Charlie Creek Water Quality Analysis Using Generalized Linear Mixed Models
Technical Memo
Revised September 2023

Kristina Deak
Environmental Flows and Levels
Southwest Florida Water Management District

Introduction

Environmental data tend to deviate from a normal distribution and can be impacted by the unique characteristics of a sampling location. Generalized linear mixed models (GLMMs) can be used to predict the probability of an outcome (including a binary response) using inputs of fixed and random effects. Fixed effects are those variables that are assumed to have a constant effect on an outcome. Random effects are characteristics unique to a given sample, such as the influence of unquantified aspects of the sampling location. Input data may be normal or non-parametric and either continuous or categorical (Bolker et al., 2009).

In this analysis, GLMMs were used to predict the probability of exceeding State water quality thresholds (per Chapter 62-302.531, F.A.C.) for Class III waters in Charlie Creek under the proposed minimum flows for the system (Ghile et al., 2023). A similar application of GLMMs was used by Janicki Environmental, Inc. (JEI) through Applied Technology & Management, Inc. (ATM) in their 2018 analysis of water quality in the Chassahowitzka River (ATM and JEI 2018).

The most recently adopted Verified List for Charlie Creek was approved on July 15, 2022. According to this list, Charlie Creek above Peace River (Florida Department of Environmental Protection (DEP) waterbody identification number (WBID) 1763A) is impaired for macrophytes (determined by linear vegetation surveys) and total phosphorus (frequent exceedances of the annual geometric mean threshold) and is a medium priority for TMDL development. The WBID has biological evidence indicating non-attainment of its designated use. Charlie Creek above Old Town Creek (WBID 1763D) is impaired for fecal coliform. Charlie Creek above Oak Creek (WBID 1763B), Little Charlie Bowlegs (WBID 1857), and Fish Branch (WBID 1928) are on the Study List for dissolved oxygen percent saturation exceedances without a known causative pollutant. Apart from a statewide total maximum daily load (TMDL) for mercury (DEP 2013), the DEP has not established a TMDL or basin management plan for any waterbody within Charlie Creek.

