



Estuarine Nutrient Enrichment Monitoring: Data Summary and Trend Analysis for Gateway National Recreation Area

January 2, 2023

Prepared by:
Christopher P. Peck
Paul Duffy, PhD

Prepared for:
Holly K. Plaisted
Protocol Lead
Northeast Coastal and Barrier I&M Network
National Park Service



Neptune and Company, Inc.
1435 Garrison Street, Suite 201
Lakewood, Colorado 80215

Estuarine Nutrient Enrichment Monitoring:

Data Summary and Trend Analysis for Gateway National Recreation Area

Christopher P. Peck

Paul Duffy, PhD

Neptune and Company, Inc.

1435 Garrison Street, Suite 201

Lakewood, Colorado 80215

Please cite this publication as:

Peck, Christopher P., Duffy, Paul. 2023. Estuarine Nutrient Enrichment Monitoring: Data Summary and Trend Analysis for Gateway National Recreation Area. Neptune and Company, Inc., Lakewood, Colorado.

Contents

	Page
Figures.....	5
Tables.....	8
Acronyms and Abbreviations.....	10
1. Background Information.....	11
2. Spatial-Station Water Quality Data.....	13
2.1 Methods.....	13
2.1.1 Area-Weighted Condition.....	13
2.1.2 General Linear Mixed Model (GLMM).....	14
2.1.3 Generalized Additive Model (GAM).....	14
2.1.4 Multivariate Mann-Kendall Trend Analysis.....	16
2.2 Area-Weighted Condition Results.....	16
2.2.1 Chlorophyll-a (near surface).....	16
2.2.2 Dissolved Oxygen (near bottom).....	16
2.2.3 PAR Attenuation K_d (SAV standard).....	16
2.2.4 pH.....	16
2.2.5 Salinity.....	16
2.2.6 Temperature (near surface).....	16
2.2.7 Turbidity.....	16
2.3 GLMM Results.....	19
2.3.1 Chlorophyll-a (near surface).....	19
2.3.2 Dissolved Oxygen (near bottom).....	19
2.3.3 PAR Attenuation K_d (SAV standard).....	19
2.3.4 Salinity (near surface).....	19
2.3.5 Temperature (near surface).....	19
2.3.6 Turbidity (near surface).....	20
2.4 GAM Results.....	33

2.4.1 Chlorophyll-a (near surface).....	33
2.4.2 Dissolved Oxygen (near bottom).....	33
2.4.3 PAR Attenuation K_d (SAV standard).....	33
2.4.4 Salinity (near surface).....	33
2.4.5 Temperature (near surface).....	33
2.4.6 Turbidity (near surface).....	33
2.5 Multivariate Mann-Kendall Trend Test Results.....	40
3. Trend-Station Water Quality Data	41
3.1 Methods	41
3.1.1 Generalized Additive Model (GAM)	41
3.1.2 Multivariate Mann-Kendall Trend Analysis	41
3.2 Summary Statistics	41
3.3 GAM Results	45
3.3.1 Chlorophyll-a (near surface).....	45
3.3.2 Dissolved Oxygen (near bottom).....	45
3.3.3 PAR Attenuation K_d (SAV standard).....	45
3.3.4 Salinity (near surface).....	45
3.3.5 Temperature (near surface).....	45
3.3.6 Turbidity (near surface).....	45
3.4 Multivariate Mann-Kendall Trend Test Results.....	53
4. Logging-Station Continuous Water Quality Data.....	54
4.1 Continuous Record Plots	54
4.2 Methods	54
4.2.1 Generalized Additive Model (GAM)	54
4.2.2 Additional Data Summaries	55
4.3 Results	55
4.3.1 Chlorophyll-a.....	55
4.3.2 Dissolved Oxygen	55

4.3.3 PAR Attenuation K_d	55
4.3.4 pH	55
4.3.5 Salinity.....	55
4.3.6 Temperature.....	56
4.3.7 Turbidity.....	56
5. Power Simulations	67
5.1 Methods	67
5.2 Simulation Results.....	69
5.3 Discussion.....	74
5.3.1 Key Question #1	74
Literature Cited	77
Appendix A.....	79

Figures

	Page
Figure 1-1. Locations of Northeast Coastal and Barrier Inventory & Monitoring (I&M) Network parks monitored for estuarine nutrient enrichment.	11
Figure 2-1. (Top): Boxplot of log surface chlorophyll- <i>a</i> concentration by year within Gateway National Recreation Area. Estuarine quality category thresholds (Table 2-1) are plotted as red dashed lines. (Bottom): Compact display of the multiple comparison test results	21
Figure 2-2. (Top): Boxplot of log bottom dissolved oxygen concentration by year within Gateway National Recreation Area. Estuarine quality category thresholds (Table 2-1) are plotted as red dashed lines. (Bottom): Compact display of the multiple comparison test results	23
Figure 2-3. (Top): Boxplot of log PAR attenuation by year within Gateway National Recreation Area. Estuarine quality category thresholds (Table 2-1) are plotted as red dashed lines. (Bottom): Compact display of the multiple comparison test results.	25
Figure 2-4. (Top): Boxplot of log surface salinity by year within Gateway National Recreation Area. (Bottom): Compact display of the multiple comparison test results.....	27
Figure 2-5. (Top): Boxplot of log surface temperature by year within Gateway National Recreation Area. Estuarine quality category thresholds (Table 2-1) are plotted as red dashed lines. (Bottom): Compact display of the multiple comparison test results	29
Figure 2-6. (Top): Boxplot of log surface turbidity by year within Gateway National Recreation Area. (Bottom): Compact display of the multiple comparison test results.....	31
Figure 2-7. Trends in Gateway National Recreation Area water quality measures based on estimated generalized additive model smooths.....	35
Figure 2-7 (continued). Trends in Gateway National Recreation Area water quality measures based on estimated generalized additive model smooths.....	36
Figure 2-8. Mean surface for the generalized additive model latitude/longitude smooth of surface chlorophyll- <i>a</i> concentration within Gateway National Recreation Area during 2003-2018.	37
Figure 2-9. Mean surface for the generalized additive model latitude/longitude smooth of bottom dissolved oxygen concentration within Gateway National Recreation Area during 2003-2018.	37
Figure 2-10. Mean surface for the generalized additive model latitude/longitude smooth of PAR attenuation within Gateway National Recreation Area during 2003-2018.....	38

Figure 2-11. Mean surface for the generalized additive model latitude/longitude smooth of surface salinity within Gateway National Recreation Area during 2003-2018.	38
Figure 2-12. Mean surface for the generalized additive model latitude/longitude smooth of surface temperature within Gateway National Recreation Area during 2003-2018.....	39
Figure 2-13. Mean surface for the generalized additive model latitude/longitude smooth of surface turbidity within Gateway National Recreation Area during 2003-2018.	39
Figure 3-1. Mean surface chlorophyll- <i>a</i> concentration ($\mu\text{g/L}$) trend at each Gateway National Recreation Area trend station based on estimated generalized additive model smooths	47
Figure 3-2. Mean bottom dissolved oxygen concentration (mg/L) trend at each Gateway National Recreation Area trend station based on estimated generalized additive model smooths.	48
Figure 3-3. Mean PAR attenuation ($K_d \text{ (m}^{-1}\text{)}$) trend at each Gateway National Recreation Area trend station based on estimated generalized additive model smooths	49
Figure 3-4. Mean surface salinity (psu) trend at each Gateway National Recreation Area trend station based on estimated generalized additive model smooths.....	50
Figure 3-5. Mean surface temperature ($^{\circ}\text{C}$) trend at each Gateway National Recreation Area trend station based on estimated generalized additive model smooths	51
Figure 3-6. Mean surface turbidity (NTU) trend at each Gateway National Recreation Area trend station based on estimated generalized additive model smooths	52
Figure 4-1. Continuous dissolved oxygen readings from sonde deployed at Gateway National Recreation Area logging station in 2012.....	54
Figure 4-2. Intra-annual trends in surface chlorophyll- <i>a</i> concentration (top), bottom dissolved oxygen concentration (middle), and PAR attenuation (bottom) at the Gateway National Recreation Area logging station based on estimated generalized additive model smooths.	58
Figure 4-3. Distribution of index period time series values for surface chlorophyll- <i>a</i> concentration (top), bottom dissolved oxygen concentration (middle), and PAR attenuation (bottom) at the Gateway National Recreation Area logging station	59
Figure 4-4. Intra-annual trends in surface salinity (top), surface temperature (middle), and surface turbidity (bottom) at the Gateway National Recreation Area logging station based on estimated generalized additive model smooths.....	60
Figure 4-5. Distribution of index period time series values for surface salinity (top), surface temperature (middle), and surface turbidity (bottom) at the Gateway National Recreation Area logging station.....	61

Figure 4-6. Intra-annual trends in pH (top) and depth (bottom) at the Gateway National Recreation Area logging station based on estimated generalized additive model smooths.....	62
Figure 4-7. Distribution of index period time series values for pH (top) and depth (bottom) at the Gateway National Recreation Area logging station.	63
Figure 4-8. Total hours per day that surface chlorophyll- <i>a</i> concentration was observed within each condition category at the Gateway National Recreation Area logging station.....	64
Figure 4-9. Total hours per day that bottom dissolved oxygen concentration was observed within each condition category at the Gateway National Recreation Area logging station.....	65
Figure 4-10. Total hours per day that PAR attenuation was observed within each condition category at the Gateway National Recreation Area logging station.	66
Figure 5-1. (A.) Process model realizations (black dots) representing samples from 30 sites within a hypothetical estuary.	68
Figure 5-2. Estimated power to detect a linear trend in mean surface chlorophyll- <i>a</i> concentration within Gateway National Recreation Area.....	69
Figure 5-3. Estimated power to detect a linear trend in mean bottom dissolved oxygen concentration within Gateway National Recreation Area.....	71
Figure 5-4. Estimated power to detect a linear trend in mean PAR attenuation within Gateway National Recreation Area.....	73
Figure A-1. (a) Data simulated from the nonlinear model given by Equation 2 and depicted in red. (b) Fitted generalized additive model smooth in black with corresponding 95% pointwise confidence interval in gray. True underlying curve provided in red for reference. (c) The six thin plate regression splines basis functions that make up the rank 6 basis expansion of $f(x)$ (black curve).	80

