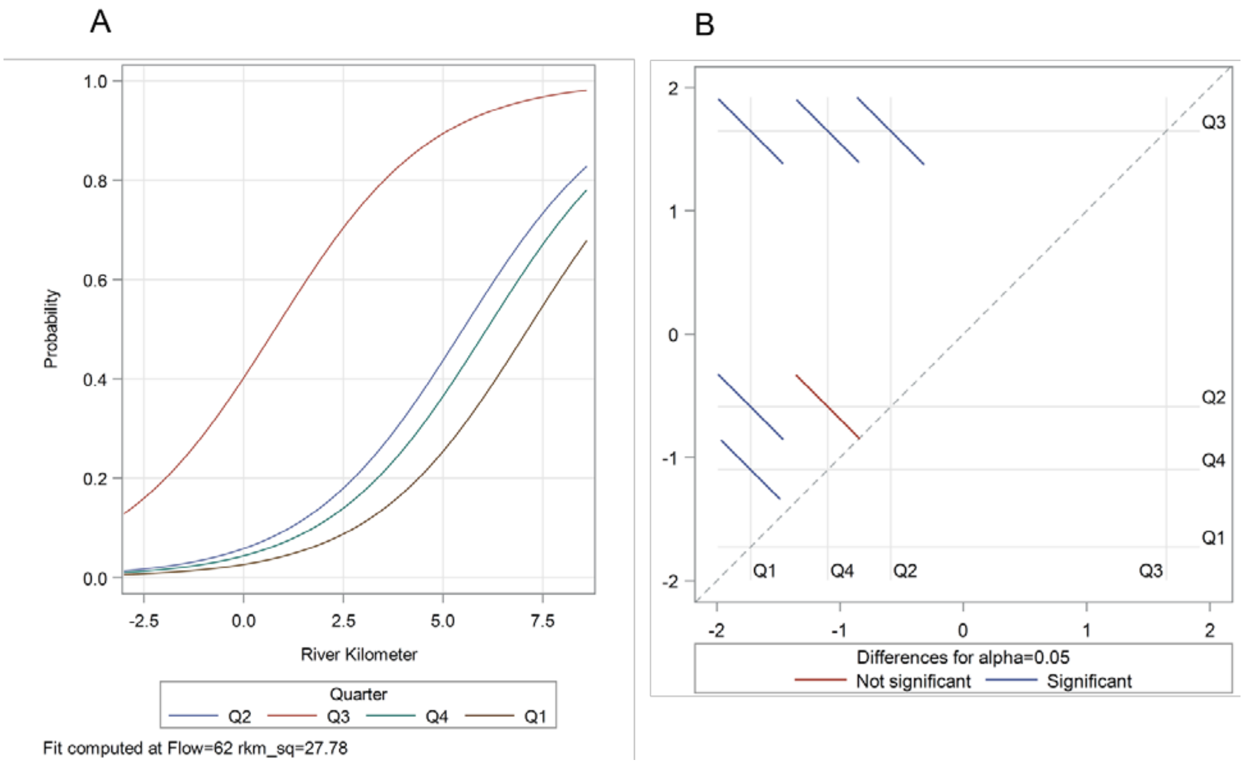


Figure 3-20. Predicted probabilities for each quarter as a function of river kilometer at a fixed daily flow of 62 cfs (A) and a diffogram of the multiple comparisons test to assess differences between quarters (B).

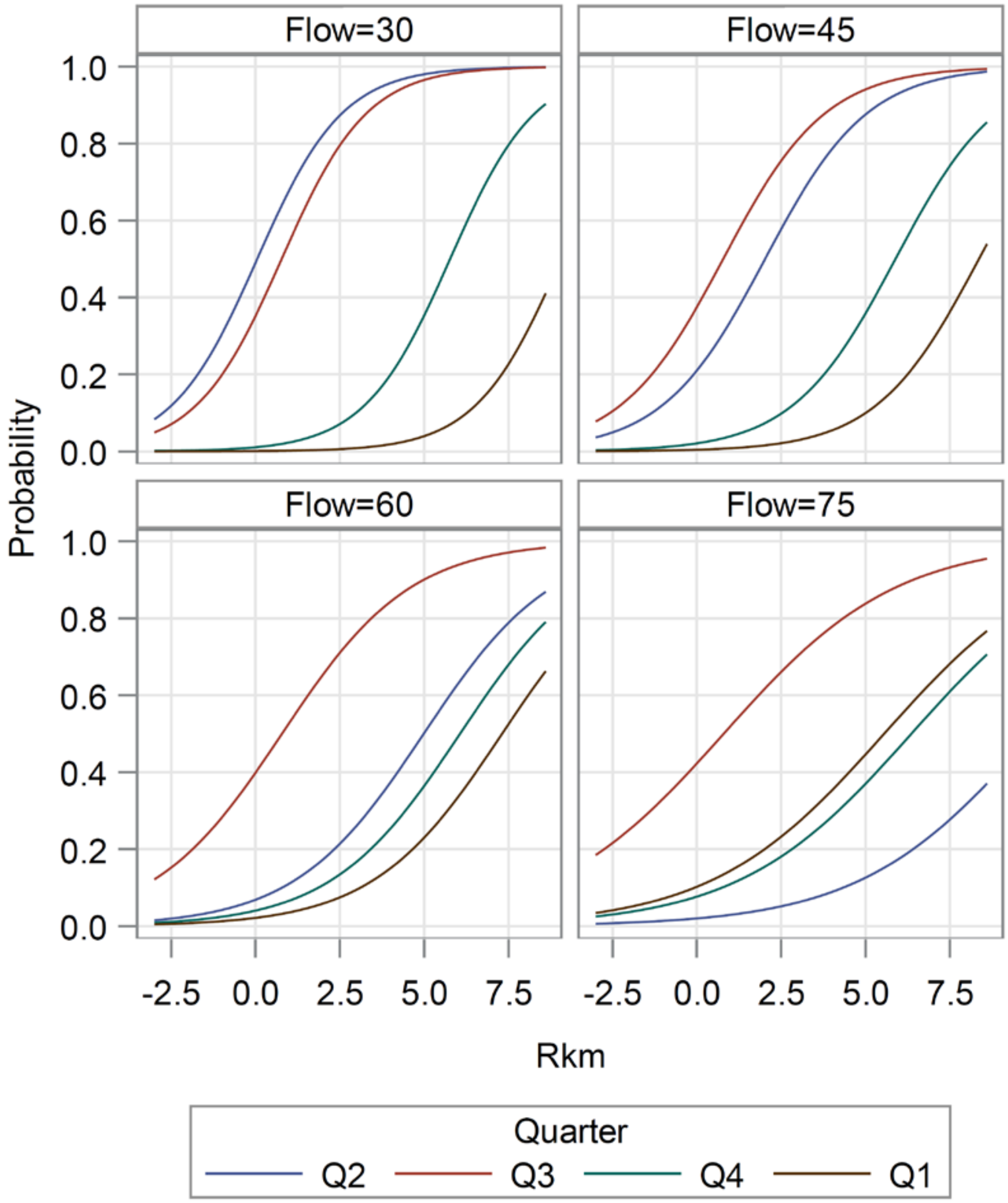


Figure 3-21. **Predicted probability of exceedance as a function of river kilometer by quarter for four fixed flow values in the Chassahowitzka River.**

In Figure 3-22, the relationship between flow and the predicted probability of exceedance by quarter is plotted at 4 different locations in the river (Rkm -3, 0, 3, 6). The effects of flow are clearly most visible upstream in Quarter 2 where decreased flows are predicted to increase the probability of an exceedance. The effects of the interaction terms are also clearly evident as the slope varies significantly by quarter and by location in the estuary. Based on these results, it is evident that, with respect to establishing minimum flows for the system, the effects of flow reductions will be most relevant in Q2, precisely when the flows tend to be at or near their annual minima.

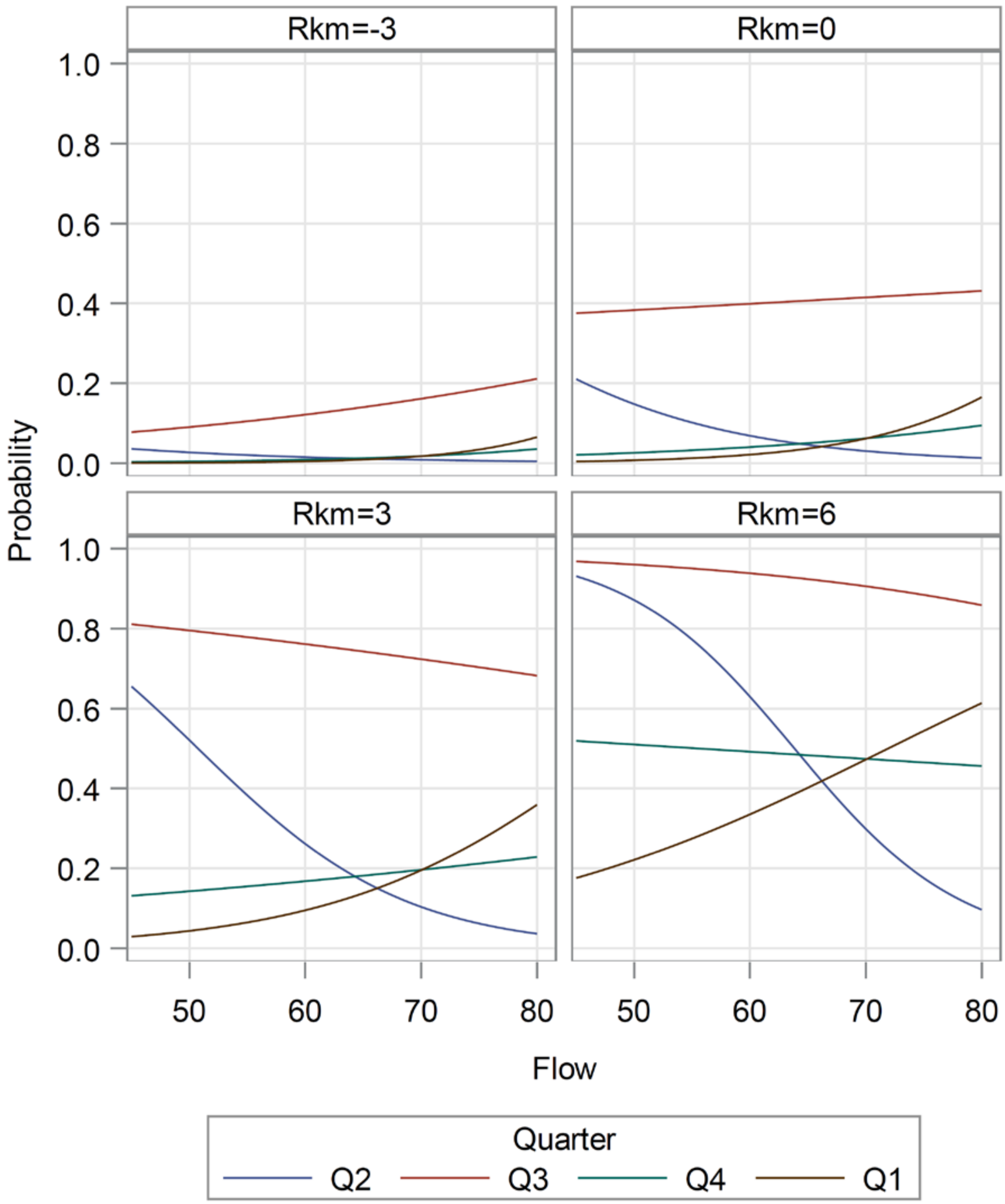


Figure 3-22. Predicted probability curves as a function of flow by quarter at 4 different locations in the Chassahowitzka River System.

This model was subsequently used to evaluate the effects of flow reductions on changes in chlorophyll a exceedance probabilities for the upper portion of the river. The results of that evaluation are described in Chapter 4.

3.4 ESTUARY

The goal of the estuary analysis was to assess relationships between flows and water quality constituents of interest for sites located outside the hydrodynamic model domain. Sites for the mainstem of the river were described by the analysis above. The "Estuary" sites include four Project COAST (P529) sampling stations, as well as three transects from the previously completed UF 5 Rivers Study (Figure 3-23). Two Project Coast sites (9, 10) were deemed too far removed from the mouth of the river to be useful for this evaluation.