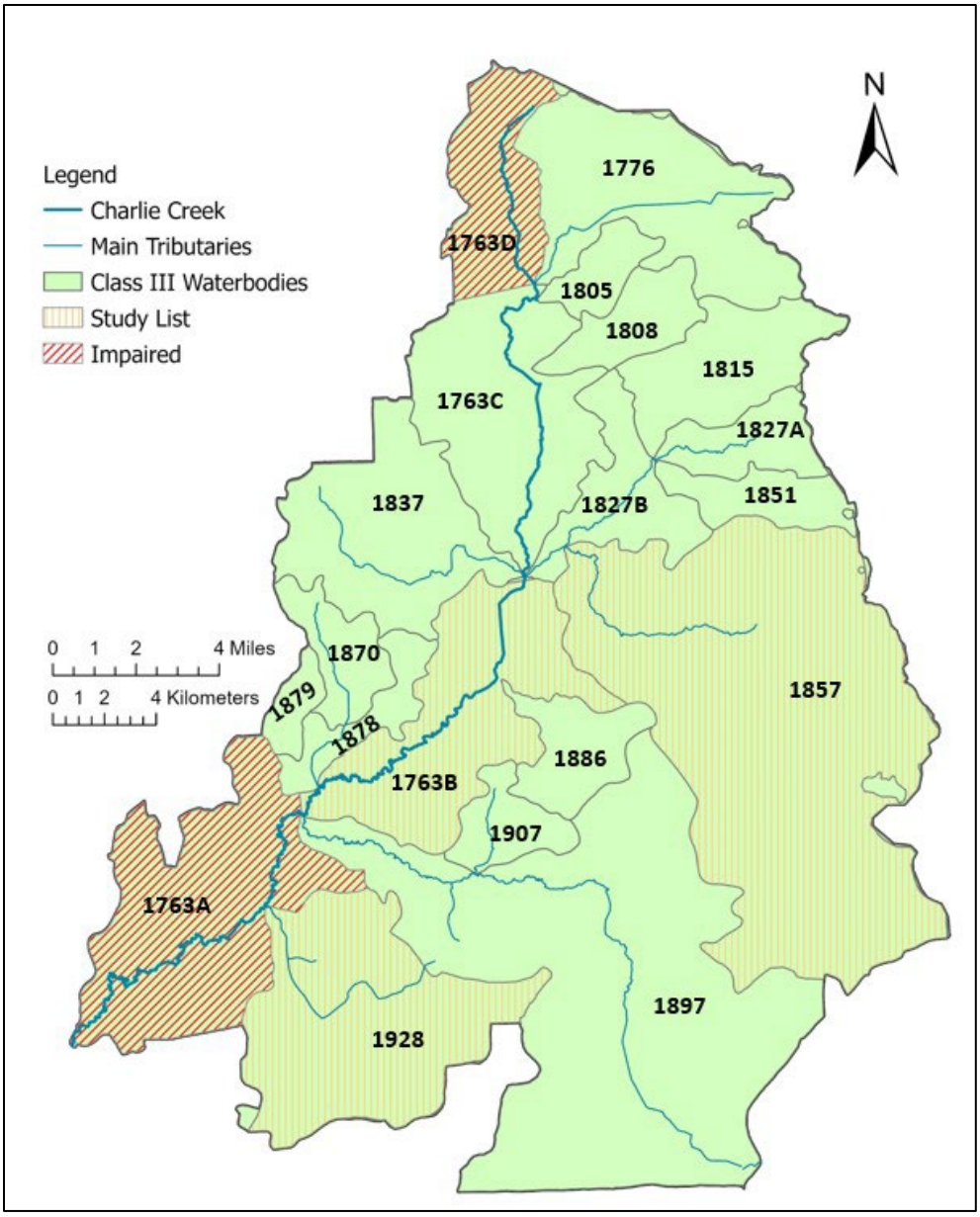


Figure 1. Location of waterbodies by waterbody identification number (WBID) within the Charlie Creek watershed, colored according to designated use classification, impairment status, and inclusion on the study list, according to the DEP's Impaired Waters Rule Run 60 and the Verified List adopted in 2022.

Data

Water quality samples from the DEP, Southwest Florida Water Management District (SWFWMD) and United States Geological Survey (USGS) were used in this analysis (Figure 2). The available period of record and frequency of sampling varied by the sampling agency (Table 1). For consistency with data quality assurance, all water quality data used were pulled from the DEP Impaired Waters Rule Run 59 database, provided by Janicki Environmental, Inc (JEI) through Applied Technology and Management, Inc (ATM and JEI 2021). Baseline (unimpacted) flow data developed for the USGS Charlie Creek near Gardner, FL (No.

02296500) were matched to dates of sample collection. The baseline flow record development is described in Deak et al., 2023.

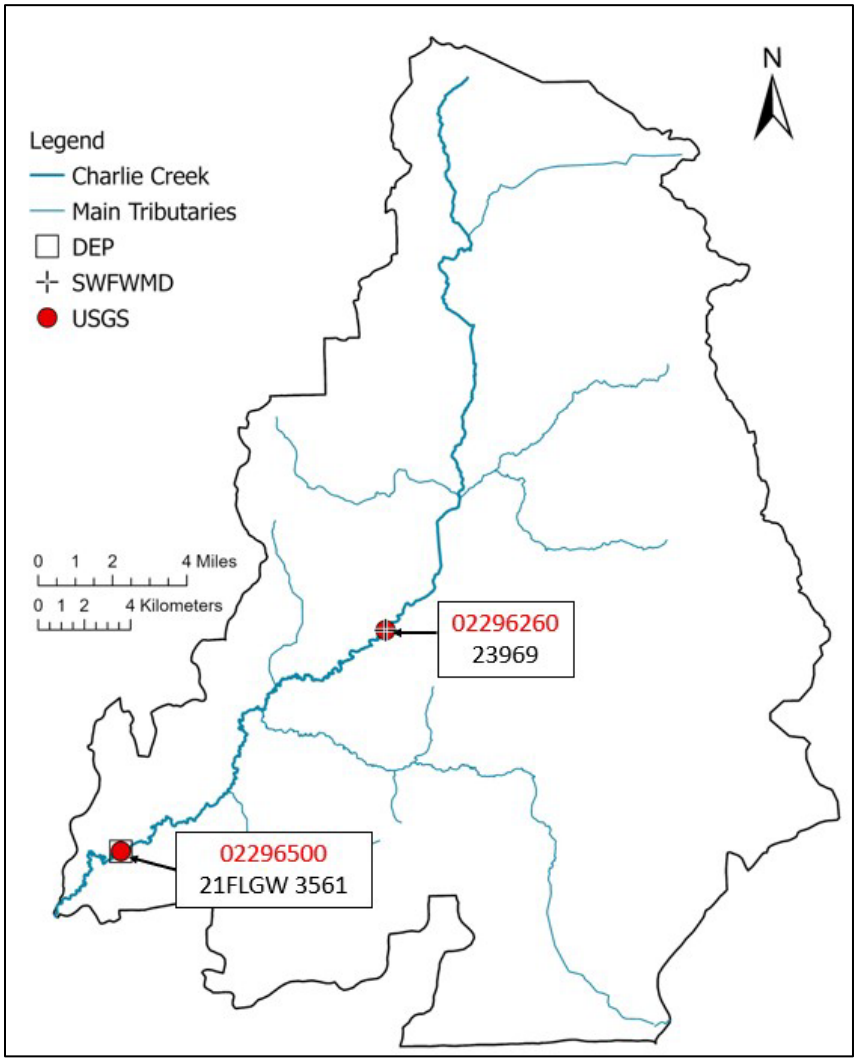


Figure 2. Locations of the water quality sampling sites throughout Charlie Creek considered during analysis. Baseline flow data was based on flows at the USGS station 02296500 (USGS Charlie Creek near Gardner, FL), adjusted for groundwater withdrawals.