Tables

	Page
Table 1-1. Sampling years for water quality monitoring within Gateway National Recreation Area.	12
Table 2-1. Threshold values associated with each measure quality category.	13
Table 2-2. Area-weighted average water quality condition at Gateway National Recreation Area during the index periods of 2003-2020.	17
Table 2-3. Proportion of monitored estuarine resources falling into each quality category at Gateway National Recreation Area during the index periods of 2003-2020.	18
Table 2-4. Sampling start and end dates for each index period within Gateway National Recreation Area.	19
Table 2-5. Analysis of variance for the general linear mixed model fit to log-transformed water quality measurements collected within Gateway National Recreation Area.	20
Table 2-6. Summary of the multiple comparison test results for all possible yearly contrasts of mean log surface chlorophyll- <i>a</i> concentration within Gateway National Recreation Area	22
Table 2-7. Summary of the multiple comparison test results for all possible yearly contrasts of mean log bottom dissolved oxygen within Gateway National Recreation Area	24
Table 2-8. Summary of the multiple comparison test results for all possible yearly contrasts of mean log PAR attenuation within Gateway National Recreation Area.....	26
Table 2-9. Summary of the multiple comparison test results for all possible yearly contrasts of mean log surface salinity within Gateway National Recreation Area.....	28
Table 2-10. Summary of the multiple comparison test results for all possible yearly contrasts of mean log surface temperature within Gateway National Recreation Area.	30
Table 2-11. Summary of the multiple comparison test results for all possible yearly contrasts of mean log surface turbidity within Gateway National Recreation Area.....	32
Table 2-12. Approximate significance of the generalized additive model smooth terms for each model.....	34
Table 2-13. Proportion of null deviance explained by the explanatory variables for each model.....	34
Table 2-14. Summary of multivariate Mann-Kendall test for monotonic trend for each measure	40

Table 3-1. Yearly summary statistics for surface chlorophyll- <i>a</i> concentration ($\mu\text{g/L}$) at Gateway National Recreation Area trend stations	42
Table 3-2. Yearly summary statistics for bottom dissolved oxygen concentration (mg/L) at Gateway National Recreation Area trend stations	42
Table 3-3. Yearly summary statistics for PAR attenuation (K_d (m^{-1})) at Gateway National Recreation Area trend stations	43
Table 3-4. Yearly summary statistics for surface salinity (psu) at Gateway National Recreation Area trend stations	43
Table 3-5. Yearly summary statistics for surface temperature ($^{\circ}\text{C}$) at Gateway National Recreation Area trend stations	44
Table 3-6. Yearly summary statistics for surface turbidity (NTU) at Gateway National Recreation Area trend stations	44
Table 3-7. Approximate significance of the generalized additive model smooth terms for each trend station model	46
Table 3-8. Summary of multivariate Mann-Kendall test for monotonic trend in each trend station measure.....	53
Table 4-1. Yearly summary statistics for measures collected at the Gateway National Recreation Area logging station.....	57
Table 5-1. Process model dispersion (ϕ) used to simulate trends in each water quality measure.	68
Table 5-2. Estimated power to detect a linear trend in mean surface chlorophyll- <i>a</i> concentration large enough to cross both quality thresholds within Gateway National Recreation Area	70
Table 5-3. Estimated power to detect a linear trend in mean bottom dissolved oxygen concentration large enough to cross both quality thresholds within Gateway National Recreation Area	72
Table 5-4. Estimated power to detect a linear trend in mean PAR attenuation large enough to cross both quality thresholds within Gateway National Recreation Area.....	74

Acronyms and Abbreviations

ENE	Estuarine nutrient enrichment
GAM	Generalized additive model
GATE	Gateway National Recreation Area
NCBN	Northeast Coastal and Barrier Network
NTU	Nephelometric turbidity units
PAR	Photosynthetically active radiation
Psu	Practical salinity units
SAV	Submerged aquatic vegetation

1. Background Information

Six parks within the Northeast Coastal and Barrier Network (NCBN) contain critical estuarine habitat (Figure 1-1). Roman et al. (2000) identified nutrient enrichment as the primary threat to the ecological integrity of estuaries within this region. Additionally, threats to coastal water quality is one of the highest priority issues identified by park managers (Stevens et al. 2005). In response, B. S. Kopp and Neckles (2009) developed a regional protocol for monitoring estuarine nutrient enrichment (ENE) that was fully implemented in NCBN parks beginning in 2007. The vital signs selected for monitoring include: estuarine water chemistry, estuarine water quality, estuarine water clarity, seagrass distribution, and seagrass condition. Previous reports by Blaine S Kopp et al. (2009) and Caldwell et al. (2015) summarized water quality condition and trends during 2003-2006 and 2006-2011, respectively. The current report summarizes conditions and trends for water quality data collected in Gateway National Recreation Area to date (Table 1-1).

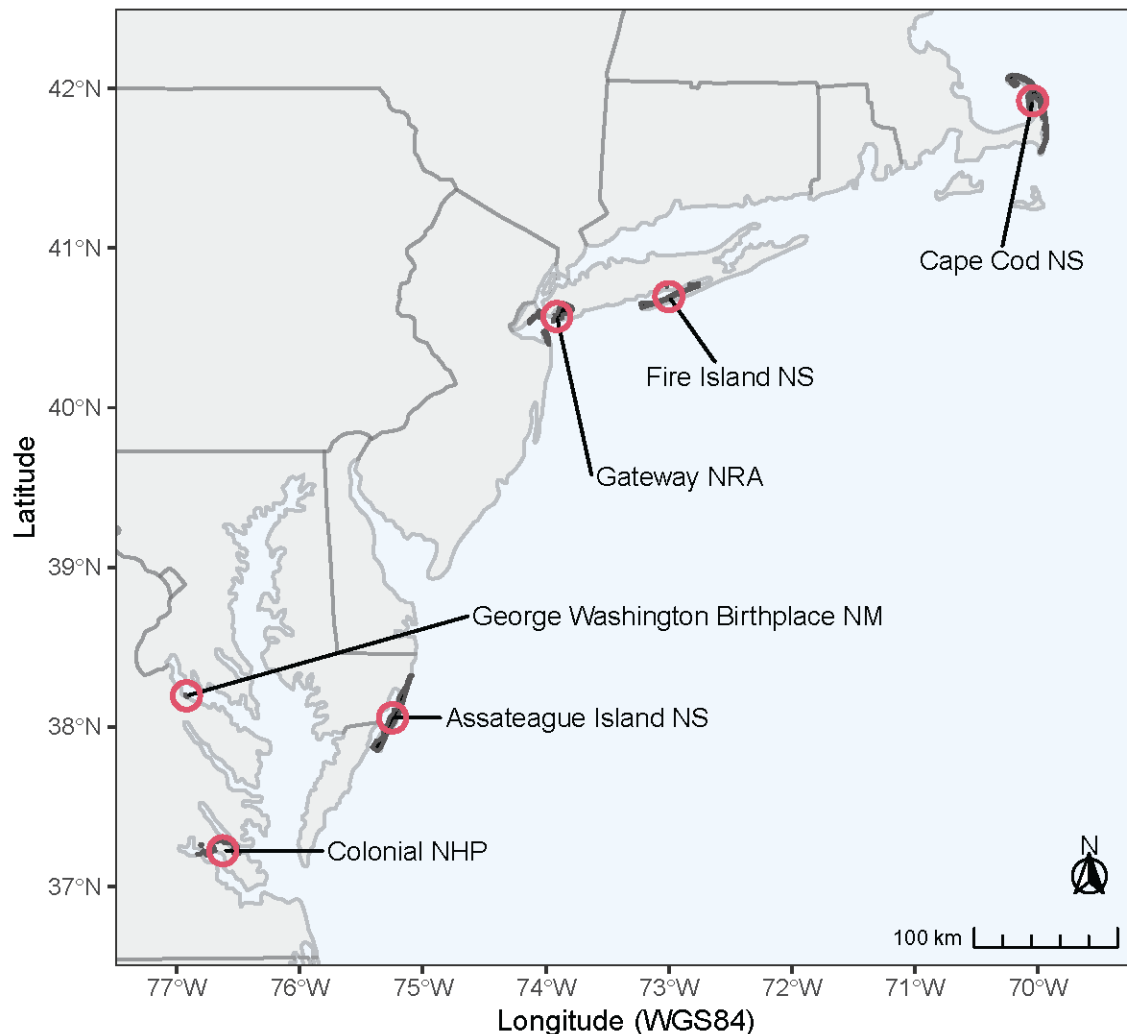


Figure 1-1. Locations of Northeast Coastal and Barrier Inventory & Monitoring (I&M) Network parks monitored for estuarine nutrient enrichment.

Table 1-1. Sampling years for water quality monitoring within Gateway National Recreation Area.

Year	Water Quality
2003	sampled
2004	not sampled
2005	sampled
2006	sampled
2007	not sampled
2008	not sampled
2009	not sampled
2010	sampled
2011	not sampled
2012	sampled
2013	not sampled
2014	sampled
2015	not sampled
2016	sampled
2017	not sampled
2018	sampled
2019	not sampled
2020	sampled

2. Spatial-Station Water Quality Data

2.1 Methods

2.1.1 Area-Weighted Condition

Water quality within each estuary is characterized by the area-weighted mean of each measure. For estuaries containing multiple strata, summaries are computed for each stratum and on a park-wide basis. The area-weighted mean ($\bar{y}$) is computed as:

$$\bar{y} = \sum_{i=1}^n w_i \cdot y_i,$$

where y_i represents the measurement within the i^{th} hexagon and w_i is the percent of the total estuarine area contained within the hexagon. The area-weighted mean is computed separately for each measured depth, as applicable. Measurements from the first successful sampling event are used for hexagons containing trend stations. Despite several parks having existing monitoring programs prior to the implementation of the protocol, no hexagons contained multiple trend stations in the provided data set.

Condition is also characterized by the percent of estuarine area within quality categories for a subset of measures. Thresholds for chlorophyll-*a*, dissolved oxygen, and water clarity are defined in B. S. Kopp and Neckles (2009) and temperature thresholds are based on eelgrass (Shields, Parrish, and Moore 2019) and blue crab (Bauer and Miller 2010a, 2010b) tolerance (Table 2-1). The proportion of estuarine area in quality category j (p_j) is computed as:

$$p_j = \sum_{i=1}^n w_i \cdot 1_{ij},$$

where 1_{ij} is an indicator for whether the measurement within the i^{th} hexagon falls within the category j values and w_i is the percent of the total estuarine area contained within the hexagon. The proportion of estuarine area containing missing observations is also provided. All computations were carried out using the R Statistical Computing Environment (R Core Team 2020).

Table 2-1. Threshold values associated with each measure quality category.

Measure	Quality Category*		
	Good	Fair	Poor
Chlorophyll- <i>a</i> (near surface)	< 5 $\mu\text{g/L}$	5-20 $\mu\text{g/L}$	> 20 $\mu\text{g/L}$
Dissolved oxygen (near bottom)	> 5 mg/L	2-5 mg/L	< 2 mg/L
PAR attenuation K_d (SAV standard)	< 0.92 m^{-1}	< 0.92-1.61 m^{-1}	> 1.61 m^{-1}
Water temperature (near surface)	< 25 $^{\circ}\text{C}$	25-28 $^{\circ}\text{C}$	> 28 $^{\circ}\text{C}$

*Good/Fair/Poor is not applicable to surface water temperature.