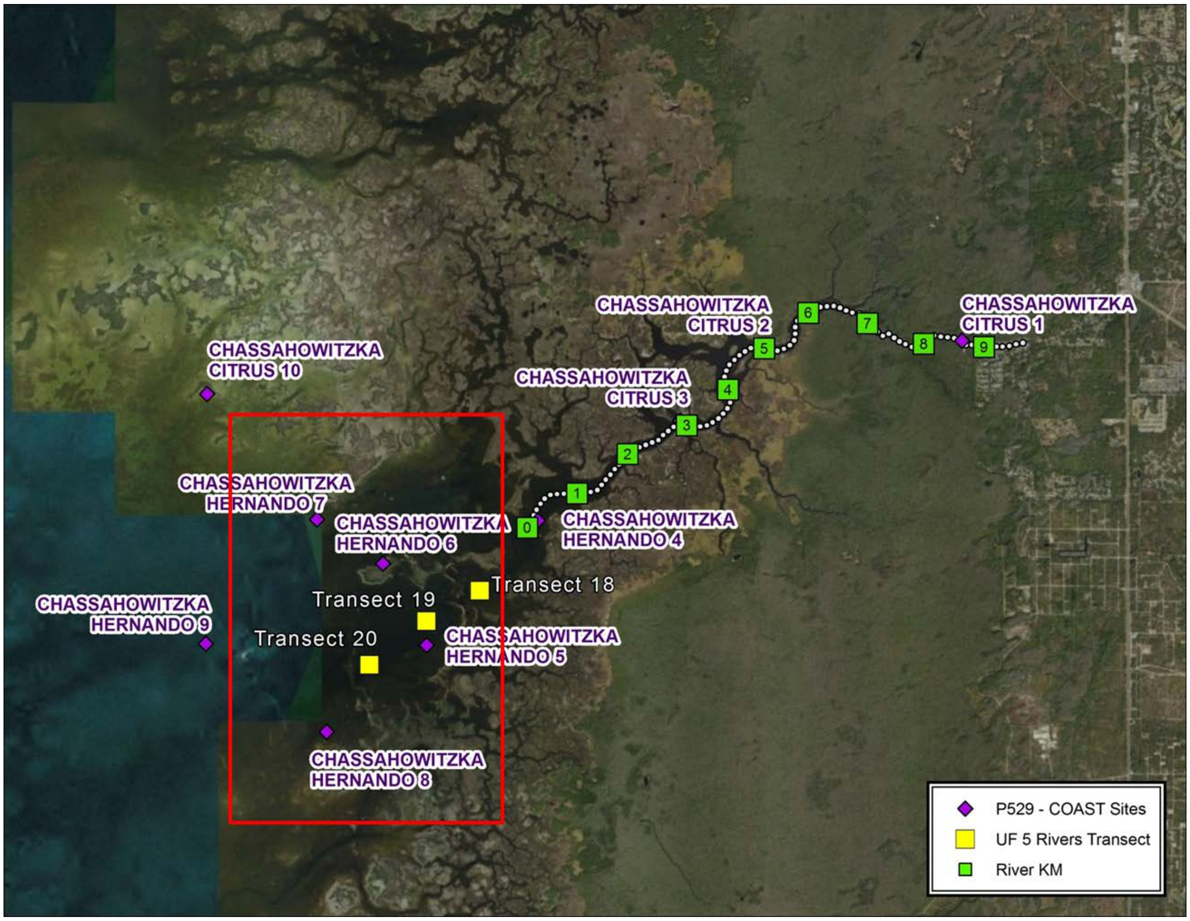


Figure 3-23. Sampling areas in the estuary outside of the hydrodynamic model domain considered for analysis for the "Estuary" (highlighted with red rectangle).

The same regression process used for the Springs data was applied to the Estuary sites. Outliers indicated by robust regression and data points with qualifiers indicating unreliable data were removed from the analyses. The parameters listed in Table 3-10 were tested for significant relationships with flow. After application of the linear regression acceptance criteria adopted for this project, salinity was the principal water quality constituent affected by springs flows which was consistently inversely related to flow in all stations except Hernando 6 . Plots for all stations and parameters analyzed are provided in Appendix E.

Table 3-10. List of water quality constituents evaluated for linear relationships with flow.

Parameter	Sampling Program(s)
Alkalinity (Total)	UF 5 Rivers
Ammonia	UF 5 Rivers
Chlorophyll a (corrected)	UF 5 Rivers
Chlorophyll a (uncorrected)	UF 5 Rivers
Chlorophyll a (Total)	P-529
Chlorophyll (Total)	P-529
Color	UF 5 Rivers, P-529
Dissolved Oxygen	UF 5 Rivers, P-529
Light Attenuation Coefficient	P-529
Nitrogen – Total	UF 5 Rivers, P-529
Nitrate	UF 5 Rivers
pH	UF 5 Rivers, P-529
Phosphorus – Total	UF 5 Rivers, P-529
Salinity	UF 5 Rivers, P-529
Secchi-vertical	P-529
Specific Conductivity	UF 5 Rivers
SRP	UF 5 Rivers
Temperature	UF 5 Rivers, P-529

In a study by Yobbi and Knochenmus (1988), the location of 25-ppt salinity isohaline in the Chassahowitzka River had a range in movement that was more than three times as great as the range in movement of the upstream extent of the zone of saltwater mixing. The authors also report that the 18-ppt salinity isohaline can be expected to be found downstream of the river mouth about 90 percent of the days; and the 25-ppt salinity can be expected to be found 3 miles or more outside the river mouth about 90 percent of days. The authors examined how a 15% reduction in spring flow (via groundwater pumping) would impact upstream movement of low salinity in the river but unfortunately did not report on the movements of higher salinities in the estuary as a result of flow reduction.

Table 3-11. Significant regression results for estuary data.						
Site Name	Parameter	Intercept	Slope	DF	R Square	P Value
CHASSAHOWITZKA HERNANDO 5	Salinity (Total)	35.4046	-0.3257	215	0.32	0.0000
CHASSAHOWITZKA HERNANDO 7	Salinity (Total)	37.9703	-0.3207	214	0.30	0.0000
CHASSAHOWITZKA HERNANDO 8	Salinity (Total)	37.9498	-0.3221	215	0.32	0.0000
Transect 18 - 3	Salinity (Total)	37.2142	-0.3780	43	0.35	0.0000
Transect 19 - 3	Salinity (Total)	39.9059	-0.3981	43	0.34	0.0000
Transect 20 - 3	Salinity (Total)	40.2606	-0.3735	43	0.32	0.0001

As a general application of the regression results for each unit change in the 3-day flow, salinity at the estuarine sites would change by between ~ 0.3 psu (at Chassahowitzka Hernando 7) to ~0.4 psu (at the Transect 19 location). Using the median Chassahowitzka flow of 62.5 cfs, a 15 % reduction in flow would result in between a 3.0-3.7 psu increase in salinity depending on station. However, there is currently no criterion value from which to identify significant harm to the estuary as a function of changes in salinity in the open estuary. Given that the estuarine area examined in the current analyses is so far removed from the springs flows, and is affected by direct rainfall, surface flows from coastal zone runoff, and wetland storage, there appears to be little utility in directly using these regressions to support the establishment of minimum flow criteria for the springs flows to the system.

3.5 CONTINUOUS WATER QUALITY MONITORING STATIONS

There are three continuous water quality monitoring sites in the Chassahowitzka River (Figure 3-24). The most upstream site collected data on salinity, temperature, dissolved oxygen, pH and nitrite+nitrate, while the downstream locations (near Mouth and Near USGS) include several additional fluorescence-based estimates including chlorophyll, fluorescence dissolved organic matter (FDOM), light spectra, nitrate, and turbidity. These latter parameters are of particular interest since the salinity and temperature parameters are modeled by the hydrodynamic model developed separately for reevaluation of the minimum flows.

Exploratory data analysis consisted of evaluating the relative contribution of within- and between-day variability on the distribution of data available within each quarter for the period of record between January 2017 and March 2018. Much of these data are still provisional and therefore the results are meant only for exploratory analysis. The coefficient of variation (i.e. CV= the standard deviation divided by the mean) was used to quantify the variability around the expected value. For the within-day variation, the result was a distribution of CV values. For the between-day variability, the average value for each date was calculated and then the CV of the daily average values was calculated resulting in a single CV value to represent the between-day variation. These results were then overlaid to evaluate the relative difference between the within- and between-day variability.

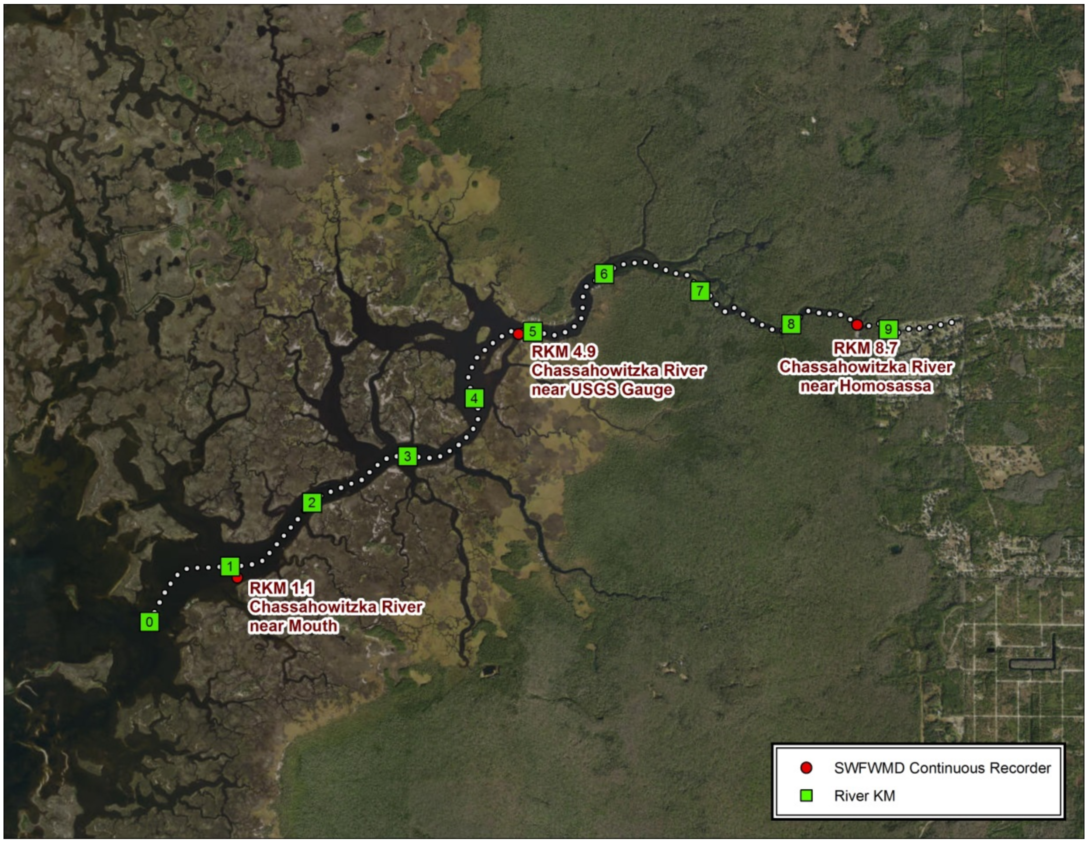


Figure 3-24. Location of continuous recorder gages in the Chassahowitzka River.

An example application of this approach is provided for the USGS gage at Chassahowitzka River near Homosassa (USGS 02310650) that includes discharge measurements from the main spring as well as continuous NO23 data. These are 15-minute data that were averaged by hour. A plot of the timeseries for discharge and nitrite+nitrate is provided in Figure 3-25. Missing data are evident in both timeseries but more prevalent in the NO23 data. The CV plots are provided in Figure 3-26 and suggest that within-day variability may be substantial relative to the between-day variability within a quarter for both flow and NO23. The CV of the daily discharge was consistently below the within-day CV in all quarters, suggesting that while the daily average flows are seasonally consistent, tidal action has considerable effect on the discharge estimates within a day. The NO23 variations are similar and quite small both within- and between-day with CV values typically less than 10% of the mean. It should be noted that the within-day CV is calculated based on the standard deviation of 24 observations while the between-day variability is calculated from 90 daily observations and, as such, the standard deviation for the within-day CVs may not represent an asymptotic value in all cases.