Table 1: Sampling information for each water quality parameter in the GLMM analysis, including station name data source, number (n) of samples, the start and end dates of the period of record (POR), and the approximate sampling frequency over the POR.

Parameter	Station	Source	Samples (n)	POR	Approximate Sampling Frequency
Total Phosphorus	21FLGW 3561	DEP	226	10/1998 - 12/2017	Monthly
	23969	SWFWMD	66	10/2009 - 3/2019	Monthly from 2009 to 2011, bimonthly to 2019
	02296500	USGS	76	5/1970 - 9/1998	Variable; from 1 to 8 samples per year
Dissolved Oxygen (% Saturation)	21FLGW 3561	DEP	221	10/1998 - 12/2017	Monthly
	23969	SWFWMD	104	6/2007 - 3/2019	Variable; from 2 to 15 samples per year
	02296500	USGS	126	5/1967 - 9/1999	Variable; from 1 to 9 samples per year

Methods

To predict the probability of State water quality threshold exceedances with flow reductions (a binomial response), GLMMs were created for total dissolved solids, sulfate, dissolved calcium, total nitrogen, total phosphorus, and dissolved oxygen percent saturation using the glmer function in the lme4 package in R programming language (Bates et al. 2015, R Core Team 2021). Models were run for each analyte, considering combinations of the continuous variables (flow) and categorical variables (season, river kilometer) and the interaction terms among them. All models considered "Station" a random effect term. "Season" is defined as the quarter of the year in which samples were taken, beginning in January. This seasonal term was intended to incorporate the impact of variables like temperature, light availability, and biological activity during traditional seasons of the year. If the model failed to converge with raw flows, the log of flows was taken. The successful model with the lowest Akaiki information criterion (AIC) score was selected for further analysis.

The predict function in R was then applied to the selected model to predict the probability of State water quality threshold exceedance at a given flow and location. low reduction scenarios were run according to the recommended minimum flow conditions (Table 2) to determine if the minimum flows would increase the likelihood of 0.5 probability threshold exceedance of State water quality criteria compared to baseline conditions.

Table 2: Proposed minimum flows for Charlie Creek based on flows at the USGS Charlie Creek near Gardner, FL (No. 02296500) gage that have been adjusted for withdrawal effects.

Flow-Based Block	If Adjusted Flow, in Cubic Feet per Second (cfs) on the Previous Day is:	Minimum Flow is:
1	≤ 27 cfs	100% of the flow on the previous day
2	> 27 cfs and ≤ 120 cfs	86% of the flow on the previous day
3a	> 120 cfs and ≤ 316 cfs	88% of the flow on the previous day
3b	> 316 cfs and ≤ 945 cfs	91% of the flow on the previous day
3c	> 945 cfs	93% of flow on the previous day

Results

Total Phosphorus

Data from DEP station 21FLGW 3561, the USGS Charlie Creek near Gardner, FL (No. 02296500) station, and SWFWMD station 23969 were utilized in the creation of GLMMs for total phosphorus. Sample distributions are presented in Figure 3. Exceedance of the State Class III water quality threshold of 0.49 mg/L total phosphorus was utilized as the binary response. Exceedance of the State water quality threshold occurred frequently over the POR for available total phosphorus data.

The GLMM model for total phosphorus with the lowest AIC (359.4) considered the log transformed baseline flows, season, river kilometer, and the interaction between river kilometer and log transformed baseline flows. When all data were visualized, the probability of exceeding the State water quality criteria for total phosphorus surpassed the 0.5 probability threshold during all flow periods for DEP station 21FLGW 3561 and the USGS 02296500 station, and during Block 3b and Block 3c at the SWFWMD station 23969 (Figure 4).

Trends were apparent in the data, when dots representing predictions for samples were compared to the smoothed polyline for each station (Figure 4). These trends reflect a seasonal influence as explored in Figure 5. At the upstream station, 23969, the probability of state water quality threshold exceedance increased with increasing flow. At the co-located stations downstream, 21FLGW 3561 and 02296500, the relationship is the opposite. Only station 21FLGW 3561 is shown in Figure 5 as the trend is the same for both stations and more samples were collected at the highlighted station. For all stations, the probability of threshold exceedance was greatest during the typically wet summer season (July-September) and lowest during the winter (January – March; Figure 5).