2.1.2 General Linear Mixed Model (GLMM)

Consistent with the previous report by Caldwell et al. (2015), a general linear mixed model is used to determine whether there are differences in water quality measures among index periods. Year is treated as a categorical fixed effect and individual hexagons are included as random intercepts. The fitted model is used to test the hypothesis: H_0 : the mean measure value is the same across years versus H_1 : at least one of the year's means is different. If the null hypothesis is rejected at the 5% significance level (i.e., p -value < 0.05), then a multiple comparison test (MCT) is performed to determine which pairs of years have statistically different means based on Tukey-adjusted p -values¹.

Residual analysis indicated that the homogeneity of variance assumption is oftentimes violated. Therefore, each measure is log-transformed prior to fitting the model and the results of the MCT are presented on the log scale. Interpretation of the MCTs on the original scale should be reported as follows: it is estimated that the median for year t_0 is $\exp(\bar{Z}_0 - \bar{Z}_1)$ times as large as the median for year t_1 , where $\bar{Z}_0$ and $\bar{Z}_1$ represent the means of the log-transformed values for years t_0 and t_1 , respectively (Ramsey and Schafer 2012 Chapter 3.5).

For example, consider the comparison of PAR attenuation between years 2005 and 2006 (Figure 2-3). In this matrix, the values for Z_{-t} are provided on the diagonal and the differences ($Z_{-i} - Z_{-j}$, $i \neq j$) are located on the upper diagonal entries. Note: a more detailed summary of the test statistics and confidence intervals associated with Figure 2-3 are provided in Table 2-8. For this comparison we would state: it is estimated that the median PAR attenuation in GATE during 2005 is $\exp(-0.412) \approx 0.66$ times the median during 2006 (95% CI equal to (0.48, 0.92)). Stated another way: with 95% confidence, it is estimated that the median PAR attenuation in GATE during 2005 is between 0.48 and 0.92 times the median in 2006. Analyses were carried out using the R Statistical Computing Environment (R Core Team 2020). The nlme package (Pinheiro et al. 2020) is used to fit the models and multiple comparison tests are carried out using the emmeans package (Lenth 2020).

2.1.3 Generalized Additive Model (GAM)

Based on initial exploratory analyses, several measures display potential nonlinear trends over time. Generalized additive models (GAMs; Wood 2017) are used to estimate trends in water quality measures as smooth functions over time. We refer the reader to Appendix A for a short primer on GAMs if desired.

Each water quality measure is modeled using a gamma distribution with log link function. This is a commonly used distribution when data are non-negative, continuous, and potentially right-skewed. A major benefit of this model over applying a log transformation to data is that we are modeling the mean of the measure on the response scale, rather than the transformed scale. This in turn simplifies interpretation of the estimated model parameters and plots of the GAM smooths represent the mean

¹ See Appendix A for additional discussion of MCT family-wise error rate and p -value adjustments.

response (as opposed to the geometric mean²). Additionally, the log link has the added benefit of ensuring that simulated values used in power simulations will be non-negative.

For each measure, we model the mean of response as a function of both time and location within the estuary. Mathematically, the model can be written as:

$$\begin{aligned} \text{response}_i &\sim \text{Gamma}(\mu_i, \phi) \\ \log(\mu_i) &= \beta_0 + f(\text{year}_i) + f(\text{longitude}_i, \text{latitude}_i), \end{aligned}$$

where $i = 1, \dots, n$ indexes the respective hexagon within the estuary. Thin plate regression splines (TPRS; rank of basis expansion equal to number of years of data) and Gaussian process splines (rank 50 basis expansion) were utilized for the year and location smooths, respectively. The spatial dependence of observations is modeled in the mean structure (i.e., first-order model specification³) through the inclusion of the interaction smooth between longitude and latitude. Attempts to model temporal autocorrelation in a majority of cases resulted in unidentifiable models. Factors likely contributing to this include: (1) the relatively short time series, (2) irregular intervals of data collection, and (3) insufficient information in the data to uniquely distinguish between the wiggly trend and autocorrelation since they can both be used to describe similar patterns in the data (Simpson 2018). Therefore, for consistency across all models, temporal autocorrelation was not modeled. Residual analysis indicated a lack of temporal autocorrelation in a majority of measures. For measures that did exhibit significant temporal autocorrelation, the lag-1 autocorrelation was generally less than 0.4.

The first derivative is approximated along the length of the year smooth using the finite difference method. Periods of significant temporal change are identified as intervals of time wherein the 95% confidence interval of the first derivative (i.e., slope of tangent line) does not contain the value of zero.

Analyses were carried out using the R Statistical Computing Environment (R Core Team 2020). Each model is fit using the `mgcv` package (Wood 2011) and results were summarized/visualized using the `gratia` package (Simpson 2020).

² The geometric mean is defined as the n th root of the product of n numbers. For example, the geometric mean of the numbers 2, 5, and 7 is equal to $(2 \cdot 5 \cdot 7)^{1/3} = 70^{1/3} \approx 4.12$.

³ See Hefley et al. (2017) for additional discussion of first-order and second-order model specifications for modeling autocorrelation in ecological and environmental data.

2.1.4 Multivariate Mann-Kendall Trend Analysis

Despite several measures exhibiting non-linear trends over time, it is still of interest to quantify whether measures are increasing/decreasing over time. The multivariate Mann-Kendall test (trend R package; (Pohlert 2020)) is used to determine whether there exists a monotonically increasing or decreasing trend in values for all sites within the estuary over time (Libiseller and Grimvall 2002). Positive (negative) values of the test statistic Z_{MK} indicates that measurements tend to increase (decrease) with time. For example, based on the summary of tests provided in Table 2-14, we'd conclude that there is evidence of decreasing measurements over time for both bottom dissolved oxygen ($Z_{MK} = -2.24$; p -value = 0.025) and surface turbidity ($Z_{MK} = -2.29$; p -value = 0.022).

2.2 Area-Weighted Condition Results

Sampling start and end dates for each index period are provided in Table 2-4. Area-weighted condition summaries for each measure are as follows:

2.2.1 Chlorophyll-*a* (near surface)

Area-weighted average near surface chlorophyll-*a* concentration ranged from 21.33 to 48.54 $\mu\text{g/L}$ and was poor (i.e., $> 20 \mu\text{g/L}$) in every sampled year (Table 2-2). The proportion of the estuary characterized as being of poor quality ranged from 44.7 to 89.9% (Table 2-3).

2.2.2 Dissolved Oxygen (near bottom)

Area-weighted average bottom dissolved oxygen concentration was good (i.e., $> 5 \text{ mg/L}$) for all sampled years and ranged from 4.99 to 7.77 mg/L . The majority of area (at least 70%) within the estuary was characterized as in good condition except for years 2006, 2014, and 2020. In 2006, approximately 18% of the area within the estuary was characterized as being poor.

2.2.3 PAR Attenuation K_d (SAV standard)

Area-weighted average PAR attenuation ranged from 0.94 to 1.67 m^{-1} , which were measured in consecutive years (2005 and 2006). Except for 2005, more than 60% of the estuary was characterized as being either fair or poor quality during each sampling period.

2.2.4 pH

Area-weighted average pH ranged from 7.75 to 8.07. Measurements of pH began in 2018.

2.2.5 Salinity

Excluding the sampling year 2005 (which appears anomalous), area-weighted average salinity ranged from 25.66 (surface depth in 2003) to 28.72 psu (bottom depth in 2016).

2.2.6 Temperature (near surface)

Area-weighted average near surface water temperature ranged from 23.89 to 25.51 $^{\circ}\text{C}$, and exceeded 25 $^{\circ}\text{C}$ during the last two sampling years (2018 and 2020). Temperatures in the 25-28 $^{\circ}\text{C}$ range were observed in more than 50% of the estuary in all sampled years except 2003, 2014, and 2016. In 2018, approximately 67% of the estuary had observed temperatures in the 25-28 $^{\circ}\text{C}$ range.

2.2.7 Turbidity

Area-weighted average turbidity ranged from 2.14 (surface in 2018) to 19.21 NTU (bottom depth in 2003). For all years except 2003, 2005, and 2020, turbidity is approximately constant across depths.

Table 2-2. Area-weighted average water quality condition at Gateway National Recreation Area during the index periods of 2003-2020.

Measure	Depth	Sample Year								
		2003	2005	2006	2010	2012	2014	2016	2018	2020
Chlorophyll <i>a</i> ($\mu\text{g/L}$)	Surface	36.87	26.61	42.12	48.54	34.18	22.73	21.33	25.18	30.55
Dissolved oxygen (% saturation)	Surface	114.52	101.09	112.23	119.82	122.58	108.70	107.24	126.97	91.84
Dissolved oxygen (% saturation)	Mid-Depth	109.51	82.18	86.97	103.40	98.57	86.13	90.33	100.48	73.52
Dissolved oxygen (% saturation)	Bottom	107.15	90.62	77.56	92.09	92.54	88.40	91.89	93.99	70.65
Dissolved oxygen (mg/L)	Surface	8.32	8.22	7.91	8.44	8.43	7.84	7.52	8.86	6.43
Dissolved oxygen (mg/L)	Mid-Depth	7.92	5.82	6.19	7.31	7.05	6.28	6.40	7.08	5.20
Dissolved oxygen (mg/L)	Bottom	7.77	7.04	5.54	6.54	6.61	6.46	6.48	6.65	4.99
PAR attenuation (K_d (m^{-1}))	Water Column	1.08	0.94	1.67	1.62	1.27	1.16	1.10	1.08	1.40
pH	Surface	NA	NA	NA	NA	NA	NA	NA	8.07	7.80
pH	Mid-Depth	NA	NA	NA	NA	NA	NA	NA	7.94	7.78
pH	Bottom	NA	NA	NA	NA	NA	NA	NA	7.90	7.75
Salinity (psu)	Surface	25.66	6.26	26.65	27.96	27.06	26.14	28.25	27.64	27.83
Salinity (psu)	Mid-Depth	27.12	26.76	27.30	28.29	27.46	26.62	28.66	28.02	28.59
Salinity (psu)	Bottom	27.44	15.37	27.49	28.36	27.71	26.02	28.72	27.98	28.38
Temperature ($^{\circ}\text{C}$)	Surface	23.89	24.34	25.15	24.97	24.91	24.06	24.98	25.51	25.42
Temperature ($^{\circ}\text{C}$)	Mid-Depth	23.73	25.24	24.56	24.63	24.26	23.37	24.40	24.85	24.66
Temperature ($^{\circ}\text{C}$)	Bottom	23.49	24.61	24.32	24.34	23.89	23.47	24.21	24.51	24.78
Turbidity (NTU)	Surface	2.96	13.91	3.55	3.02	2.92	2.88	2.51	2.14	3.05
Turbidity (NTU)	Mid-Depth	4.07	3.18	3.44	2.77	2.20	2.49	2.39	2.20	3.68
Turbidity (NTU)	Bottom	19.21	9.06	3.36	2.92	2.75	3.70	2.55	2.78	7.08

Table 2-3. Proportion of monitored estuarine resources falling into each quality category at Gateway National Recreation Area during the index periods of 2003-2020.