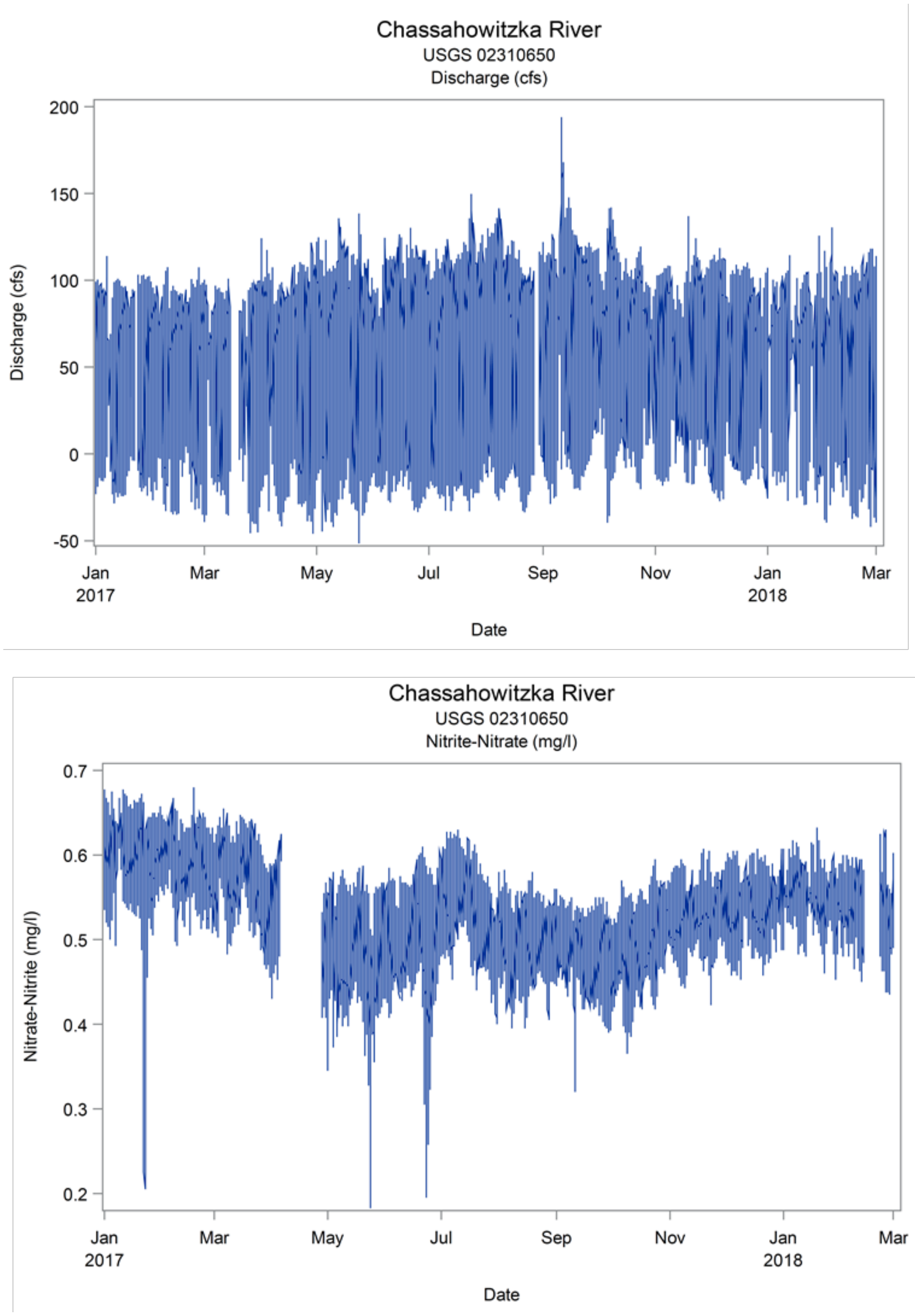


Figure 3-25. Timeseries of Discharge (top) and nitrite+nitrate (bottom) from the continuous recorder gage Chassahowitzka River near Homosassa (USGS 02310650).

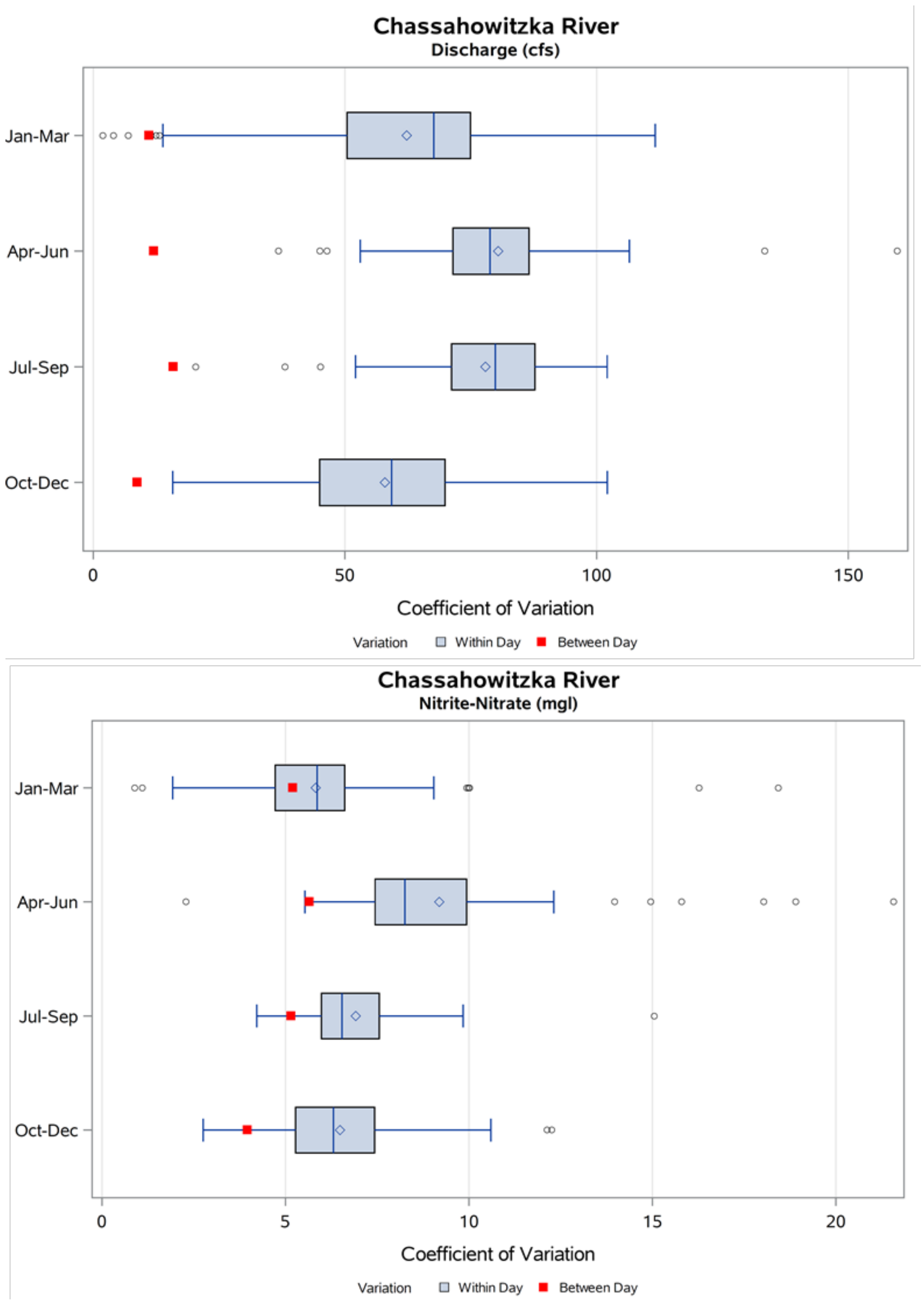


Figure 3-26. Coefficient of variation plots for discharge (top) and nitrite+nitrate (bottom) from the continuous recorder gage Chassahowitzka River near Homosassa (USGS 02310650).

A scatter plot of the relationship between discharge and NO23 for this gage is provided in Figure 3-27 and suggests no correspondence between flows and NO23; the data are centered around 0.52 mg/l irrespective of flows from the main spring.

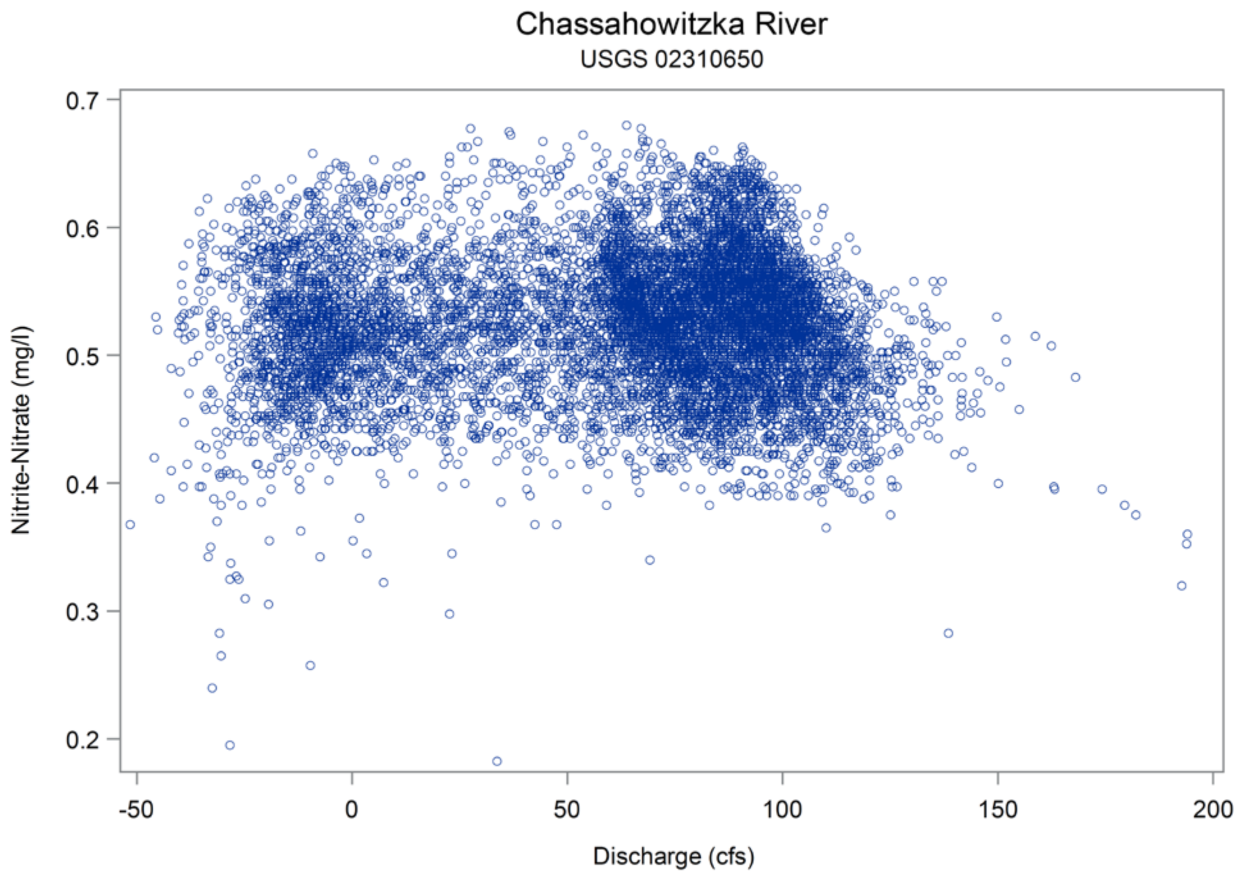


Figure 3-27. Relationship between discharge and NO23 from the continuous recorder gage Chassahowitzka River near Homosassa (USGS 02310650).

The hourly distribution of NO23 is also quite consistent, though there seems to be a tendency for NO23 concentrations to be slightly lower in the early afternoon (Figure 3-28).