The recommended minimum flows did not change the number of samples surpassing the 0.5 probability threshold for exceedance of the State water quality criteria for total phosphorus for stations 21FLGW 3561 or 02296500 in any flow block (Table 3). For station 23969, the recommended minimum flow for Block 2 reduced the number of samples expected to surpass the 0.5 probability threshold by one, from four of 28 samples tested predicted to exceed the threshold under baseline conditions to three samples under the recommended minimum flows. The same result was observed during Block 3a, although only 10 samples were available to be tested in this block (Table 3).

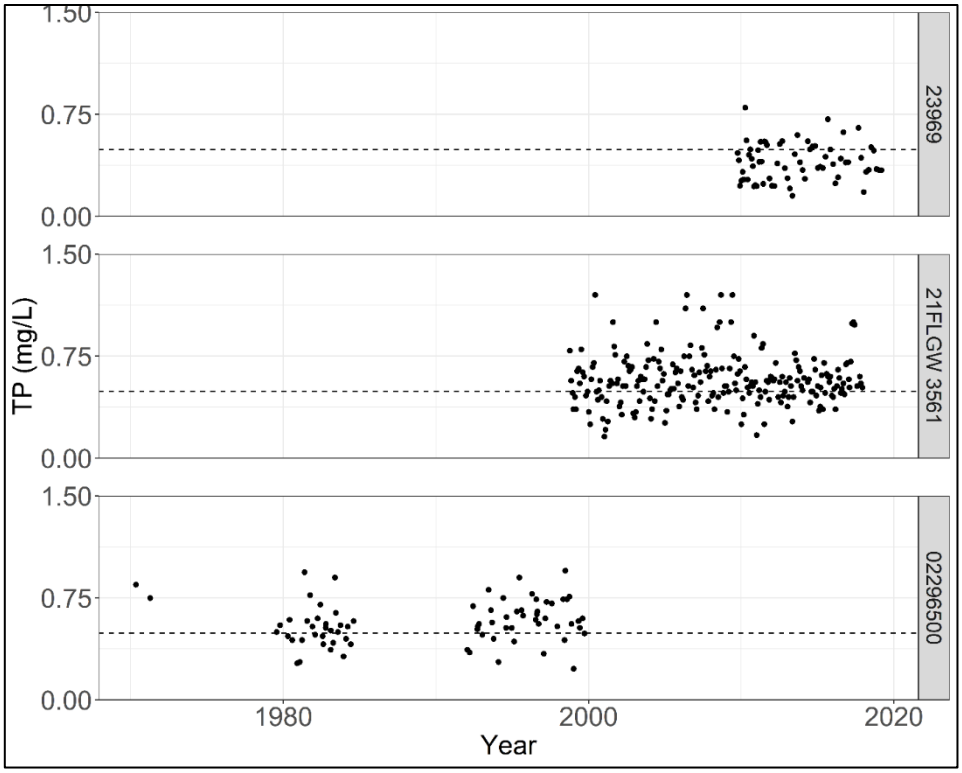


Figure 3. Sample distribution for total phosphorus (TP) at SWFWMD (23969), DEP (21FLGW 3561), and USGS (02296500) stations, listed from upstream to downstream in the watershed. The dashed line indicates the total State water quality threshold for Class III waterbodies (0.49 mg/L). Note this threshold is established for the annual geometric mean rather than for individual samples.