Measure	Condition	Sampling Year								
		2003	2005	2006	2010	2012	2014	2016	2018	2020
Bottom dissolved oxygen (mg/L)	Good	94.8	84.5	44.6	71.9	79.8	60.7	77.3	76.8	52.6
Bottom dissolved oxygen (mg/L)	Fair	5.2	15.5	37.2	21.9	17.8	39.3	18.0	15.0	44.7
Bottom dissolved oxygen (mg/L)	Poor	0.0	0.0	18.2	6.2	2.3	0.0	4.7	8.2	2.7
Bottom dissolved oxygen (mg/L)	Missing	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Chlorophyll a ($\mu\text{g/L}$)	Good	14.8	5.4	9.8	0.0	3.4	10.2	0.0	11.9	0.0
Chlorophyll a ($\mu\text{g/L}$)	Fair	24.2	26.9	24.0	12.2	6.6	45.1	46.5	39.1	43.7
Chlorophyll a ($\mu\text{g/L}$)	Poor	61.1	59.2	66.2	87.8	89.9	44.7	53.5	49.0	56.3
Chlorophyll a ($\mu\text{g/L}$)	Missing	0.0	8.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PAR attenuation (K_d (m^{-1}))	Good	40.5	56.0	17.8	14.6	14.4	25.5	15.3	26.8	3.4
PAR attenuation (K_d (m^{-1}))	Fair	59.5	37.9	56.8	33.1	66.0	61.8	82.7	72.1	72.5
PAR attenuation (K_d (m^{-1}))	Poor	0.0	6.1	25.4	52.2	15.4	12.7	2.0	1.0	24.1
PAR attenuation (K_d (m^{-1}))	Missing	0.0	0.0	0.0	0.0	4.2	0.0	0.0	0.0	0.0
Surface Temperature ($^{\circ}\text{C}$)	<25	100.0	46.5	33.8	37.8	42.0	74.1	52.6	31.3	43.2
Surface Temperature ($^{\circ}\text{C}$)	25-28	0.0	53.5	66.2	60.6	58.0	25.9	45.2	66.9	56.8
Surface Temperature ($^{\circ}\text{C}$)	>28	0.0	0.0	0.0	1.6	0.0	0.0	2.2	1.8	0.0
Surface Temperature ($^{\circ}\text{C}$)	Missing	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table 2-4. Sampling start and end dates for each index period within Gateway National Recreation Area.

Year	Start	End
2003	2003-07-24	2003-07-24
2005	2005-07-29	2005-08-24
2006	2006-07-27	2006-08-23
2010	2010-06-22	2010-07-12
2012	2012-07-02	2012-07-24
2014	2014-07-02	2014-07-25
2016	2016-07-05	2016-07-26
2018	2018-07-03	2018-07-24
2020	2020-08-05	2020-08-24

2.3 GLMM Results

For each measure, the analysis of variance based on the fitted GLMM indicates that there is at least one year that is statistically different from the others, assuming a significance level equal to 0.05 (Table 2-5). Notable observations from the multiple comparison tests (MCT) for each measure are as follows

2.3.1 Chlorophyll-*a* (near surface)

The MCT results indicate that mean log near surface chlorophyll-*a* concentrations in 2010 were statistically different from most other index periods (Figure 2-1; Table 2-6).

2.3.2 Dissolved Oxygen (near bottom)

The MCT results indicate that mean log bottom dissolved oxygen in 2005 is statistically different compared to 2020 (Figure 2-2; Table 2-7). There is no strong evidence of statistical differences between the other sampled year pairs.

2.3.3 PAR Attenuation K_d (SAV standard)

The MCT results indicate that mean log PAR attenuation in 2006 and 2010 are statistically different compared to most other years (Figure 2-3; Table 2-8). Additionally, 2020 is statistically different from both 2005 and 2018.

2.3.4 Salinity (near surface)

The MCT results indicate that mean log surface salinity in 2014 is statistically different compared to 2014 and all more recent sampling events (Figure 2-4; Table 2-9). There is no strong evidence of statistical differences between the other sampled year pairs.

2.3.5 Temperature (near surface)

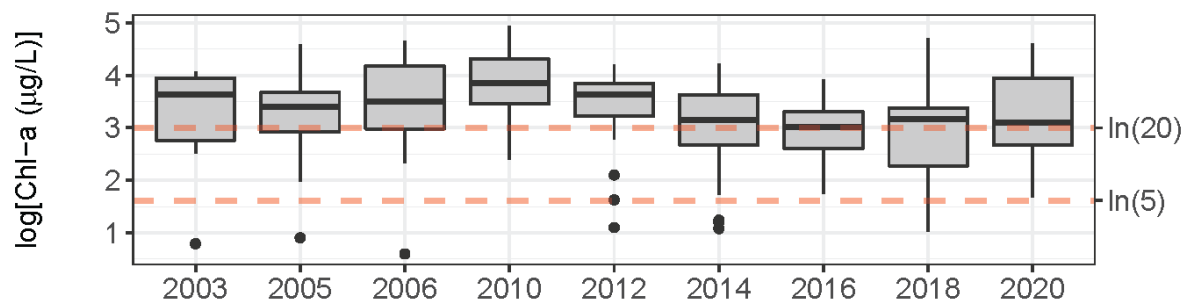
The MCT results indicate that mean log surface water temperature in 2014 is statistically different from most other index periods (Figure 2-5; Table 2-10). Additionally, 2005 is statistically different from both 2018 and 2020.

2.3.6 Turbidity (near surface)

The MCT results indicate that mean log surface turbidity in 2005 is statistically different compared to all other sampled years (Figure 2-6; Table 2-11), and is estimated to be higher in each instance. Additionally, 2018 is statistically different from 2010.

Table 2-5. Analysis of variance for the general linear mixed model fit to log-transformed water quality measurements collected within Gateway National Recreation Area. Each respective F -statistic and p -value corresponds to the test of the hypotheses: H_0 : the mean of the log-transformed measure values are the same in each year versus H_1 : at least one of the year's means is different. Hyperlinked references will re-direct to the results of the multiple comparison tests (MCT), if applicable.

Measure	Depth(s)	D.F.	F -value	p -value	MCT Results
Chlorophyll a ($\mu\text{g/L}$)	Surface	(8, 177)	6.09	< 0.005	Figure 2-1
Dissolved oxygen (mg/L)	Bottom	(8, 179)	2.09	0.04	Figure 2-2
PAR attenuation (K_d (m^{-1}))	Water Column	(8, 178)	7.35	< 0.005	Figure 2-3
Salinity (psu)	Surface	(7, 158)	3.25	< 0.005	Figure 2-4
Temperature ($^{\circ}\text{C}$)	Surface	(8, 179)	4.97	< 0.005	Figure 2-5
Turbidity (NTU)	Surface	(8, 178)	16.47	< 0.005	Figure 2-6



2020	0.112	-0.011	0.144	0.580	0.117	-0.264	-0.297	-0.317	3.270
2018	0.428	0.305	0.460	0.896	0.434	0.053	0.019	2.950	0.503
2016	0.409	0.286	0.441	0.877	0.415	0.033	2.970	1.000	0.590
2014	0.376	0.253	0.408	0.844	0.381	3.000	1.000	1.000	0.734
2012	-0.006	-0.129	0.026	0.462	3.380	0.248	0.157	0.116	0.998
2010	-0.468	-0.591	-0.436	3.850	0.085	0.000	0.000	0.000	0.009
2006	-0.032	-0.155	3.410	0.540	1.000	0.621	0.515	0.455	0.999
2005	0.123	3.260	0.999	0.028	0.998	0.875	0.779	0.712	1.000
2003	3.380	1.000	1.000	0.753	1.000	0.910	0.862	0.829	1.000
	2003	2005	2006	2010	2012	2014	2016	2018	2020

p-value 0.00 0.25 0.50 0.75 1.00

Figure 2-1. (Top): Boxplot of log surface chlorophyll-a concentration by year within Gateway National Recreation Area. Estuarine quality category thresholds (Table 2-1) are plotted as red dashed lines. (Bottom): Compact display of the multiple comparison test results. The bolded diagonal elements contain the mean log surface chlorophyll-a concentration within each year. Upper diagonal cells contain the difference between the means of the years indicated by the row and column names (earlier minus latter). Lower diagonals contain the Tukey-adjusted p-value from a two-sided test that the means of the years indicated by the row and column names are equal (i.e., $H_0: \mu_1 = \mu_2$ versus $H_1: \mu_1 \neq \mu_2$). A more detailed summary of the test statistics and confidence intervals is provided in Table 2.6.

Table 2-6. Summary of the multiple comparison test results for all possible yearly contrasts of mean log surface chlorophyll-a concentration within Gateway National Recreation Area. Note: all estimates and confidence interval bounds are reported on the log scale. [SE = standard error; CI = confidence interval]

Contrast	Estimate	SE	t_{177}	p-value	95% CI
2003 - 2005	0.123	0.288	0.427	1.000	(-0.774, 1.020)
2003 - 2006	-0.032	0.310	-0.102	1.000	(-0.996, 0.933)
2003 - 2010	-0.468	0.280	-1.669	0.753	(-1.341, 0.405)
2003 - 2012	-0.006	0.279	-0.020	1.000	(-0.874, 0.863)
2003 - 2014	0.376	0.279	1.346	0.910	(-0.493, 1.244)
2003 - 2016	0.409	0.279	1.465	0.862	(-0.460, 1.278)
2003 - 2018	0.428	0.279	1.535	0.829	(-0.440, 1.297)
2003 - 2020	0.112	0.279	0.401	1.000	(-0.757, 0.981)
2005 - 2006	-0.155	0.230	-0.675	0.999	(-0.870, 0.560)
2005 - 2010	-0.591	0.179	-3.311	0.028	(-1.147, -0.036)
2005 - 2012	-0.129	0.176	-0.732	0.998	(-0.676, 0.419)
2005 - 2014	0.253	0.176	1.436	0.875	(-0.295, 0.800)
2005 - 2016	0.286	0.176	1.624	0.779	(-0.262, 0.833)
2005 - 2018	0.305	0.176	1.734	0.712	(-0.242, 0.853)
2005 - 2020	-0.011	0.176	-0.064	1.000	(-0.559, 0.536)
2006 - 2010	-0.436	0.220	-1.986	0.540	(-1.120, 0.248)
2006 - 2012	0.026	0.218	0.120	1.000	(-0.653, 0.705)
2006 - 2014	0.408	0.218	1.869	0.622	(-0.271, 1.086)
2006 - 2016	0.441	0.218	2.021	0.516	(-0.238, 1.119)
2006 - 2018	0.460	0.218	2.110	0.454	(-0.219, 1.139)
2006 - 2020	0.144	0.218	0.658	0.999	(-0.535, 0.822)
2010 - 2012	0.462	0.158	2.920	0.085	(-0.030, 0.955)
2010 - 2014	0.844	0.158	5.328	0.000	(0.351, 1.337)
2010 - 2016	0.877	0.158	5.538	0.000	(0.384, 1.370)
2010 - 2018	0.896	0.158	5.660	0.000	(0.403, 1.389)
2010 - 2020	0.580	0.158	3.662	0.009	(0.087, 1.073)
2012 - 2014	0.381	0.155	2.457	0.248	(-0.102, 0.865)
2012 - 2016	0.415	0.155	2.671	0.157	(-0.069, 0.898)
2012 - 2018	0.434	0.155	2.796	0.117	(-0.049, 0.917)
2012 - 2020	0.117	0.155	0.757	0.998	(-0.366, 0.601)
2014 - 2016	0.033	0.155	0.214	1.000	(-0.450, 0.516)
2014 - 2018	0.053	0.155	0.339	1.000	(-0.431, 0.536)
2014 - 2020	-0.264	0.155	-1.700	0.734	(-0.747, 0.219)
2016 - 2018	0.019	0.155	0.125	1.000	(-0.464, 0.502)
2016 - 2020	-0.297	0.155	-1.914	0.590	(-0.780, 0.186)
2018 - 2020	-0.317	0.155	-2.039	0.503	(-0.800, 0.167)