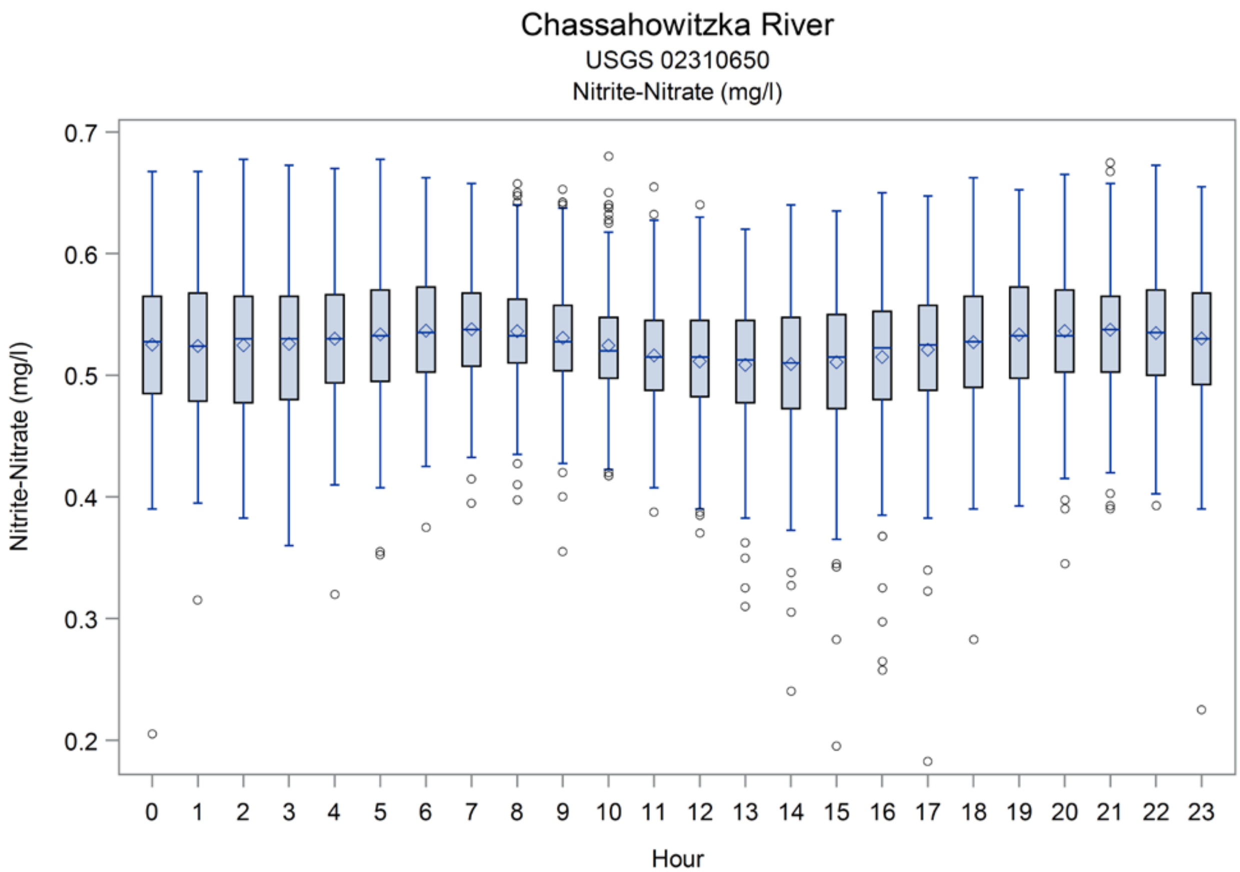
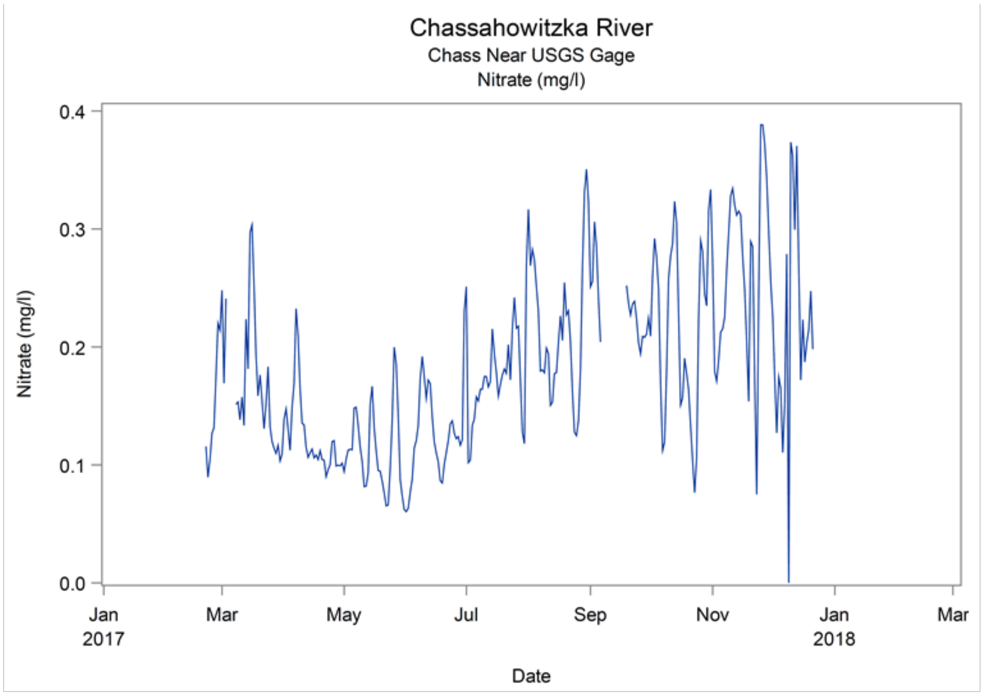


Figure 3-28. Distribution of hourly nitrite+nitrate concentrations (mg/l) from the continuous recorder gage Chassahowitzka River near Homosassa (USGS 02310650).

The two downstream locations included more parameters but also more missing data. Timeseries for NO23 at the two downstream continuous recorders is provided in Figure 3-29 and suggest considerable seasonal differences when data are available. The CV plots (Figure 3-30) confirm that the between-day CV tends to be larger than the within-day distribution of CVs. Note that the x-axis scales are different in these plots as well.



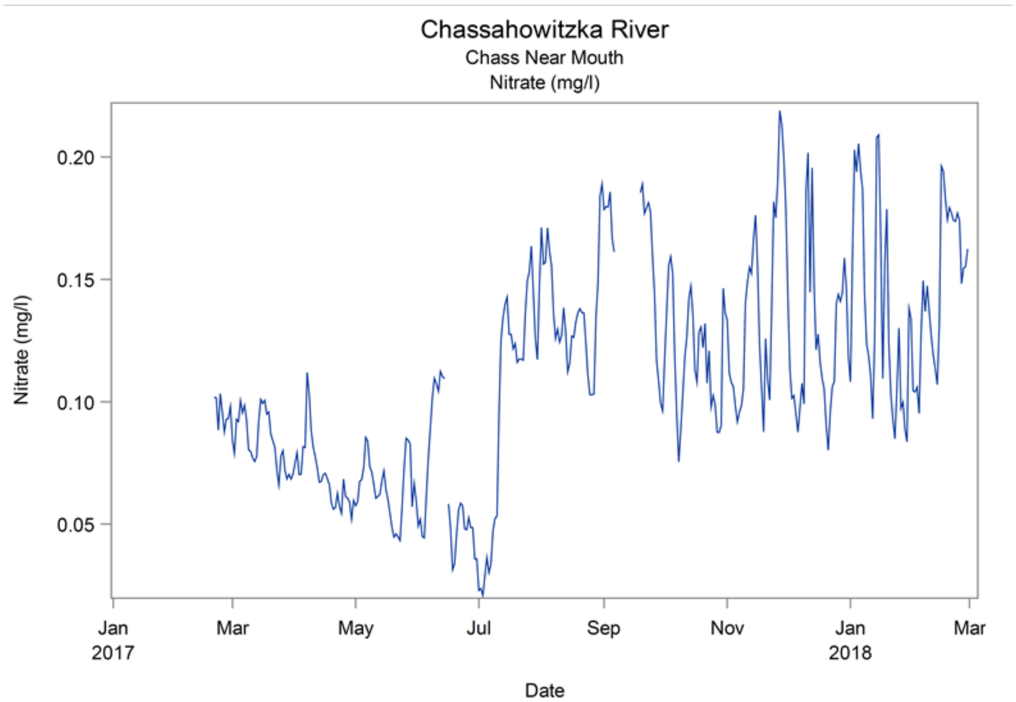


Figure 3-29. **Timeseries plots for nitrite+nitrate at the two downstream continuous recorder sites, near USGS Gage (Top) and Near Mouth (Bottom).**

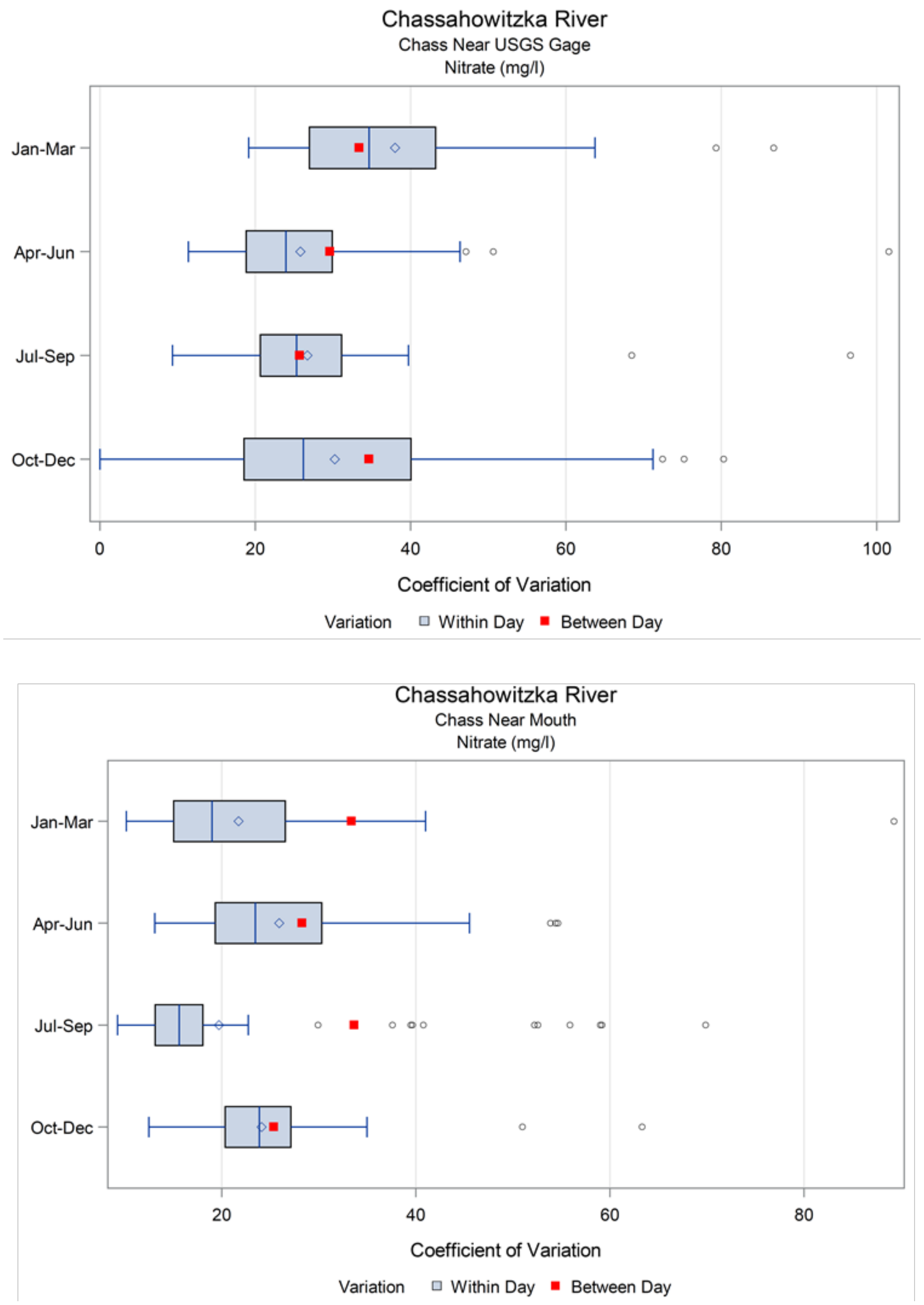


Figure 3-30. Coefficient of variation plots for nitrate-nitrate for the two downstream continuous recorder sites, near USGS Gage (Top) and Near Mouth (Bottom).

Plots for all constituents are provided in Appendix F. Part of the work effort also included spectral decomposition analysis to identify principal dominant frequencies in the continuous recorder data. Spectral decomposition analysis does not allow for missing values which hampers the ability to evaluate the data for dominant frequencies. However, an example of the analysis was performed using a short segment of data for dissolved oxygen (DO) at the near

USGS Gage site to investigate the ability of the spectral decomposition approach to detect short term seasonal signals in the continuous recorder data. The data were hourly DO measurements in mg/l. The analysis cannot include missing values so we used the timeseries between January and September of 2017. DO measurements over that period ranged from 0.01 mg/l to 15.03 mg/l. Spectral analysis suggested the first dominant frequency was 24 indicating a diurnal signal (a frequency = the number of observations required to complete a cycle); this is termed "seasonal" irrespective of the frequency so, in this case, "season" is a day. Defining the dominant frequency for decomposition yields Figure 3-31 that includes the observed data; the 24 hour frequency; the de-seasonalized trend, and the residuals. The time axis is displayed as the number of the cycles so, in this case, days.