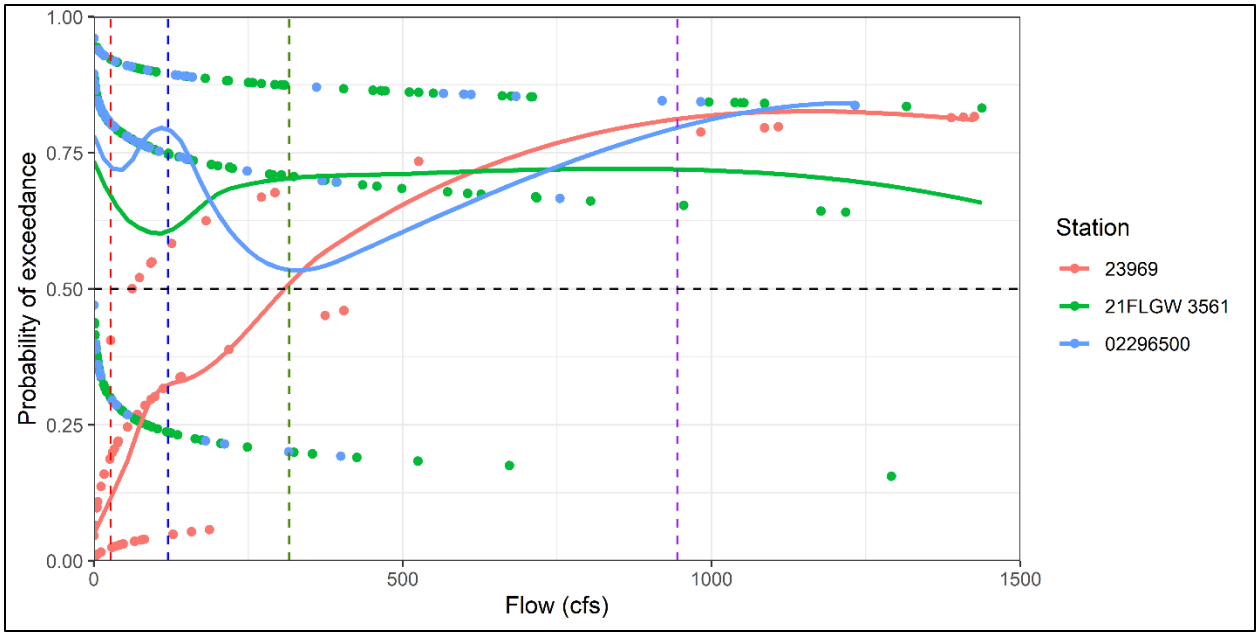
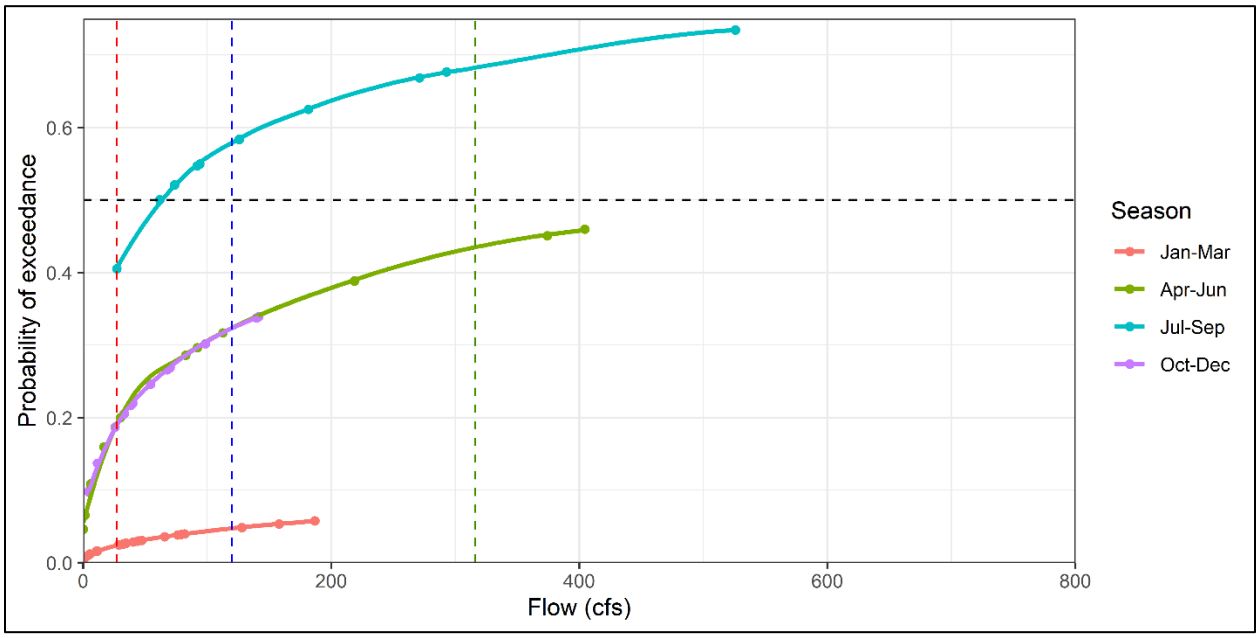


Figure 4: Predicted probability of exceedance of the State water quality threshold for total phosphorus in Class III waters (0.49 mg/L) compared to baseline flow at the USGS Charlie Creek near Gardner, FL gage (No. 02296500). The dashed black line indicates the 0.5 probability of exceedance threshold. The dashed red line indicates the border of Block 1 and Block 2 (27 cfs), the dashed blue line delineates the border between Block 2 and Block 3a (120 cfs), the green dashed line shows the border between Block 3a and Block 3b (316 cfs), and the purple dashed line delineates the border between Block 3b and Block 3c (945 cfs). Dots show predicted probabilities for individual samples under baseline flow conditions and locally estimated scatterplot smoothing (LOESS) polylines are provided for each station. Stations are listed in the key from upstream to downstream.