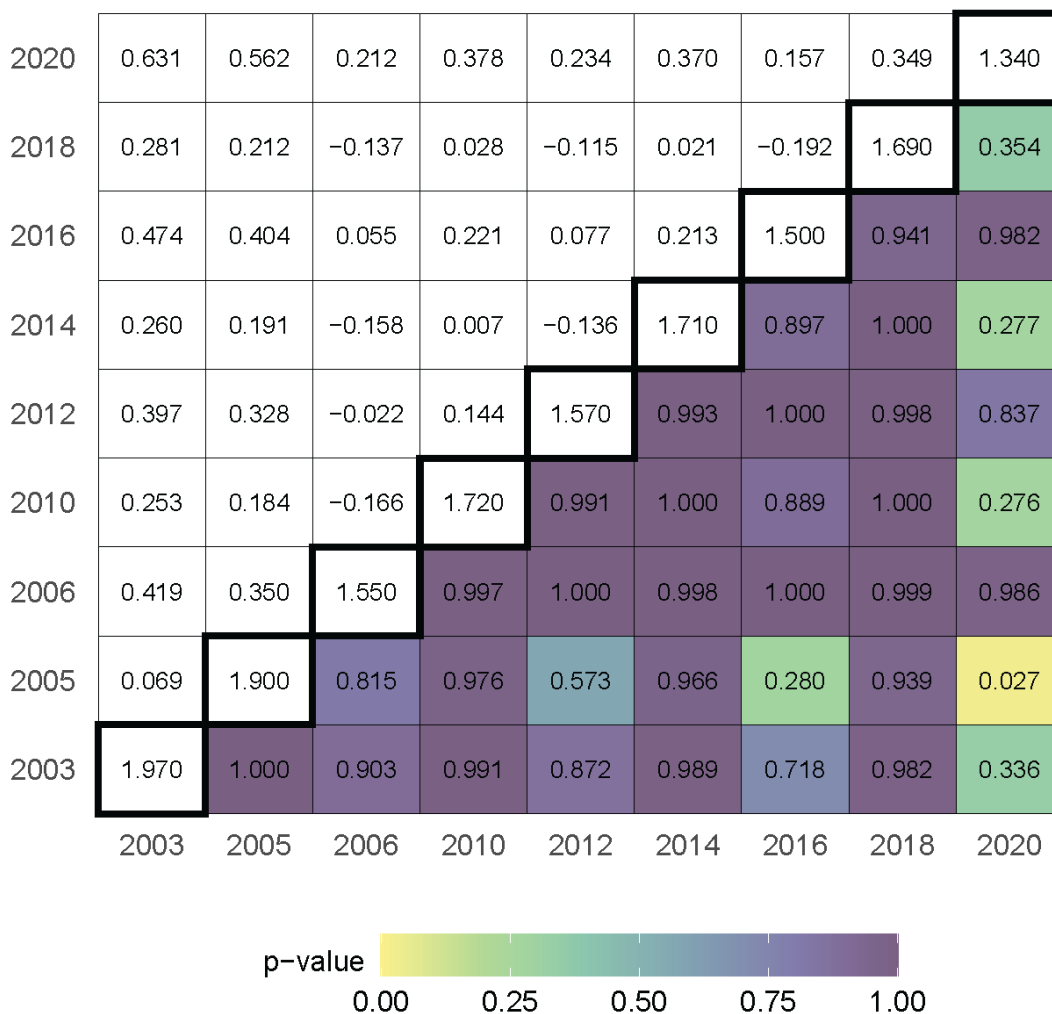
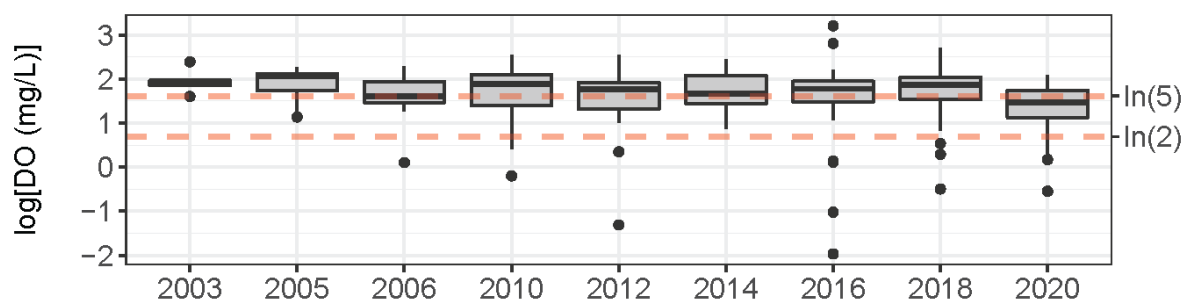
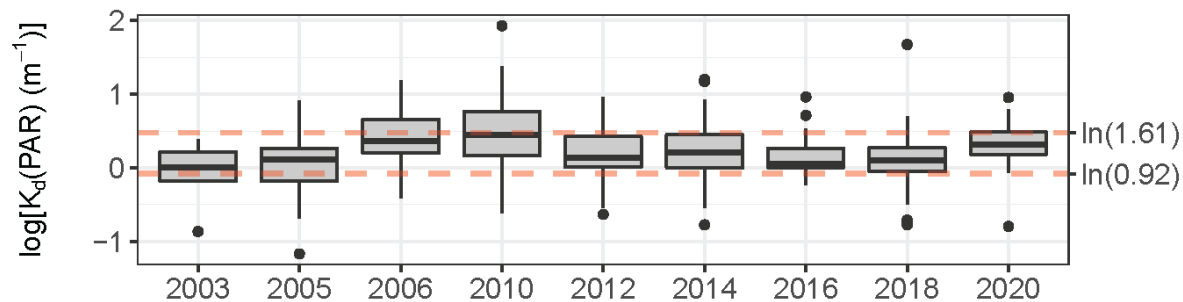


Figure 2-2. (Top): Boxplot of log bottom dissolved oxygen concentration by year within Gateway National Recreation Area. Estuarine quality category thresholds (Table 2-1) are plotted as red dashed lines. (Bottom): Compact display of the multiple comparison test results. The bolded diagonal elements contain the mean log bottom dissolved oxygen concentration within each year. Upper diagonal cells contain the difference between the means of the years indicated by the row and column names (earlier minus latter). Lower diagonals contain the Tukey-adjusted p-value from a two-sided test that the means of the years indicated by the row and column names are equal (i.e., $H_0: \mu_1 = \mu_2$ versus $H_1: \mu_1 \neq \mu_2$). A more detailed summary of the test statistics and confidence intervals is provided in Table 2-7.

Table 2-7. Summary of the multiple comparison test results for all possible yearly contrasts of mean log bottom dissolved oxygen within Gateway National Recreation Area. Note: all estimates and confidence interval bounds are reported on the log scale. [SE = standard error; CI = confidence interval]

Contrast	Estimate	SE	t_{179}	p -value	95% CI
2003 - 2005	0.069	0.282	0.245	1.000	(-0.809, 0.947)
2003 - 2006	0.419	0.307	1.365	0.904	(-0.538, 1.375)
2003 - 2010	0.253	0.276	0.916	0.991	(-0.608, 1.114)
2003 - 2012	0.397	0.275	1.445	0.871	(-0.460, 1.253)
2003 - 2014	0.260	0.275	0.948	0.989	(-0.596, 1.117)
2003 - 2016	0.474	0.275	1.725	0.718	(-0.383, 1.330)
2003 - 2018	0.281	0.275	1.024	0.982	(-0.575, 1.138)
2003 - 2020	0.631	0.275	2.297	0.336	(-0.226, 1.487)
2005 - 2006	0.350	0.224	1.562	0.814	(-0.349, 1.048)
2005 - 2010	0.184	0.171	1.071	0.976	(-0.351, 0.719)
2005 - 2012	0.328	0.169	1.939	0.573	(-0.199, 0.854)
2005 - 2014	0.191	0.169	1.132	0.966	(-0.336, 0.718)
2005 - 2016	0.404	0.169	2.394	0.281	(-0.122, 0.931)
2005 - 2018	0.212	0.169	1.256	0.938	(-0.315, 0.739)
2005 - 2020	0.562	0.169	3.325	0.027	(0.035, 1.089)
2006 - 2010	-0.166	0.217	-0.766	0.997	(-0.842, 0.510)
2006 - 2012	-0.022	0.215	-0.103	1.000	(-0.692, 0.648)
2006 - 2014	-0.158	0.215	-0.737	0.998	(-0.829, 0.512)
2006 - 2016	0.055	0.215	0.255	1.000	(-0.615, 0.725)
2006 - 2018	-0.137	0.215	-0.640	0.999	(-0.808, 0.533)
2006 - 2020	0.212	0.215	0.987	0.986	(-0.458, 0.882)
2010 - 2012	0.144	0.157	0.915	0.991	(-0.347, 0.634)
2010 - 2014	0.007	0.157	0.047	1.000	(-0.483, 0.498)
2010 - 2016	0.221	0.157	1.404	0.889	(-0.270, 0.711)
2010 - 2018	0.028	0.157	0.181	1.000	(-0.462, 0.519)
2010 - 2020	0.378	0.157	2.404	0.276	(-0.113, 0.868)
2012 - 2014	-0.136	0.154	-0.884	0.993	(-0.617, 0.345)
2012 - 2016	0.077	0.154	0.499	1.000	(-0.404, 0.558)
2012 - 2018	-0.115	0.154	-0.748	0.998	(-0.597, 0.366)
2012 - 2020	0.234	0.154	1.518	0.837	(-0.247, 0.715)
2014 - 2016	0.213	0.154	1.383	0.897	(-0.268, 0.694)
2014 - 2018	0.021	0.154	0.136	1.000	(-0.460, 0.502)
2014 - 2020	0.370	0.154	2.402	0.276	(-0.111, 0.852)
2016 - 2018	-0.192	0.154	-1.247	0.941	(-0.673, 0.289)
2016 - 2020	0.157	0.154	1.019	0.982	(-0.324, 0.638)
2018 - 2020	0.349	0.154	2.266	0.354	(-0.132, 0.831)