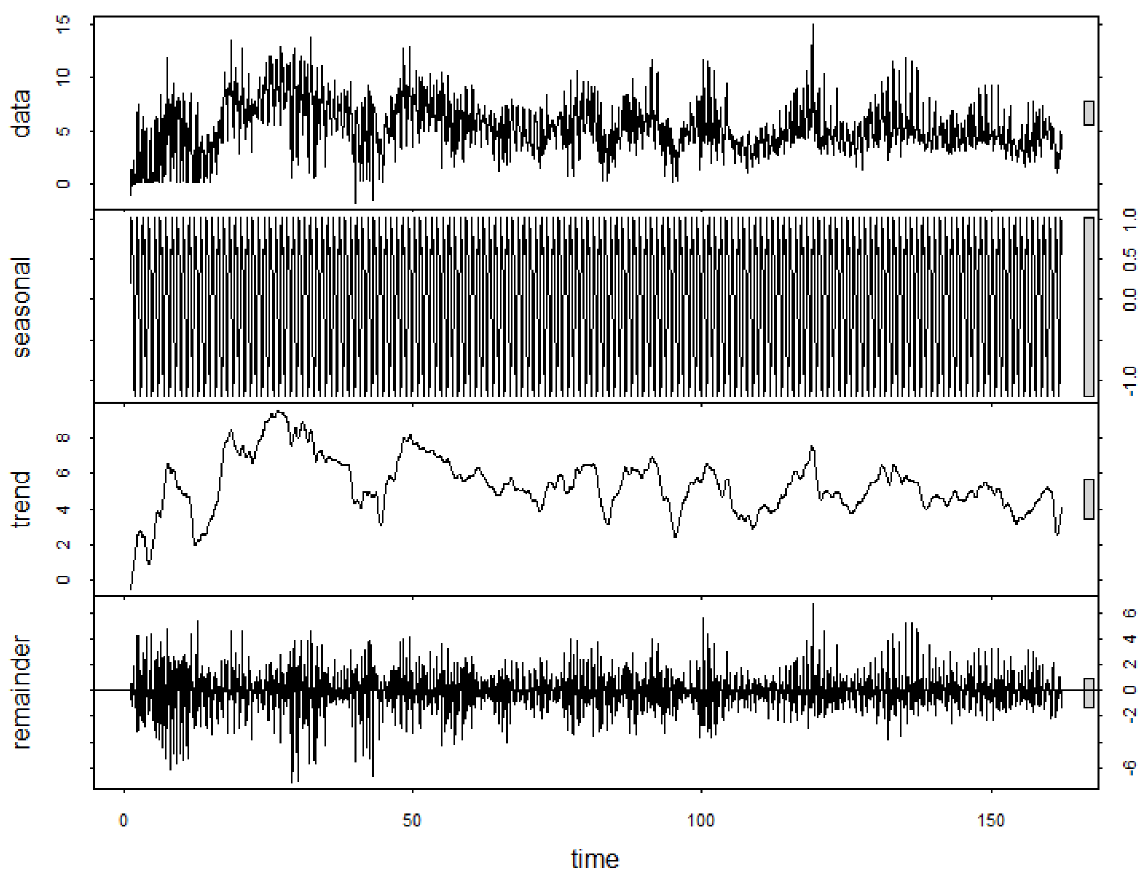


Figure 3-31. Seasonal decomposition of dissolved oxygen timeseries including raw data (top), seasonal cycle identified (second from top), the de-seasonalized timeseries trend, (second from bottom), and the residual (bottom).

To discover an additional frequency in the data we performed spectral analysis on the residuals resulting from the first analysis. This yielded a 12-hour frequency indicating the potential of a tidal signal (plot not shown). These outcomes make sense for dissolved oxygen dynamics in tidal systems including a diurnal component associated with production and respiration and a tidal component associated with mixing of fresh and salt waters. It seemed possible that lunar cycle might also have an effect though no additional dominant frequencies were identified by spectral decomposition. There is a way to specify multiple frequencies into a decomposition method. It is much more complex and has a number of complex embedded functions that perform ARIMA modeling on residuals and box cox transformations, all automatically. We used this method and specified tidal (12-hour), diurnal (24-hour), and lunar (672-hour) frequencies. The results are provided in Figure 3-32. The "observed" plot (top) is simply the raw data. The "level" and "slope" plots below that present the step changes between observations from one time step to another. The three "seasonal" signals from top to bottom are 12, 24, and 672 hour frequencies. More investigation would need to be conducted to determine if the lunar cycle imposed on the data was reasonable.

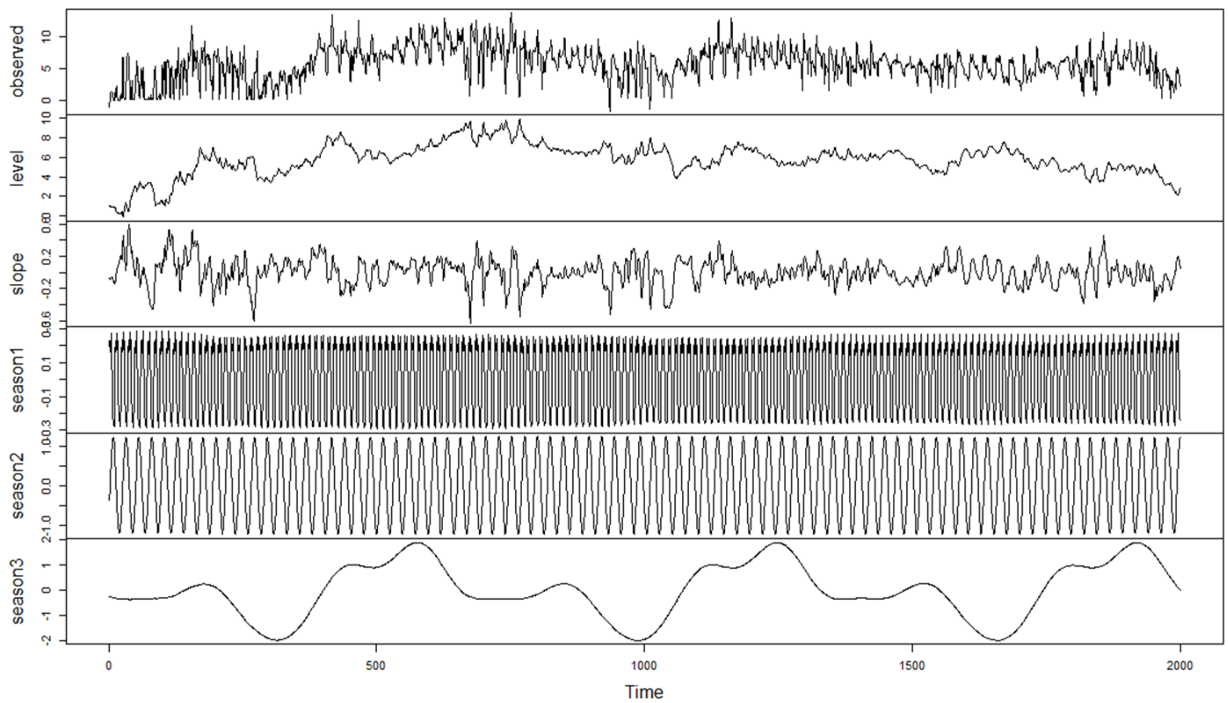


Figure 3-32. Specification of multiple (12, 24, and 672 hour) frequencies in a spectral decomposition of the dissolved oxygen timeseries for the near USGS Gage site.

The results suggest that spectral decomposition may be a valuable approach to identify short term cycles in continuous recorder data, but more data are needed, and an approach to handle missing data needs to be developed in order to accurately identify dominant wave forms in the longer-term continuous recorder data.

4.0 APPLICATION TO MINIMUM FLOWS ASSESSMENT

This work effort was completed to support the District's consideration of the water quality environmental value in its reevaluation of minimum flows for the Chassahowitzka River System. The tools developed as part of this work effort may be used in future analyses to support various aspects of flow management in these systems. In an effort to evaluate the efficacy of these tools for future assessments of the effects of flows on water quality, a summary of the minimum flow evaluation process is provided, and the chlorophyll model developed for the river mainstem was used to assess the potential for this type of model to be used in future assessments.

The goal of a minimum flows determination is to protect the resource from significant harm due to withdrawals. This goal was broadly defined in the enacting legislation as "the limit at which further withdrawals would be significantly harmful to the water resources or ecology of the area." What constitutes "significant harm" was not defined. In the absence of specific stressor-response threshold values identifying significant harm, a 15% reduction in a beneficial attribute of a resource of concern has been identified as a prescriptive standard by which significant harm has been defined. This 15% threshold has been used and supported in the development of the majority of minimum flows developed for Southwest Florida Water Management District which have been peer reviewed and subsequently adopted into Florida Administrative Code. The identification of the threshold values relies on a "percent of flow" approach in which predictive equations or mechanistic models are used in an iterative fashion to evaluate the effects of daily flow reduction scenarios of various increasing percentages of flow until the response threshold is achieved.

Results of the analysis described in Section 3.4 suggested that the model developed to assess the response of chlorophyll distributions to changes in flows had potential to provide supporting evidence to evaluate the water quality environmental value as part of the minimum flows reevaluation for the Chassahowitzka River, though the model would require validation before implementation as a regulatory tool. The sections below detail how the model could be implemented and presents results of that implementation for a hypothetical set of flow reduction evaluations under the assumption that the model is predictive of future conditions.

4.1 FLOW REDUCTION SCENARIOS

To apply results of the chlorophyll a model to evaluate the effects of flow reductions on increases in the exceedance frequency of the chlorophyll threshold, 15 flow reduction scenarios corresponding to 1% to 15% reductions from the baseline flow record for the Chassahowitzka River, in 1% flow-change increments, were developed. The period from 1998 through 2017, generally corresponding to when gauged flows were available for the system, was identified as the period of record for the analyses. Season (i.e. Quarter) was assigned to each date based on month such that January-March was defined as Q1, April-June as "Q2", July-September as "Q3" and October-December as "Q4". In addition, after initial discussion of the model results, the area between sites 1 and 10 was identified as the focus area for analysis since this portion of the system is most likely to be directly influenced by spring flows to the system though the model was developed for the entire system (Figure 4-1). This spatial area is referred to for the remainder of this document as the “inset stratum”. Below site 10 (i.e. Rkm 4.9) the morphometrics of the river change dramatically with the potential for influence from additional tributaries and sheet flow from expansive marsh areas in the lower section of the River. This is evidenced by a plot showing the thermal effects of springs flows during cold events (Figure 4-2) provided in Heyl et al 2012. This “inset stratum” portion of the system is also important low salinity habitat that can serve as nursery areas for juvenile fishes or recreational and commercial value.

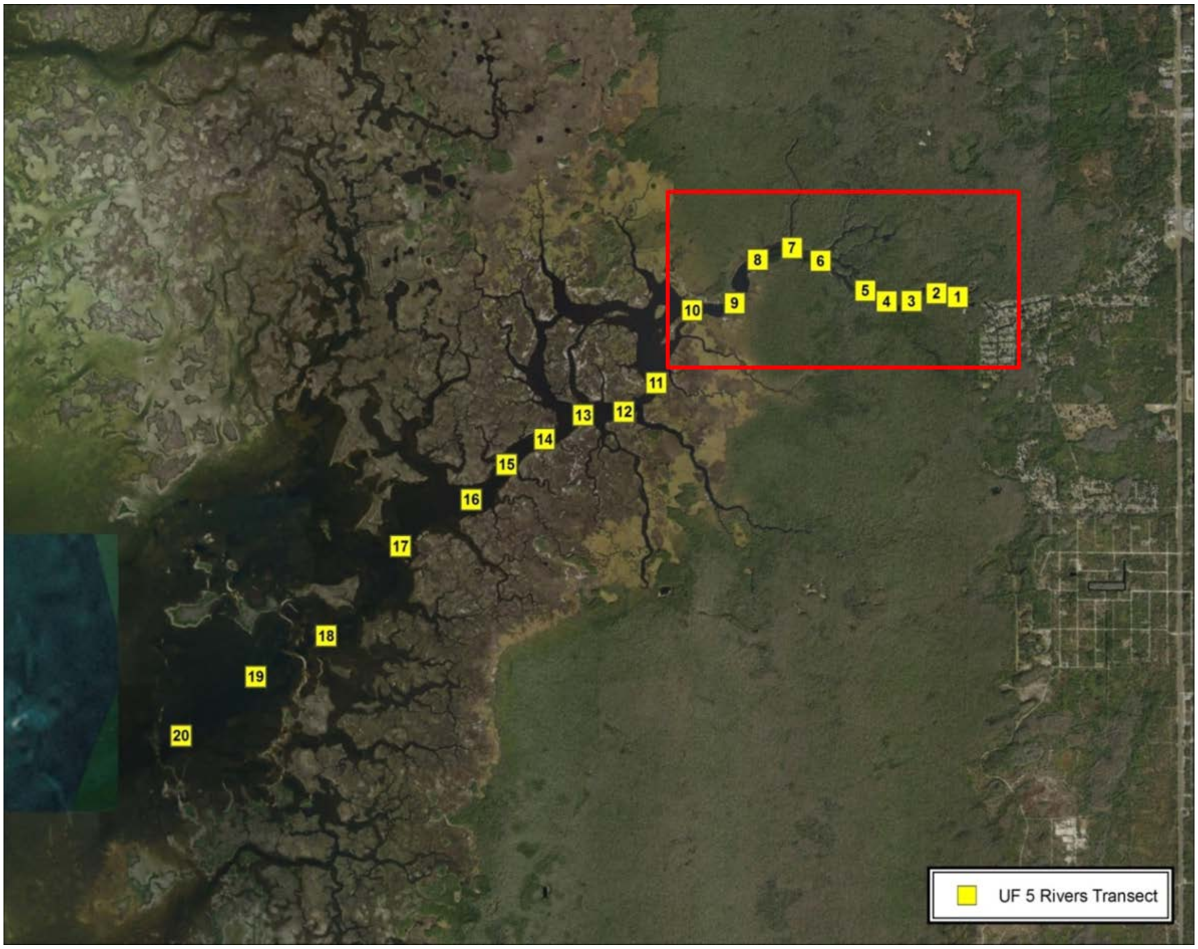


Figure 4-1. **Station locations in the Chassahowitzka River displaying the “inset stratum” relative to morphometric and landscape characteristics of the system.**