a) 23969



b) 21FLGW 3561

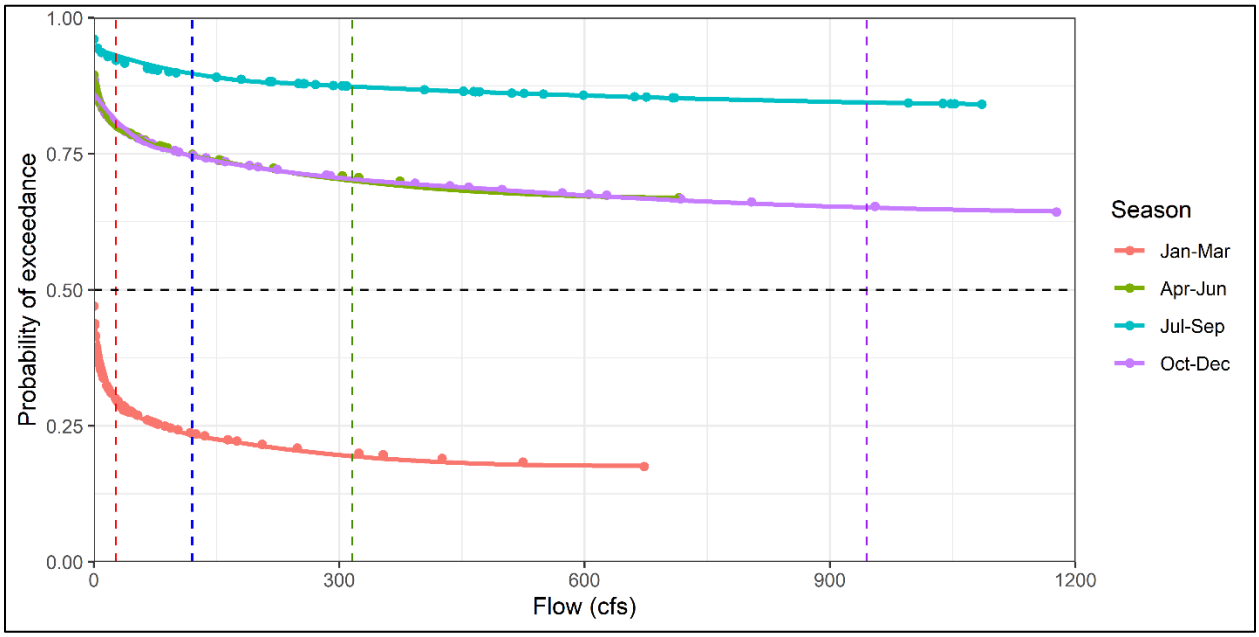


Figure 5: Predicted probability of exceedance of the State water quality threshold for total phosphorus in Class III waters (0.49 mg/L) compared to baseline flow at the USGS Charlie Creek near Gardner, FL gage (No. 02296500) for station a) 23969 and b) 21FLGW 3561. The dashed black line indicates the 0.5 probability of exceedance threshold. The dashed red line indicates the border of Block 1 and Block 2 (27 cfs), the dashed blue line delineates the border between Block 2 and Block 3a (120 cfs), the green dashed line shows the border between Block 3a and Block 3b (316 cfs), and the dashed purple line in Figure 5b separates Block 3b from Block 3c (945 cfs). Dots show predicted probabilities for individual samples under baseline flow conditions and locally estimated scatterplot smoothing (LOESS) polylines are provided for each season.

Table 3: The number of total phosphorus samples predicted to surpass the 0.5 probability threshold for State water quality criteria exceedance under baseline flow conditions and proposed minimum flows, as predicted using generalized linear mixed models. Stations and flow blocks where probabilities of exceedance were less than 0.5 are not shown.

Station	Block	Samples (n)	Samples Exceeding 0.5 Probability of Threshold Exceedance (n)		Difference Between Baseline and MFL (n)
			Baseline	MFL	
23969	2	28	4	3	-1
	3a	10	4	3	-1
	3b	3	3	3	0
	3c	8	8	8	0
21FLGW3561	2	56	37	37	0
	3a	33	27	27	0
	3b	30	25	25	0
	3c	18	17	17	0

Dissolved Oxygen

Data from DEP station 21FLGW 3561, the USGS Charlie Creek near Gardner, FL (No. 02296500) station, and SWFWMD station 23969 were utilized in the creation of GLMMs for dissolved oxygen percent saturation. Their sample distribution is presented in Figure 5. Sample values below the DEP Class III water quality threshold of 38% were utilized as the binary response. Over the POR for available dissolved oxygen percent saturation data, exceedance of State water quality criteria for this parameter occurred in approximately 36% of samples at SWFWMD station 23969 (Figure 6).

The GLMM model for dissolved oxygen percent saturation with the lowest AIC (107.4) considered log transformed baseline flow and river kilometer. When all data were visualized, the probability of exceeding the State water quality criteria increased with increasing flows at station 23969 and surpassed the 0.5 probability threshold in Blocks 3a and 3b at this station (Figure 7). The probability of exceedance was very low for the other stations during all flow blocks.