2020	-0.229	-0.279	0.133	0.169	-0.152	-0.103	-0.183	-0.236	0.330
2018	0.007	-0.043	0.370	0.405	0.084	0.134	0.053	0.094	0.028
2016	-0.046	-0.096	0.316	0.352	0.031	0.080	0.148	0.998	0.202
2014	-0.127	-0.176	0.236	0.272	-0.049	0.228	0.968	0.624	0.876
2012	-0.077	-0.127	0.285	0.321	0.179	0.999	1.000	0.960	0.460
2010	-0.399	-0.448	-0.036	0.500	0.001	0.007	0.000	0.000	0.323
2006	-0.363	-0.412	0.464	1.000	0.110	0.308	0.046	0.009	0.918
2005	0.049	0.052	0.003	0.000	0.795	0.372	0.949	1.000	0.014
2003	0.101	1.000	0.211	0.056	1.000	0.986	1.000	1.000	0.682
	2003	2005	2006	2010	2012	2014	2016	2018	2020



Figure 2-3. (Top): Boxplot of log PAR attenuation by year within Gateway National Recreation Area. Estuarine quality category thresholds (Table 2-1) are plotted as red dashed lines. (Bottom): Compact display of the multiple comparison test results. The bolded diagonal elements contain the mean log PAR attenuation within each year. Upper diagonal cells contain the difference between the means of the years indicated by the row and column names (earlier minus latter). Lower diagonals contain the Tukey-adjusted p -value from a two-sided test that the means of the years indicated by the row and column names are equal (i.e., $H_0: \mu_1 = \mu_2$ versus $H_1: \mu_1 \neq \mu_2$). A more detailed summary of the test statistics and confidence intervals is provided in Table 2-8.

Table 2-8. Summary of the multiple comparison test results for all possible yearly contrasts of mean log PAR attenuation within Gateway National Recreation Area. Note: all estimates and confidence interval bounds are reported on the log scale. [SE = standard error; CI = confidence interval]

Contrast	Estimate	SE	<i>t</i> ₁₇₈	<i>p</i> -value	95% CI
2003 - 2005	0.049	0.132	0.375	1.000	(-0.362, 0.460)
2003 - 2006	-0.363	0.143	-2.536	0.211	(-0.808, 0.083)
2003 - 2010	-0.399	0.129	-3.078	0.055	(-0.802, 0.005)
2003 - 2012	-0.077	0.129	-0.599	1.000	(-0.480, 0.325)
2003 - 2014	-0.127	0.129	-0.983	0.986	(-0.528, 0.275)
2003 - 2016	-0.046	0.129	-0.359	1.000	(-0.448, 0.355)
2003 - 2018	0.007	0.129	0.053	1.000	(-0.395, 0.408)
2003 - 2020	-0.229	0.129	-1.780	0.682	(-0.631, 0.172)
2005 - 2006	-0.412	0.105	-3.942	0.003	(-0.738, -0.086)
2005 - 2010	-0.448	0.080	-5.607	0.000	(-0.697, -0.199)
2005 - 2012	-0.127	0.079	-1.598	0.795	(-0.374, 0.120)
2005 - 2014	-0.176	0.079	-2.238	0.371	(-0.421, 0.069)
2005 - 2016	-0.096	0.079	-1.216	0.949	(-0.341, 0.149)
2005 - 2018	-0.043	0.079	-0.541	1.000	(-0.288, 0.203)
2005 - 2020	-0.279	0.079	-3.543	0.014	(-0.524, -0.034)
2006 - 2010	-0.036	0.101	-0.354	1.000	(-0.352, 0.280)
2006 - 2012	0.285	0.101	2.823	0.109	(-0.030, 0.600)
2006 - 2014	0.236	0.101	2.344	0.307	(-0.078, 0.550)
2006 - 2016	0.316	0.101	3.143	0.046	(0.003, 0.630)
2006 - 2018	0.370	0.101	3.671	0.009	(0.056, 0.683)
2006 - 2020	0.133	0.101	1.324	0.918	(-0.180, 0.447)
2010 - 2012	0.321	0.074	4.355	0.001	(0.091, 0.551)
2010 - 2014	0.272	0.073	3.722	0.007	(0.044, 0.499)
2010 - 2016	0.352	0.073	4.823	0.000	(0.125, 0.580)
2010 - 2018	0.405	0.073	5.551	0.000	(0.178, 0.633)
2010 - 2020	0.169	0.073	2.317	0.324	(-0.058, 0.397)
2012 - 2014	-0.049	0.072	-0.681	0.999	(-0.274, 0.176)
2012 - 2016	0.031	0.072	0.431	1.000	(-0.194, 0.256)
2012 - 2018	0.084	0.072	1.166	0.960	(-0.141, 0.309)
2012 - 2020	-0.152	0.072	-2.102	0.461	(-0.377, 0.073)
2014 - 2016	0.080	0.072	1.123	0.968	(-0.143, 0.303)
2014 - 2018	0.134	0.072	1.866	0.624	(-0.089, 0.357)
2014 - 2020	-0.103	0.072	-1.434	0.876	(-0.326, 0.120)
2016 - 2018	0.053	0.072	0.742	0.998	(-0.170, 0.276)
2016 - 2020	-0.183	0.072	-2.557	0.201	(-0.406, 0.040)
2018 - 2020	-0.236	0.072	-3.300	0.029	(-0.459, -0.013)

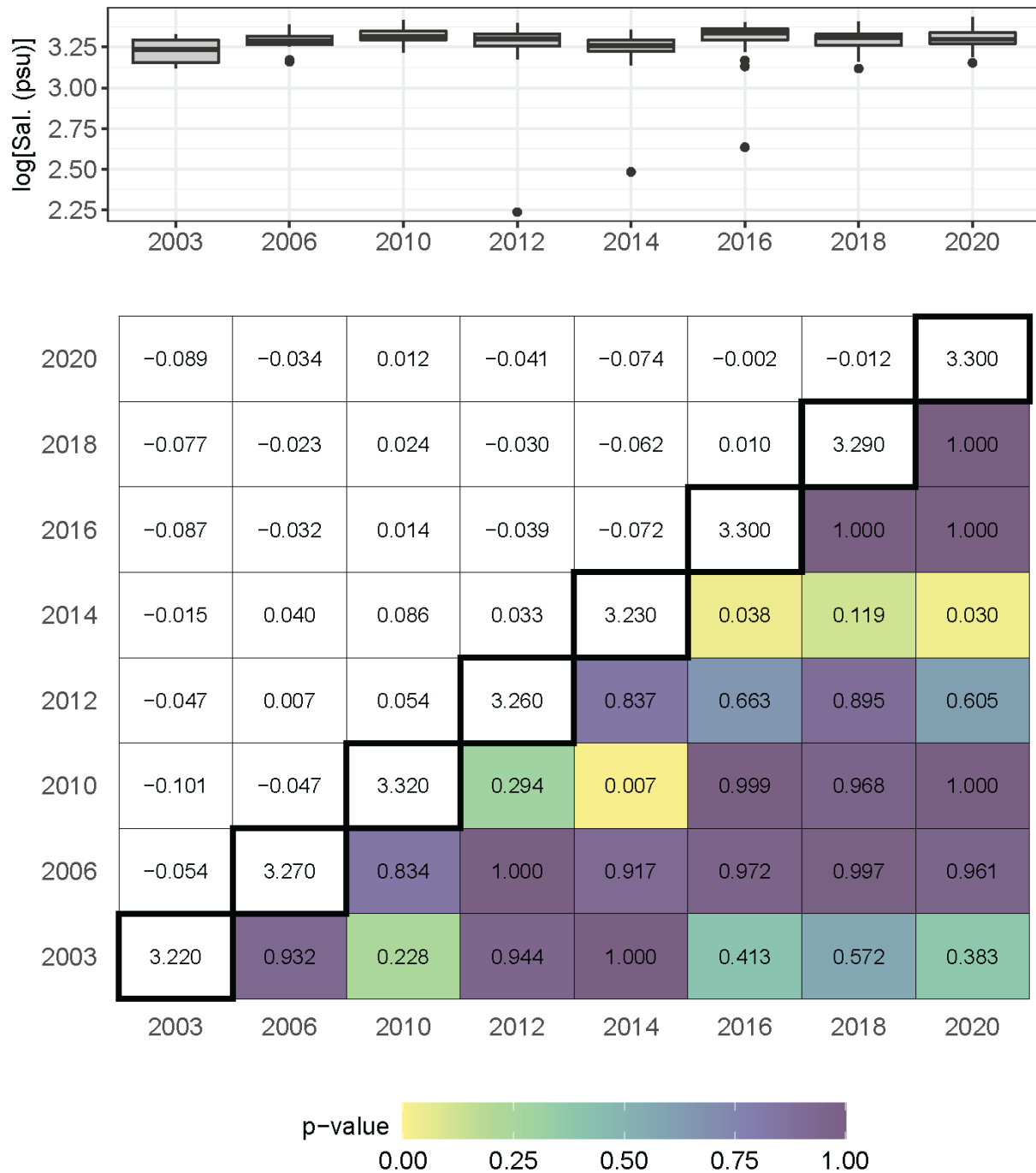


Figure 2-4. (Top): Boxplot of log surface salinity by year within Gateway National Recreation Area. (Bottom): Compact display of the multiple comparison test results. The bolded diagonal elements contain the mean log surface salinity within each year. Upper diagonal cells contain the difference between the means of the years indicated by the row and column names (earlier minus latter). Lower diagonals contain the Tukey-adjusted p-value from a two-sided test that the means of the years indicated by the row and column names are equal (i.e., $H_0: \mu_1 = \mu_2$ versus $H_1: \mu_1 \neq \mu_2$). A more detailed summary of the test statistics and confidence intervals is provided in Table 2-9.