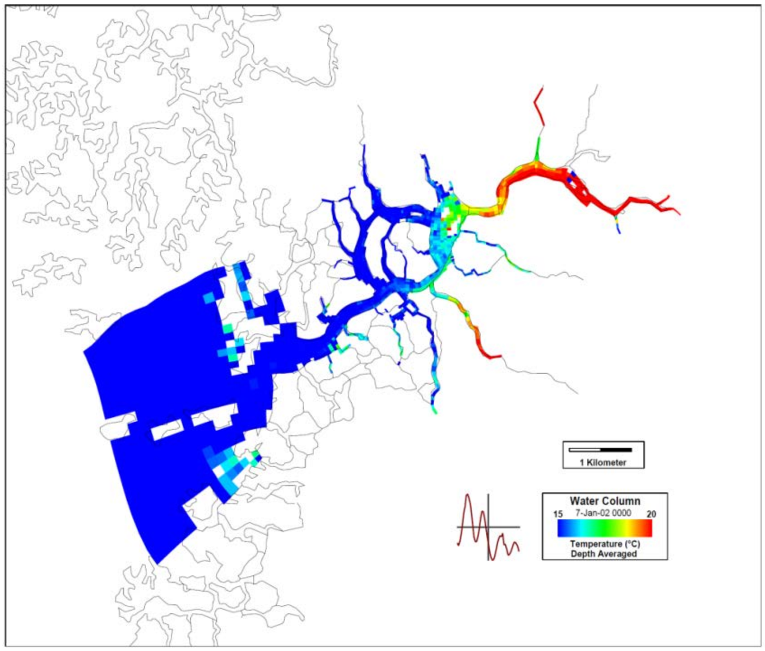


Figure 4-2. Thermal gradient displaying effects of spring flows on temperature as described in Heyl et al. 2012 (red is warmer, spring water).

The GLMM outputs two types of predictions; "marginal" prediction also known as Best Linear Unbiased Estimates (BLUEs), or population level predictions, and "Conditional" estimates (i.e. conditional on the random "site" effects) known as Best Linear Unbiased Predictions (BLUPs). Both predictions are valid and the choice of which to use depends on the question being addressed. For example, if the expectation for the population average at any point in the river is desired, the BLUEs might be chosen to infer the model results to the entire model domain. However, if site-specific characteristics are important to the inference, then the conditional estimates (BLUPs) might be chosen to ensure an adequate representation of the response at the sampled locations (Littell et al. 1996). We consider both these estimates as potential outcomes and describe the differences associated with each outcome.

To evaluate the effects of the flow reduction scenarios, a cutpoint had to be defined to identify whether or not a predicted probability of exceedance would be classified as an exceedance of the site-specific value. A cutpoint value of 0.50 was chosen based on its common use as a standard for a logistic curve of predicted probabilities that approach 1 at some point along the gradient. A plot of the effect of potential alternative cutpoint values on the model fit suggested there was not a clear alternative that would improve the model accuracy relative to the empirical data. The final model (Parameterization 3) was used to predict the probability of exceedance for each date in the timeseries at each station location above Rkm 4.9 and those predicted probabilities were converted to presence/absence identifiers for each scenario using the cutpoint value. The predicted exceedance frequencies were then summed across the entire time period for each flow reduction scenario and summary statistics were generated to evaluate the results. There are several statistics commonly used to evaluate outcomes of logistic regressions that have analogous applications for describing the predicted relative difference in the number of exceedances between the Baseline and a flow reduction scenarios including:

Risk of Exceedance: The percent of the values expected to exceed the standard for a particular scenario.

Risk Difference: Expressed as the difference between the Baseline risk of exceedance and the risk predicted by a flow reduction scenario.

Relative Risk (or Risk Ratio): the risk in the scenario group divided by the risk in the Baseline group.

Odds Ratio: the ratio of the odds of an exceedance in the scenario group divided by the odds of exceedance in the baseline group.

These statistics were used to evaluate the predicted effects of the flow reduction scenarios on the exceedance frequencies and to identify the flow reduction scenario that resulted in a 15% change from the Baseline exceedance rate. The results for the relative risk calculations are presented for each flow reduction scenario in Figure 4-3 for both the BLUP and BLUE estimates. When considering the total predicted exceedance rate for the segment of the river above Rkm 4.9, the 12% flow reduction scenario resulted in a Relative Risk of 1.14, equivalent to a 15% increase in risk of exceedance relative to the Baseline run. The Risk Difference was 6.3% and the Odds Ratio was 1.24, indicating that the odds of exceedance was 1.29 as likely under the 12% flow reduction scenario compared to the Baseline.

The BLUE results were more conservative suggesting any more than an 8 percent reduction in flows would exceed the 15% threshold (Figure 4-3b). To compare the predictions at each sampling site, a panel of paired box plots of the predicted probabilities by quarter under the Baseline and critical reduction scenario are presented for the BLUP and BLUE estimates in Figure 4-4. The effects of the interaction terms is apparent in the plots with the flow reductions resulting in a reduced probability of exceedance in Q1 (Winter) and an increased probability in Q2 (May). In Q3, the Rkm interaction results in flow reductions increasing the probability upstream and decreasing it downstream. The difference between the BLUP and BLUE predictions is also apparent in the site-specific difference for Q2 where the BLUP predictions drop dramatically for both the Baseline and Reduction scenario while the BLUE predictions are a smoother transition as a function of river kilometer.

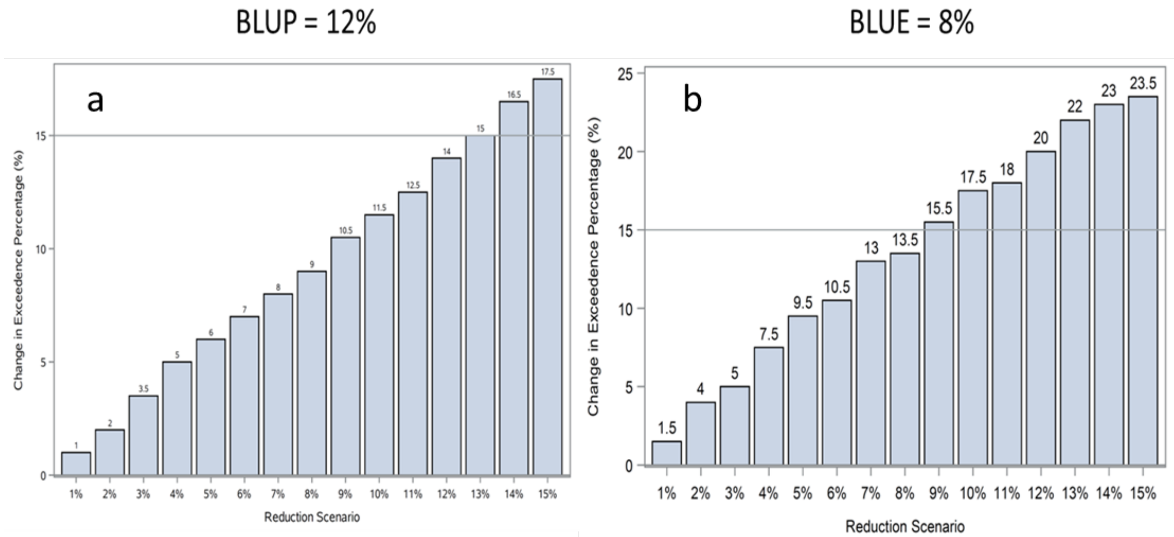


Figure 4-3. **Results of flow reduction scenarios on increase in relative risk of exceeding state water quality standard for chlorophyll a in the inset stratum of the Chassahowitzka River. Numbers above bars represent the relative risk compared to Baseline for each scenario.**

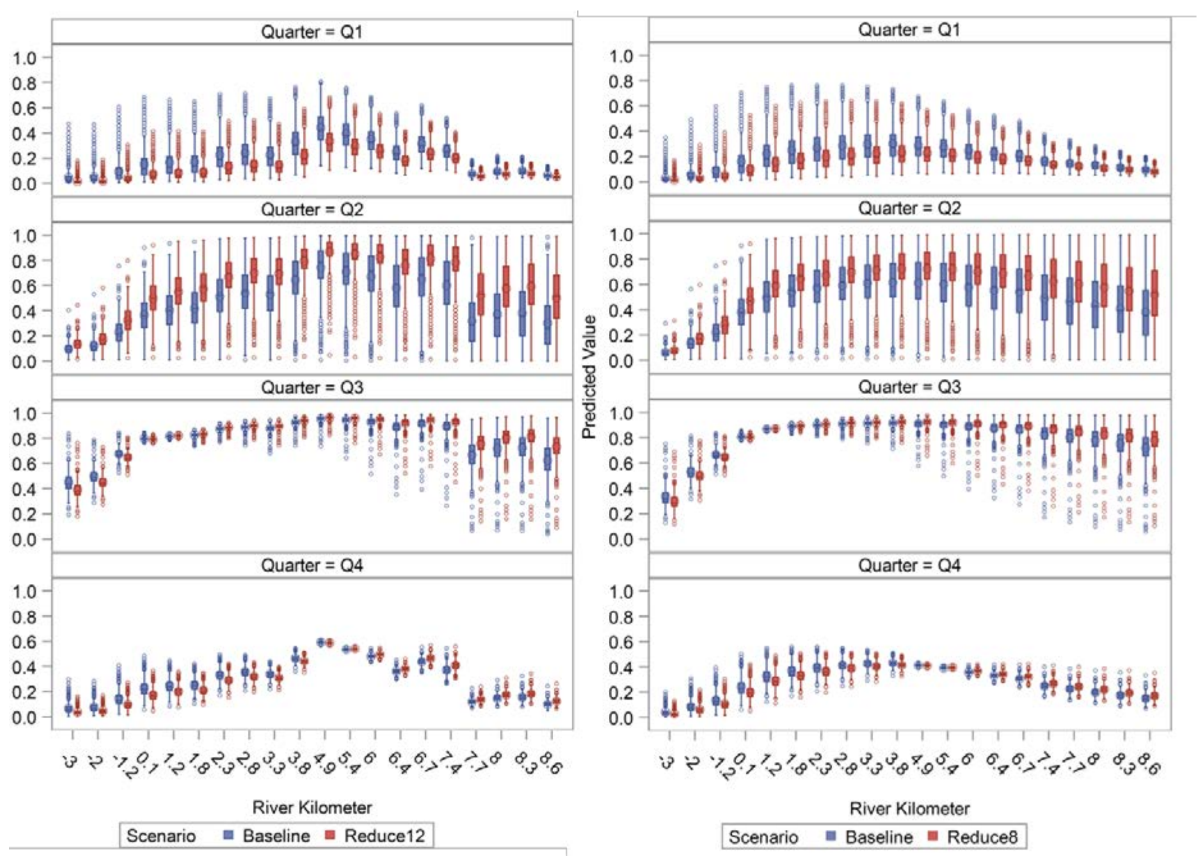


Figure 4-4. Distribution of predicted probabilities by site for the BLUP (left) and BLUE (right) predictions.

4.2 MONTE CARLO SIMULATION

The analysis described in this investigation for the Chassahowitzka River was designed to assess the increased risk of an exceedance of applicable water quality standards to hypothetical flow reduction scenarios and used a 15% change from the Baseline condition as a prescriptive standard by which to identify significant harm. However, the assessment did not, and was not intended to, directly evaluate whether the flow reduction scenario would result in a violation of the site-specific chlorophyll threshold. The state standard is expressed as an annual geometric average and the evaluation was based on a chlorophyll value exceeding that AGM on a particular date. The AGM is used as a regulatory statistic to minimize the effects of data that can be skewed by high values when calculating summary statistics such as the arithmetic mean. By taking the logarithm of a distribution of data that exhibit tendencies to be positively skewed, such as chlorophyll, the transformed data exhibit the bell shaped pattern associated with the normal distribution. This is a common and convenient method used in data analysis to reduce the influence of extreme values on statistical analysis. Since the mean and the median of a normal distribution are nearly equivalent, the AGM generally represents the median of the log normal distribution (Helsel and Hirsch 2002).   If one considers the AGM as the median then 50% of the data should lie above the median and 50% should lie below the median. These properties of the distribution were used to simulate the effects additional exceedances of the standard would have on the AGMs as described below.