Under the proposed minimum flows, three samples of the 17 tested at station 23969 were predicted to surpass the 0.5 probability threshold for exceedance of the State water quality criteria for dissolved oxygen percent saturation, as opposed to seven samples under baseline conditions. There was no predicted change in the number of samples surpassing the 0.5 probability threshold at this station during Block 3b at station 23969.

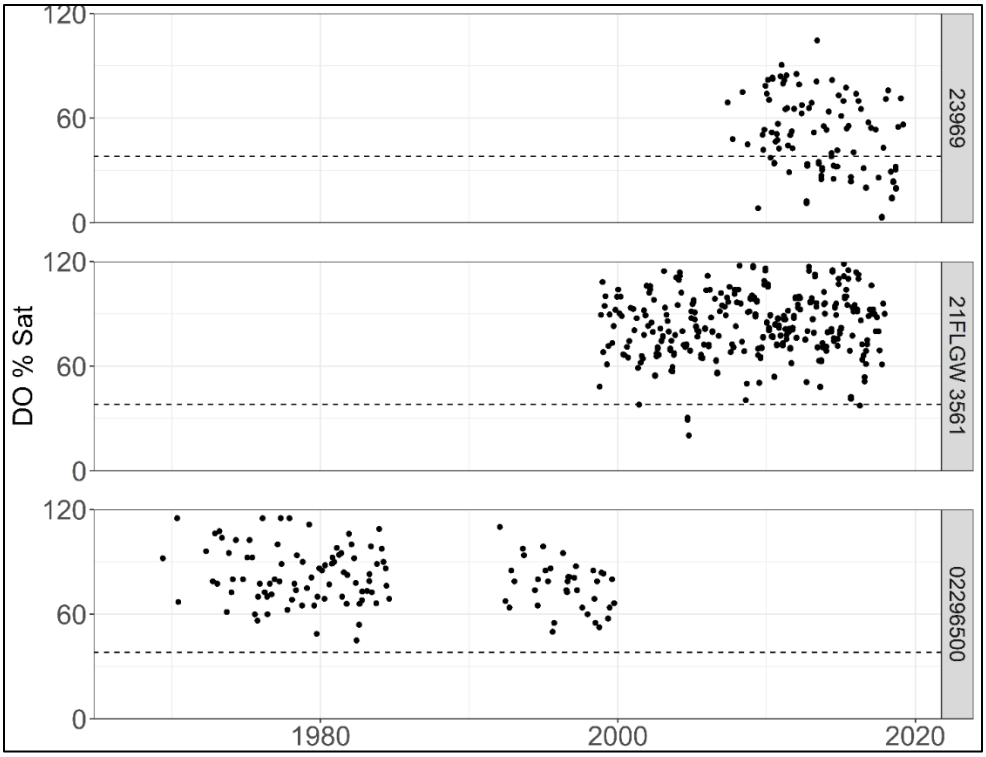


Figure 6. Sample distribution for dissolved oxygen percent saturation (DO % Sat) at SWFWMD (23949), DEP (21FLGW 3561), and USGS (02296500) stations, listed from upstream to downstream in the watershed. The dashed line indicates the total State water quality threshold for Class III waterbodies (38%).