Table 2-9. Summary of the multiple comparison test results for all possible yearly contrasts of mean log surface salinity within Gateway National Recreation Area. Note: all estimates and confidence interval bounds are reported on the log scale. [SE = standard error; CI = confidence interval]

Contrast	Estimate	SE	t_{158}	p -value	95% CI
2003 - 2006	-0.054	0.046	-1.179	0.932	(-0.194, 0.086)
2003 - 2010	-0.101	0.042	-2.417	0.228	(-0.228, 0.026)
2003 - 2012	-0.047	0.042	-1.136	0.944	(-0.174, 0.079)
2003 - 2014	-0.015	0.042	-0.349	1.000	(-0.141, 0.112)
2003 - 2016	-0.087	0.042	-2.086	0.413	(-0.213, 0.040)
2003 - 2018	-0.077	0.042	-1.850	0.572	(-0.203, 0.050)
2003 - 2020	-0.089	0.042	-2.133	0.383	(-0.215, 0.038)
2006 - 2010	-0.047	0.033	-1.427	0.834	(-0.146, 0.053)
2006 - 2012	0.007	0.032	0.218	1.000	(-0.092, 0.106)
2006 - 2014	0.040	0.032	1.227	0.917	(-0.059, 0.139)
2006 - 2016	-0.032	0.032	-0.999	0.972	(-0.131, 0.066)
2006 - 2018	-0.023	0.032	-0.697	0.997	(-0.121, 0.076)
2006 - 2020	-0.034	0.032	-1.060	0.961	(-0.133, 0.064)
2010 - 2012	0.054	0.023	2.286	0.293	(-0.018, 0.125)
2010 - 2014	0.086	0.023	3.680	0.007	(0.015, 0.158)
2010 - 2016	0.014	0.023	0.605	0.999	(-0.057, 0.086)
2010 - 2018	0.024	0.023	1.023	0.968	(-0.047, 0.096)
2010 - 2020	0.012	0.023	0.521	0.999	(-0.059, 0.084)
2012 - 2014	0.033	0.023	1.422	0.836	(-0.037, 0.103)
2012 - 2016	-0.039	0.023	-1.716	0.663	(-0.110, 0.031)
2012 - 2018	-0.030	0.023	-1.289	0.895	(-0.100, 0.040)
2012 - 2020	-0.041	0.023	-1.802	0.605	(-0.112, 0.029)
2014 - 2016	-0.072	0.023	-3.138	0.038	(-0.142, -0.002)
2014 - 2018	-0.062	0.023	-2.711	0.119	(-0.132, 0.008)
2014 - 2020	-0.074	0.023	-3.224	0.030	(-0.144, -0.004)
2016 - 2018	0.010	0.023	0.427	1.000	(-0.060, 0.080)
2016 - 2020	-0.002	0.023	-0.086	1.000	(-0.072, 0.068)
2018 - 2020	-0.012	0.023	-0.513	1.000	(-0.082, 0.058)
2003 - 2006	-0.054	0.046	-1.179	0.932	(-0.194, 0.086)
2003 - 2010	-0.101	0.042	-2.417	0.228	(-0.228, 0.026)
2003 - 2012	-0.047	0.042	-1.136	0.944	(-0.174, 0.079)
2003 - 2014	-0.015	0.042	-0.349	1.000	(-0.141, 0.112)
2003 - 2016	-0.087	0.042	-2.086	0.413	(-0.213, 0.040)
2003 - 2018	-0.077	0.042	-1.850	0.572	(-0.203, 0.050)
2003 - 2020	-0.089	0.042	-2.133	0.383	(-0.215, 0.038)
2006 - 2010	-0.047	0.033	-1.427	0.834	(-0.146, 0.053)

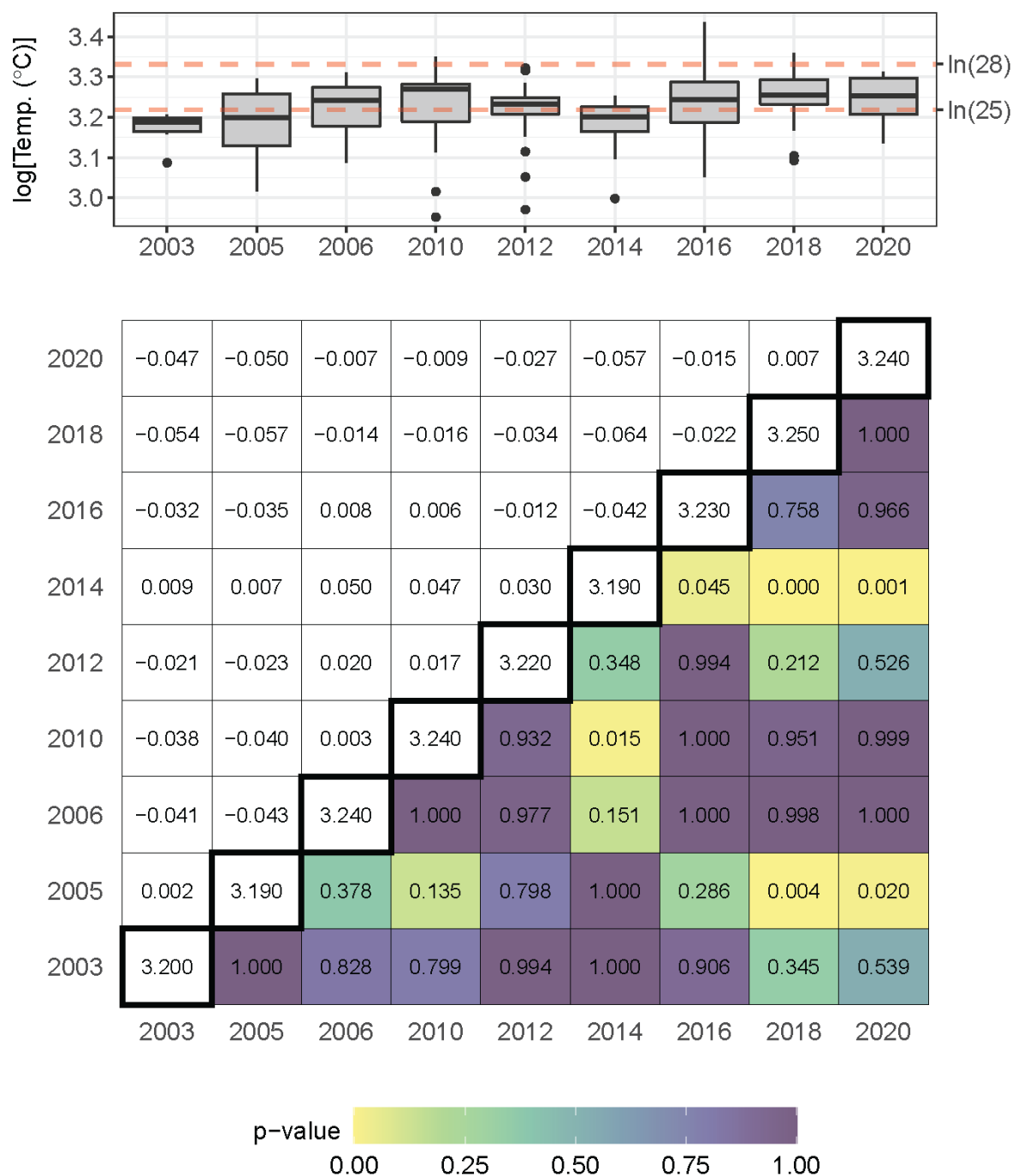
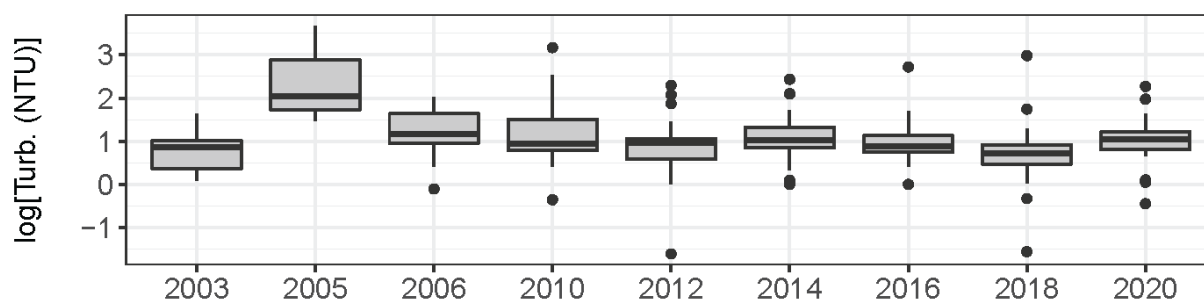


Figure 2-5. (Top): Boxplot of log surface temperature by year within Gateway National Recreation Area. Estuarine quality category thresholds (Table 2-1) are plotted as red dashed lines. (Bottom): Compact display of the multiple comparison test results. The bolded diagonal elements contain the mean log surface temperature within each year. Upper diagonal cells contain the difference between the means of the years indicated by the row and column names (earlier minus latter). Lower diagonals contain the Tukey-adjusted p-value from a two-sided test that the means of the years indicated by the row and column names are equal (i.e., $H_0: \mu_1 = \mu_2$ versus $H_1: \mu_1 \neq \mu_2$). A more detailed summary of the test statistics and confidence intervals is provided in Table 2-10.

Table 2-10. Summary of the multiple comparison test results for all possible yearly contrasts of mean log surface temperature within Gateway National Recreation Area. Note: all estimates and confidence interval bounds are reported on the log scale. [SE = standard error; CI = confidence interval]

Contrast	Estimate	SE	t_{179}	p -value	95% CI
2003 - 2005	0.002	0.024	0.098	1.000	(-0.073, 0.078)
2003 - 2006	-0.041	0.026	-1.536	0.828	(-0.123, 0.042)
2003 - 2010	-0.038	0.024	-1.590	0.799	(-0.112, 0.036)
2003 - 2012	-0.021	0.024	-0.873	0.994	(-0.095, 0.053)
2003 - 2014	0.009	0.024	0.394	1.000	(-0.065, 0.083)
2003 - 2016	-0.032	0.024	-1.357	0.906	(-0.106, 0.042)
2003 - 2018	-0.054	0.024	-2.281	0.345	(-0.128, 0.020)
2003 - 2020	-0.047	0.024	-1.987	0.539	(-0.121, 0.027)
2005 - 2006	-0.043	0.019	-2.226	0.379	(-0.103, 0.017)
2005 - 2010	-0.040	0.015	-2.737	0.134	(-0.086, 0.006)
2005 - 2012	-0.023	0.015	-1.593	0.797	(-0.068, 0.022)
2005 - 2014	0.007	0.015	0.480	1.000	(-0.038, 0.052)
2005 - 2016	-0.035	0.015	-2.386	0.285	(-0.080, 0.011)
2005 - 2018	-0.057	0.015	-3.898	0.004	(-0.102, -0.011)
2005 - 2020	-0.050	0.015	-3.417	0.020	(-0.095, -0.004)
2006 - 2010	0.003	0.019	0.137	1.000	(-0.056, 0.061)
2006 - 2012	0.020	0.019	1.067	0.976	(-0.038, 0.078)
2006 - 2014	0.050	0.019	2.688	0.151	(-0.008, 0.108)
2006 - 2016	0.008	0.019	0.446	1.000	(-0.050, 0.066)
2006 - 2018	-0.014	0.019	-0.736	0.998	(-0.072, 0.044)
2006 - 2020	-0.007	0.019	-0.360	1.000	(-0.065, 0.051)
2010 - 2012	0.017	0.013	1.278	0.932	(-0.025, 0.059)
2010 - 2014	0.047	0.013	3.511	0.015	(0.005, 0.089)
2010 - 2016	0.006	0.013	0.424	1.000	(-0.036, 0.048)
2010 - 2018	-0.016	0.013	-1.204	0.951	(-0.058, 0.026)
2010 - 2020	-0.009	0.013	-0.687	0.999	(-0.051, 0.033)
2012 - 2014	0.030	0.013	2.278	0.347	(-0.011, 0.071)
2012 - 2016	-0.012	0.013	-0.872	0.994	(-0.053, 0.030)
2012 - 2018	-0.034	0.013	-2.533	0.213	(-0.075, 0.008)
2012 - 2020	-0.027	0.013	-2.005	0.527	(-0.068, 0.015)
2014 - 2016	-0.042	0.013	-3.150	0.045	(-0.083, -0.000)
2014 - 2018	-0.064	0.013	-4.811	0.000	(-0.105, -0.022)
2014 - 2020	-0.057	0.013	-4.283	0.001	(-0.098, -0.015)
2016 - 2018	-0.022	0.013	-1.662	0.758	(-0.063, 0.019)
2016 - 2020	-0.015	0.013	-1.134	0.966	(-0.056, 0.026)
2018 - 2020	0.007	0.013	0.528	1.000	(-0.034, 0.048)