The overall distribution of chlorophyll values on the natural scale is provided in Figure 4-5 with a median value of 3.40 ug/l and an arithmetic average of 5.40 ug/l. The distribution of the natural log transformed values is provided in Figure 4-5b and shows how the transformation leads to a bell shaped curve. The mean of the transformed data is 1.25. Exponentiation of that number provides the overall geometric mean value of 3.51 ug/l, close to but not exactly the median value.

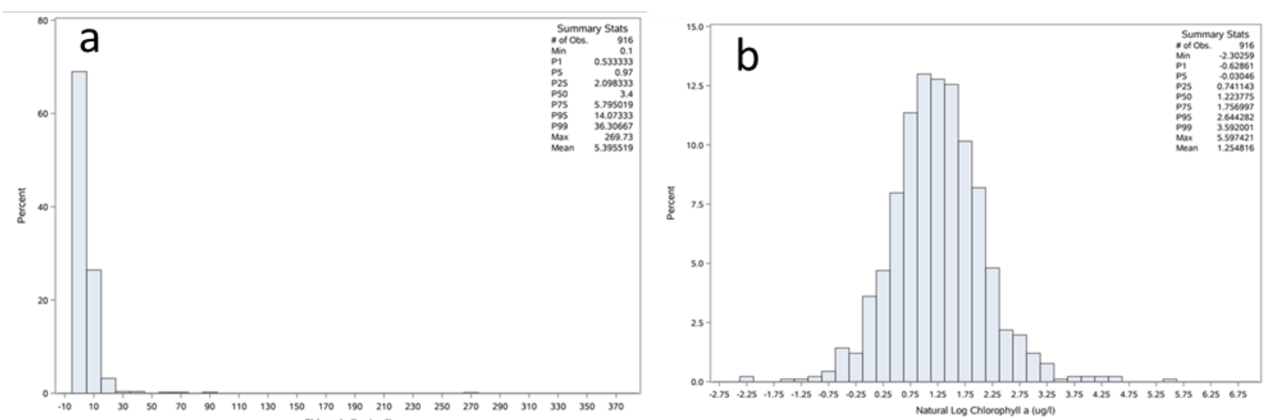


Figure 4-5. Distribution of chlorophyll a on the natural scale (a) and natural log scale (b).

The properties of the empirical distribution can be used to generate an extremely large dataset via monte-carlo simulation and that dataset can be used to evaluate the effects of changes in the exceedance frequency on the annual geometric average. The process involved the following steps:

- Generate the monte-carlo data pool using the properties of the empirical distribution
- Calculate percent exceedance under the existing condition (e.g, 45% exceedance)
- Generate a representative annual distribution for a given year (e.g. 40 samples) by selecting samples at random at the empirical exceedance frequency (i.e. 55% of the samples at or below the standard (3.9 ug/l) and 45% of the samples above the standard.
- For each replicate, calculate the AGM
- Repeat 1000 times

The results yield a distribution of simulated AGMs that represent the existing condition and the approach can be applied to any flow reduction scenario. For example, the final model results above yielded a risk difference of 6.3%. So to run the simulation for that scenario, 51.4% of the samples for each replicate would be selected from the distribution above 3.9 ug/l and 48.6% would be selected from the distribution of values 3.9 or below.

The simulation pool was constructed by using the distributional statistics for each site to generate independent distributions for each site within the system, which were then combined into a single large dataset. A sample size of 40 was chosen to represent an individual year for each replicate sample (i.e. quarterly samples from 10 sites). For each replicate, an AGM was calculated. The distribution of AGMs was then evaluated to define the increased risk of exceeding the standard under the critical flow management scenario identified above. The simulation was performed for the Inset stratum including sites 1 through 10. The empirical AGMs are shown in Figure 4-6a while the monthly flow timeseries is provided in Figure 4-6b. The horizontal reference line on the plot of AGMs denotes the NNC of 3.9 ug/l which was exceeded in several of the more recent years in the timeseries. This time period corresponded to a period of reduced flows in the Chassahowitzka relative to the long term record.

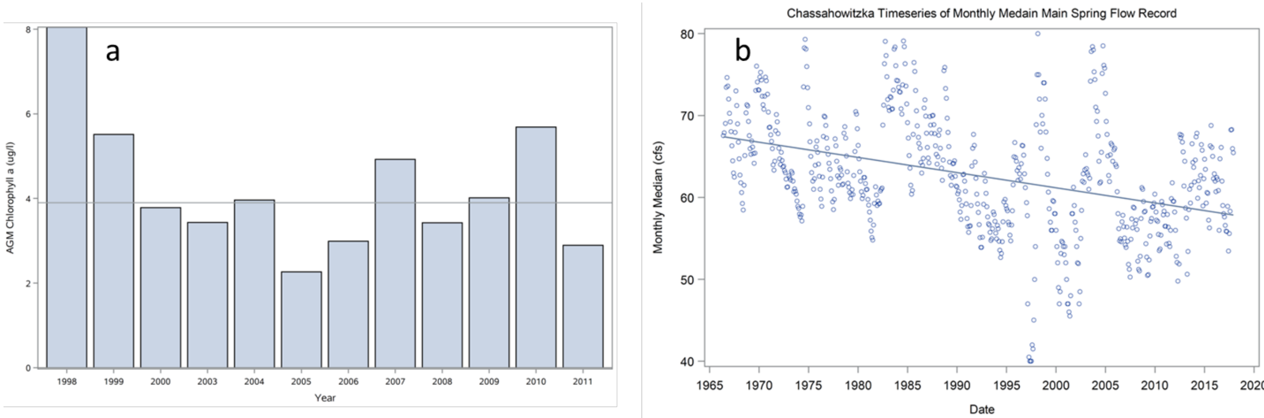


Figure 4-6. Annual geometric average chlorophyll a concentrations in the Chassahowitzka River above Rkm 4.9 based on the UF transect data (left) and flow timeseries for the Chassahowitzka River (right).

The overall geometric mean for this time period for the Inset stratum was 3.88 ug/l. The distribution of AGMs from the simulation of the existing condition is provided in Figure 4-7. The mean and median of this distribution are nearly identical at 3.8 ug/ and very close to the empirical distribution. The vertical reference line in Figure 4-7 indicates the 3.9 ug/l, and the simulated distribution exceeds the site-specific standard approximately 45% of the time. Adjusting the exceedance frequency of the individual samples to represent that in the critical flow scenario of 12% based on the BLUPs increases the expected exceedance frequency of individual samples from 45% percent to 51% (Figure 4-8a). However, because the criterion value of 3.9 is near the median of a (log) normal distribution, sliding the curve to the right results in a large increase in the proportion of AGMs over the criterion value of 3.9; from 45% to over 80% (Figure 4-8b).

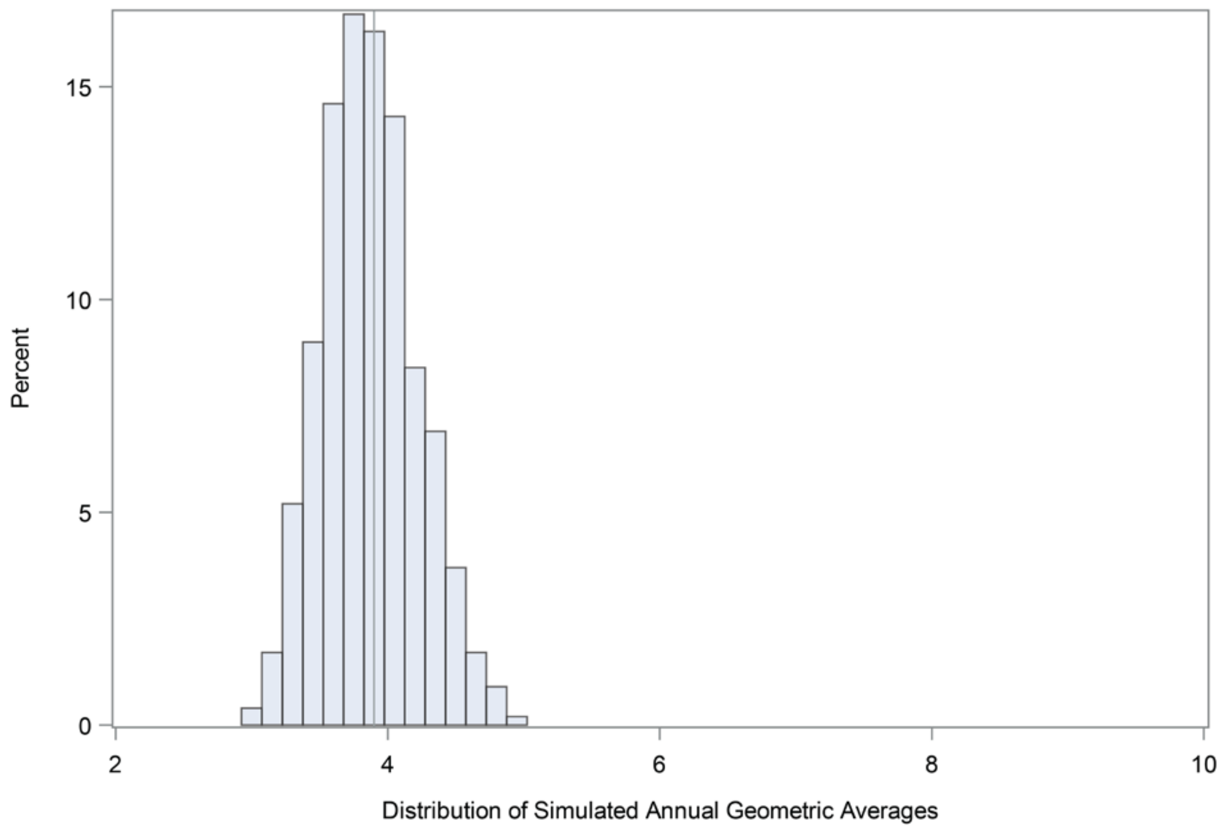


Figure 4-7. **Distribution of simulated annual geometric averages for existing condition above Rkm 7 in the Chassahowitzka River. Vertical reference line is 3.9 ug/l.**

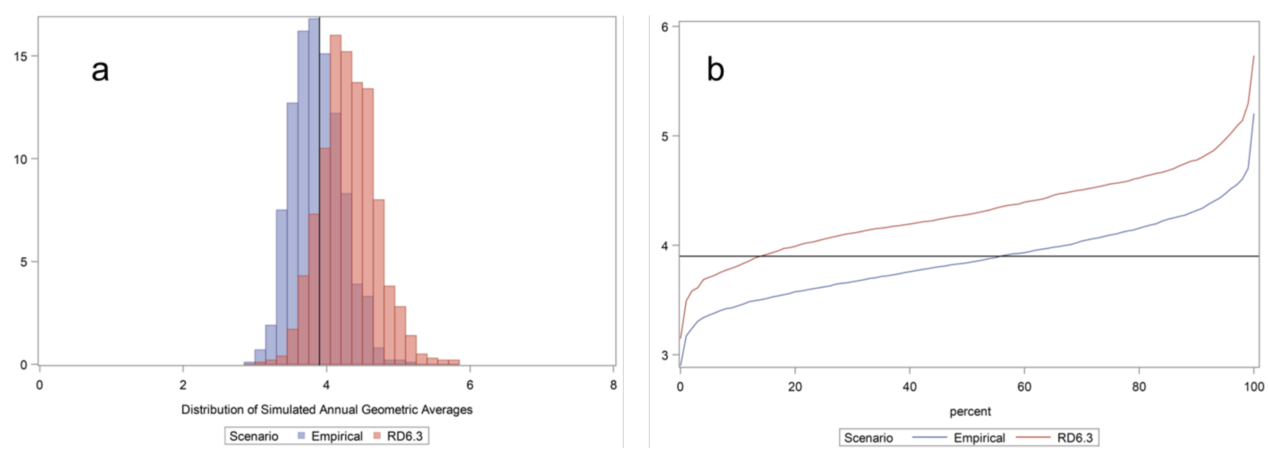


Figure 4-8. **Distributions of simulated annual geometric average chlorophyll a concentrations for the existing and 6% flow reduction condition shown as a frequency histogram (a) and a cumulative distribution plot (b), both with 3.9 ug/l reference lines.**

The results of this monte-carlo simulation illustrate how changes in the exceedance frequency as modeled for the flow reduction scenarios are related to the actual chlorophyll concentration distributions in the Chassahowitzka River. Increased exceedance frequencies of individual samples will increase the overall AGM value but without the monte-carlo approach it was difficult to determine the magnitude to which the actual concentrations are changed. The results of this analysis suggest that, on average, the AGM would be expected to increase from 3.8 to 4.3 ug/l due to the 12% reduction scenario, equating to an estimated percent change in concentration of 13% in the AGM concentration, similar results to the change in exceedance frequency.