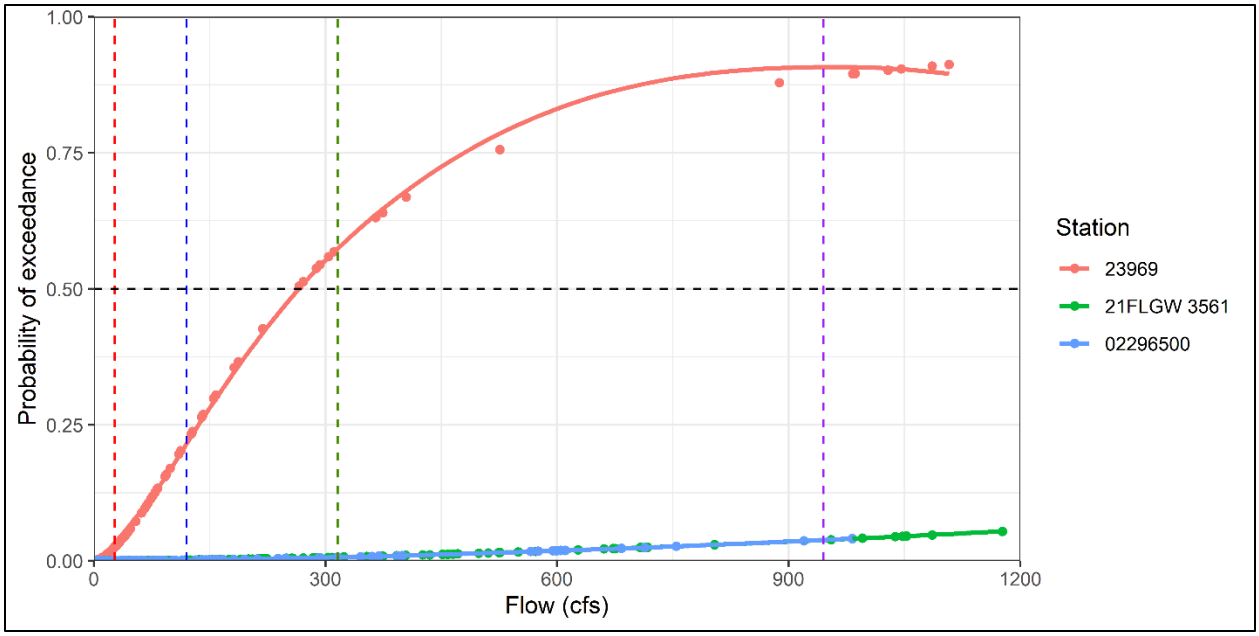


Figure 7: Predicted probability of exceedance of the State water quality threshold for dissolved oxygen percent saturation in Class III waters (38%) compared to baseline flow at the USGS Charlie Creek near Gardner, FL gage (No. 02296500). The dashed black line indicates the 0.5 probability of exceedance threshold. The dashed red line indicates the border of Block 1 and Block 2 (27 cfs), the dashed blue line delineates the border between Block 2 and Block 3a (120 cfs), the green dashed line shows the border between Block 3a and Block 3b (316 cfs), and the dashed purple line delineates the boundary between Block 3b and Block 3c (945 cfs). Dots show predicted probabilities for individual samples under baseline flow conditions and locally estimated scatterplot smoothing (LOESS) polylines are provided for each station. Stations are listed in the key from upstream to downstream.

Table 4: The number of dissolved oxygen percent saturation samples predicted to surpass the 0.5 probability threshold for State water quality criteria exceedance under baseline flow conditions and proposed minimum flows, as predicted using generalized linear mixed models. Stations and flow blocks where probabilities of exceedance were less than 0.5 are not shown.

Station	Block	Samples (n)	Samples Exceeding 0.5 Probability of Threshold Exceedance (n)		Difference Between Baseline and MFL (n)
			Baseline	MFL	
23969	2	31	0	0	0
	3a	17	7	3	-4
	3b	8	8	8	0
	3c	23	23	23	0

Summary

The State water quality threshold for total phosphorus is based upon the annual geometric mean of samples frequently exceeding the threshold in a three-year period. The State water quality threshold for dissolved oxygen percent saturation is exceeded if more than 10% of the daily average dissolved oxygen percent saturation values are below 38%. While the statistics calculated below are modeled off available sample data, and therefore, reflect the probability of exceedance on a per sample basis, it is assumed that if the number of samples exceeding the threshold is not substantially increased by flow reduction, the probability of exceeding the State water quality threshold once an annual geometric mean or 10% of daily averages is calculated would also not increase.

Based upon results from this analysis, the proposed minimum flows for Charlie Creek are not anticipated to increase the 0.5 probability of exceeding the State water quality thresholds for total phosphorus or dissolved oxygen percent saturation at the evaluated water quality stations (Tables 3 and 4).

Of note, water quality data in Charlie Creek are limited and there was little overlap in the period of record or sampling frequency in the available data. Robust sampling over a longer period of record with additional water quality stations along the length of the river would improve confidence in the model outputs.

Works Cited

Applied Technology and Management, Inc. and Janicki Environmental, Inc. (ATM and JEI) 2018. Exploratory evaluation of water quality and flow relationships for the Chassahowitzka River in support of minimum flows reevaluation. Prepared for the Southwest Florida Water Management District. Brooksville, Florida.

Applied Technology and Management, Inc. and Janicki Environmental, Inc. (ATM and JEI) 2021. Charlie Creek water quality assessment. Prepared for the Southwest Florida Water Management District. Brooksville, Florida.

Bates, D. M. Machler, B. Bolker, and S. Walker. 2015. Fitting linear mixed effects models using lme4. Journal of Statistical Software. 67:1-48.

Bolker, B.M., M.E. Brooks, C.J. Clark, S.W. Grange, J.R. Poulsen, H.H. Stevens and J-S. S. White. Generalized linear mixed models: a practical guide for ecology and evolution. Trends in Ecology and Evolution. 24:119-174.

Deak, K., Y. Ghile, G. Herrick, and D. Leeper. 2023. Proposed minimum flows for Charlie Creek. Southwest Florida Water Management District. Brooksville, Florida.

R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.