4.3 LIMITATIONS AND FUTURE RESEARCH

To date, phytoplankton distributions have not been previously used within the District as the principal determinant of minimum flows for District tidal rivers. Chlorophyll concentrations have been used to support the low-flow threshold established for the Lower Alafia River (Flannery et al. 2008) and have recently been used by the South Florida Water Management District in comparison to state water quality standards as one line of evidence in support the derivation of the revised minimum flow for the Caloosahatchee River Estuary (SFWMD 2018). While phytoplankton distributions are important indicators of riverine and estuarine condition they are notoriously difficult to model due to complex interactions between nutrient availability, light availability, and residence times. That being said, in recent years, Florida has experienced several high profile algae blooms including a protracted red tide event in 2018 and a blue green algae bloom that has affect Lake Okeechobee and its receiving waterbodies on both the east and west coasts. These blooms have attracted much media attention and raised awareness as to the potential negative effects of excessive algal production in both fresh and estuarine systems. In addition, the Chassahowitzka River has an established TMDL to reduce effects of increased nutrient loads to the system that has resulted in excessive primary production and nuisance algal mats in the upper part of the river. While data collected in the Chassahowitzka River have not indicated consequential negative impacts associated with phytoplankton bloom conditions, elevated chlorophyll concentrations that are indicative of bloom potential have been observed in the data.

The NNC threshold value used here was developed for a different section of the river. Therefore, the results presented here were not intended to be used as a direct assessment of whether or not changes in flow would result in compromises to the river's "Designated Use" as defined in State statute. Rather, the modeling effort was developed to illustrate the utility of this type of modeling to assess the sensitivity of phytoplankton (expressed here as chlorophyll a uncorrected for Phaeophytin) concentrations in the upper 4 kilometers of the river to changes in flows. The model results predict that flow reductions, especially in the spring season when flows tend towards their annual minimum, would increase the probability of exceeding a value of 3.9 ug/l. This is a novel model application in support of minimum flows and more research should be completed before this approach can be used in direct support of establishing minimum flows for the Chassahowitzka River. The location of the geographic boundary for evaluating water quality against regulatory criteria and the spatial distribution of chlorophyll in the system complicated the interpretation of the results of flow reduction scenarios, but despite these limitations, the results can be used to provide supporting information to other lines of evidence more directly applicable to the definition of significant harm.

The reported differences associated with the flow reduction scenarios are fairly small in terms of overall risk difference and it is unlikely that one could state with certainty that the reported differences represent statistically different conditions. Unfortunately, there is no standard way of evaluating the statistical certainty of the predictions when evaluating management scenarios such as this. However, it is important to consider that uncertainty exists in the predicted probabilities, and this uncertainty is not accounted for when evaluating the results of changes associated with the flow reduction scenarios. Several tributaries contribute flow and nutrients to the system which are unaccounted for by the flow records or the chlorophyll modeling efforts. Flow records for these sites are lacking long-term records and are presumed to covary with the long-term flow record developed for the minimum flows reevaluation. While these limitations are important to note, they do not obviate the need for protective limits to protect the system from degradation of water quality. The modeling effort has focused on the effects of flow and assumes that other factors that may affect the distribution of chlorophyll in the system are at a stable state for the flow reduction evaluations. In this sense, the results provide the best estimate of the effects of flow on the probability of exceeding the site-specific chlorophyll threshold as applied, but do not imply that there are no other factors that might also affect the distribution of phytoplankton in the system.

Future research should consider the utility of developing nitrate loadings from the head springs. Using nitrate loads as an explanatory variable would eliminate the potentially confounding effects of nutrient dynamics in the downstream portion of the river, and provide a truly independent variable for modeling chlorophyll concentrations. However, considerable additional effort would be required to develop the long-term timeseries of daily nitrate loads needed to simulate the effects of flow reductions on the chlorophyll a response. In addition, evaluating the efficacy of using the downstream, site-specific chlorophyll criterion value of 3.9 ug/l established for WBID 1361 as a management threshold for the entire upper portion of the river (above Rkm 5.9) should be considered. Currently, the WBID boundary bisects the peak of the spatial distribution of chlorophyll. This fact, combined with the fact that the upstream WBID does not have a site-specific chlorophyll threshold, results in a disconnect between the existing criterion and the system dynamics. Evaluating whether or not the current downstream criterion is applicable to upstream portion of the river, or developing an alternative criterion value to protect the upstream portion of the river, would result in a more site-specific protective standard for the portion of the river most affected by variation in spring discharge. There are also alternative modeling choices that could be made that consider the actual chlorophyll concentrations as a response variable; however, to apply those models to flow reduction scenarios associated with the development of minimum flows, an appropriate chlorophyll concentration would need to be developed based on an established threshold for significant harm.

5.0 RECOMMENDATIONS

This study has shown that chlorophyll concentrations are related to flows in the Chassahowitzka River System. Chlorophyll concentrations have not been previously used as a criterion for establishing minimum flows and the modeling approach summarized in this report is a novel approach that should be further investigated. Further research is needed to determine if the site-specific chlorophyll a threshold established for the downstream portion of the river is applicable to the entire upstream portion of the system, and for model validation. In the meantime, the results summarized in this report support consideration of water quality as part of the environmental values assessment associated with reevaluation of the minimum flow established for the Chassahowitzka River System.

6.0 REFERENCES

Burghart, S.E., D.L. Jones and E. Peebles. 2013. Variation in estuarine consumer communities along an assembled eutrophication gradient: indications for trophic instability. Estuaries and Coasts. 36(5). 951-965. https://doi.org/10.1007/s12237-013-9611-6

Burnham, K. P. and D. R. Anderson. 2002. *Model Selection and Multimodel Inference*. Springer-Verlag: New York.

Chen, C. 2002. *Robust Regression and Outlier Detection with the ROBUSTREG Procedure*. Proceedings of the 27th SAS Users Group International Conference, Cary NC: SAS Institute, Inc.

Dodson, J. K. Bridger and G. Maddox. Nutrient TMDL for Chassahowitzka Springs Group. Report of the Florida Department of Environmental Protection. Tallahassee, FL.

Downing, B.D., Bergamaschi, B.A., and Kraus, T.E.C., 2017, Synthesis of data from high-frequency nutrient and associated biogeochemical monitoring for the Sacramento–San Joaquin Delta, northern California: U.S. Geological Survey Scientific Investigations Report 2017–5066, 28 p., https://doi.org/10.3133/sir20175066

FDEP 2013. Numeric Nutrient Criteria for Springs Coast. Report of the Florida Department of Environmental Protection, Tallahassee, FL.

Flannery, M.S. , X. Chen, M. Heyl, A. Munson and M. Dachsteiner. 2008. The determination of minimum flows for the lower Alafia River estuary. Technical publication of the Southwest Florida Water Management District, Brooksville Florida.

Frazer, T.K., M.V.Hoyer, S.K. Notestein, J.A.Hale and D.E.Canfield, D.E., Jr. 2001. Final report. Physical, chemical and vegetative characteristics of five Gulf Coast rivers. SWFWMD contract number 98CON000077. University of Florida Institute of Food and Agricultural Sciences. Gainesville, Florida. Prepared for the Southwest Florida Water Management District. Brooksville, Florida.

Frazer, T.K., S.K. Notestein and W.E. Pine, Jr. 2006. Final report, Changes in the physical, chemical and vegetative characteristics of the Homosassa, Chassahowitzka and Weeki Wachee Rivers. SWFWMD Contract Number 03CON000038. University of Florida Institute of Food and Agricultural Sciences. Gainesville, Florida. Prepared for the Southwest Florida Water Management District. Brooksville, Florida.

Helsel, D.R., and Hirsch, R.M., 2002, Statistical methods in water resources: U.S. Geological Survey Techniques of Water Resources Investigations, book 4, chap. A3, 522 p., http://pubs.usgs.gov/twri/twri4a3/.

Helsel, D.R., Mueller, D.K., and Slack, J.R., 2006, Computer program for the Kendall family of trend tests: U.S. Geological Survey Scientific Investigations Report 2005–5275, 8 p., http://pubs.usgs.gov/sir/2005/5275/.

Heyl, M. 2010. Technical Memorandum dated March 19, 2010 – Estimation of historical Chassahowitzka River flows. Southwest Florida Water Management District. Brooksville, Florida. (Revised April 2012)

Heyl, M. 2012. Technical Memorandum dated February 29, 2012 - Impact of flow on NO3+No2 concentrations in six Florida spring discharges. Southwest Florida Water Management District. Brooksville, Florida. (Revised April 6, 2012).

Heyl, M. D. Leeper, R. Basso., and M. Kelly 2012. Recommended minimum flows for the Chassahowitzka River System. Technical publication of the Southwest Florida Water Management District, Brooksville, FL.

Hirsch, R.M., Moyer, D.L., and Archfield, S.A., 2010, Weighted regressions on time, discharge, and season (WRTDS), with an application to Chesapeake Bay river inputs: Journal of the American Water Resources Association, v. 46, no. 5, p. 857–880.

Hirsch, R.M., and Slack, J.R., 1984, A nonparametric trend test for seasonal data with serial dependence: Water Resources Research, v. 20, p. 727–732.

Hoaglin, D C; F. Mosteller, and J.W. Tukey (Eds). 1983. Understanding Robust and Exploratory Data Analysis. Wiley, New York ISBN 0-471-09777-2.

Hosmer, D.W. and S. Lemeshow. 2000. Applied Logistic Regression. 2nd Edition. John Wiley and Sons Inc. New York, New York

Hothorn T, Hornik K, Zeileis A (2006). Unbiased Recursive Partitioning: A Conditional Inference Framework. *Journal of Computational and Graphical Statistics*. 15(3):651–674, 2006.

Hyndman R, Athanasopoulos G, Bergmeir C, Caceres G, Chhay L, O'Hara-Wild M, Petropoulos F, Razbash S, Wang E, Yasmeen F (2018). *forecast: Forecasting functions for time series and linear models*. R package version 8.4, http://pkg.robjhyndman.com/forecast.

Littell, R. C., G. A. Milliken, W. W. Stroup, and R. D. Wolfinger. 1996. *SAS System for Mixed Models.* Cary, NC: SAS Institute, Inc.

Piegorsch, Walter W. and A. John Bailer. 2005. *Analyzing Environmental Data*. Wiley.

R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Richter, B.D., M.M. Davis, C. Apse, and C. Konrad. 2011. A presumptive standard for environmental flow protection. River Research and Applications.

SAS Institute, Inc. 2016. SAS Institute Inc. 2016. SAS/STAT® 14.2 User's Guide. Cary, NC
Southwest Florida Water Management District. 1997. Water Quality and Hydrology of the Homosassa, Chassahowitzka, Weeki Wachee, and Aripeka Spring Complexes, Citrus and Hernando Counties, Florida – Origin of Increasing Nitrate Concentrations.

SFWMD (South Florida Water Management District) 2018. Technical document to support reevaluation of the minimum flow criteria for the Caloosahatchee river estuary. South Florida Water Management District. West Palm Beach, FL.

SWFWMD (Southwest Florida Water Management District) 2001. The Hydrology and Water Quality of Select Springs in the Southwest Florida Water Management District.

Upchurch, S. B., J. Chen, C. R. Cain, and SDII Global Corporation. 2008. Relationships of Nitrate to Flow in Springs of the Suwannee River Water Management District, Florida. Prepared for the Suwannee River Water Management District Live Oak. Florida.

Yobbi, D. K. and L. A. Knochenmus 1988. Salinity and Flow Relations and Effects of Reduced Flow in the Chassahowitzka River and Homosassa River Estuaries, Southwest Florida. Prepared in Cooperation with the Southwest Florida Water Management District. U.S. Geological Survey Water-Resources Investigations Report 88-4044.

Yobbi, D. K. 1992. Effects of Tidal Stage and Ground-Water Levels on the Discharge and Water Quality of Springs in Coastal Citrus and Hernando Counties, Florida. US. Geological Survey in cooperation with the Southwest Florida Water Management District.

Zar, J. H. 1984. *Biostatistical Analysis*. Prentice Hall. Englewood Cliffs, New Jersey. 07632

Zuur, A. F. E. N. Ieno, N. J. Walker, A. A. Saveliev, G. M. Smith. 2009. Mixed Models with Extensions in Ecology with R. Springer. Spring Street. New York