2020	-0.068	1.286	0.258	0.153	-0.201	0.097	-0.034	-0.310	1.005
2018	0.242	1.595	0.568	0.463	0.109	0.406	0.275	0.695	0.392
2016	-0.033	1.320	0.292	0.188	-0.167	0.131	0.971	0.557	1.000
2014	-0.164	1.189	0.161	0.056	-0.298	1.102	0.990	0.091	0.999
2012	0.134	1.487	0.459	0.354	0.804	0.461	0.958	0.997	0.883
2010	-0.221	1.133	0.105	1.158	0.251	1.000	0.922	0.036	0.976
2006	-0.326	1.028	1.263	1.000	0.321	0.996	0.851	0.091	0.921
2005	-1.353	2.291	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2003	0.938	0.000	0.960	0.994	1.000	0.999	1.000	0.988	1.000
	2003	2005	2006	2010	2012	2014	2016	2018	2020



Figure 2-6. (Top): Boxplot of log surface turbidity by year within Gateway National Recreation Area. (Bottom): Compact display of the multiple comparison test results. The bolded diagonal elements contain the mean log surface turbidity within each year. Upper diagonal cells contain the difference between the means of the years indicated by the row and column names (earlier minus latter). Lower diagonals contain the Tukey-adjusted p -value from a two-sided test that the means of the years indicated by the row and column names are equal (i.e., $H_0: \mu_1 = \mu_2$ versus $H_1: \mu_1 \neq \mu_2$). A more detailed summary of the test statistics and confidence intervals is provided in Table 2-11.

Table 2-11. Summary of the multiple comparison test results for all possible yearly contrasts of mean log surface turbidity within Gateway National Recreation Area. Note: all estimates and confidence interval bounds are reported on the log scale. [SE = standard error; CI = confidence interval]

Contrast	Estimate	SE	<i>t</i> ₁₇₈	<i>p</i> -value	95% CI
2003 - 2005	-1.353	0.257	-5.259	0.000	(-2.155, -0.552)
2003 - 2006	-0.326	0.280	-1.164	0.960	(-1.197, 0.546)
2003 - 2010	-0.221	0.252	-0.874	0.993	(-1.007, 0.566)
2003 - 2012	0.133	0.252	0.530	1.000	(-0.651, 0.918)
2003 - 2014	-0.164	0.251	-0.654	0.999	(-0.947, 0.618)
2003 - 2016	-0.033	0.251	-0.132	1.000	(-0.816, 0.749)
2003 - 2018	0.242	0.251	0.964	0.988	(-0.540, 1.024)
2003 - 2020	-0.067	0.251	-0.269	1.000	(-0.850, 0.715)
2005 - 2006	1.028	0.204	5.032	0.000	(0.391, 1.664)
2005 - 2010	1.133	0.156	7.247	0.000	(0.646, 1.620)
2005 - 2012	1.487	0.155	9.576	0.000	(1.003, 1.971)
2005 - 2014	1.189	0.154	7.725	0.000	(0.710, 1.669)
2005 - 2016	1.320	0.154	8.576	0.000	(0.841, 1.800)
2005 - 2018	1.596	0.154	10.364	0.000	(1.116, 2.075)
2005 - 2020	1.286	0.154	8.353	0.000	(0.806, 1.766)
2006 - 2010	0.105	0.198	0.530	1.000	(-0.512, 0.722)
2006 - 2012	0.459	0.198	2.323	0.321	(-0.157, 1.075)
2006 - 2014	0.161	0.196	0.822	0.996	(-0.450, 0.773)
2006 - 2016	0.292	0.196	1.489	0.851	(-0.319, 0.904)
2006 - 2018	0.568	0.196	2.891	0.092	(-0.044, 1.180)
2006 - 2020	0.258	0.196	1.315	0.921	(-0.354, 0.870)
2010 - 2012	0.354	0.144	2.452	0.251	(-0.096, 0.804)
2010 - 2014	0.056	0.143	0.395	1.000	(-0.389, 0.502)
2010 - 2016	0.187	0.143	1.310	0.922	(-0.258, 0.633)
2010 - 2018	0.463	0.143	3.234	0.036	(0.017, 0.909)
2010 - 2020	0.153	0.143	1.070	0.976	(-0.293, 0.599)
2012 - 2014	-0.298	0.142	-2.101	0.461	(-0.739, 0.144)
2012 - 2016	-0.167	0.142	-1.177	0.957	(-0.608, 0.275)
2012 - 2018	0.109	0.142	0.767	0.997	(-0.333, 0.550)
2012 - 2020	-0.201	0.142	-1.418	0.883	(-0.642, 0.240)
2014 - 2016	0.131	0.140	0.933	0.990	(-0.306, 0.568)
2014 - 2018	0.406	0.140	2.896	0.091	(-0.031, 0.844)
2014 - 2020	0.097	0.140	0.689	0.999	(-0.340, 0.534)
2016 - 2018	0.275	0.140	1.962	0.557	(-0.162, 0.713)
2016 - 2020	-0.034	0.140	-0.244	1.000	(-0.471, 0.403)
2018 - 2020	-0.310	0.140	-2.206	0.392	(-0.747, 0.128)

2.4 GAM Results

The generalized additive model summaries indicate that there exists statistically significant (assuming a 0.05 significance level) spatial and temporal trends in all water quality measures (Table 2-12). Notable observations from the fitted smooths for each measure are as follows:

2.4.1 Chlorophyll-*a* (near surface)

The temporal trend in chlorophyll-*a* is non-linear and exhibited a statistically significant decrease that occurred during the period of 2012 to 2014 (Figure 2-7). Concentrations of chlorophyll-*a* are highest along the northern shore near Mill Basin Outlet, Canarsie Pier, and North Channel (Figure 2-8). Concentrations are estimated to be lowest in Rockaway Inlet. Based on the fitted model, both time and location account for 46.7% of the variability in chlorophyll-*a* concentrations (Table 2-13).

2.4.2 Dissolved Oxygen (near bottom)

Bottom dissolved oxygen has exhibited a statistically significant linear decrease since 2003 (Figure 2-7). Concentrations are estimated to be lowest near John F. Kennedy International Airport (Figure 2-9). Based on the fitted model, both time and location account for 21.6% of the variability in bottom dissolved oxygen concentrations (Table 2-13).

2.4.3 PAR Attenuation K_d (SAV standard)

The temporal trend in PAR attenuation exhibited a significant increase around 2006 followed by a significant decrease between 2010-2012 (Figure 2-7). Spatially, PAR attenuation is estimated to be highest along the northern shore near Mill Basin Outlet, Canarsie Pier, North Channel, and John F. Kennedy International Airport (Figure 2-10). PAR attenuation is estimated to be lowest in Rockaway Inlet. Based on the fitted model, both time and location account for 63.3% of the variability in PAR attenuation (Table 2-13).

2.4.4 Salinity (near surface)

Surface salinity has been nearly constant over time with the exception of 2005, which had extremely low salinity (Figure 2-7). Spatially, salinity is estimated to be lowest along the northern shore near Mill Basin Outlet, Canarsie Pier, North Channel, and John F. Kennedy International Airport (Figure 2-11). Salinity is estimated to be highest in Rockaway Inlet. Based on the fitted model, both time and location account for 37.9% of the variation in salinity (Table 2-13).

2.4.5 Temperature (near surface)

The temporal trend in surface water temperature is non-linear and exhibited a significant increase around 2016 (Figure 2-7 (continued)). Surface temperature is estimated to be highest along the northern shore near Mill Basin Outlet, Canarsie Pier, North Channel, and John F. Kennedy International Airport and lowest in Rockaway Inlet (Figure 2-12). Based on the fitted model, both time and location account for 59.4% of the variability in surface water temperature (Table 2-13).

2.4.6 Turbidity (near surface)

The temporal trend in surface turbidity has been relatively constant except for significant increase that occurred during 2005 (Figure 2-7 (continued)). Spatially, turbidity is estimated to be highest along the north/northwestern shore around Mill Basin Outlet and Canarsie Pier (Figure 2-13). Time and location account for 57.8% of the variation in turbidity (Table 2-13).

Table 2-12. Approximate significance of the generalized additive model smooth terms for each model. Each respective F -statistic and p -value corresponds to the test of the hypotheses: $H_0: f(x) = 0$ versus $H_1: f(x) \neq 0$. Note: p -values are computed without considering uncertainty in the smoothing parameter estimates and may need to be interpreted with caution when near the α level. [edf = effective degrees of freedom]

Measure	Depth	Model Term	edf	F-value	p-value
Chlorophyll <i>a</i> ($\mu\text{g/L}$)	Surface	s(year)	4.12	4.47	< 0.001
Chlorophyll <i>a</i> ($\mu\text{g/L}$)	Surface	s(long, lat)	9.08	3.72	< 0.001
Dissolved oxygen (mg/L)	Bottom	s(year)	0.79	0.57	0.016
Dissolved oxygen (mg/L)	Bottom	s(long, lat)	9.50	1.28	< 0.001
PAR attenuation (K_d (m^{-1}))	Water Column	s(year)	6.16	4.83	< 0.001
PAR attenuation (K_d (m^{-1}))	Water Column	s(long, lat)	12.54	4.56	< 0.001
Salinity (psu)	Surface	s(year)	3.70	1.88	0.005
Salinity (psu)	Surface	s(long, lat)	9.01	2.60	< 0.001
Temperature ($^{\circ}\text{C}$)	Surface	s(year)	4.73	4.80	< 0.001
Temperature ($^{\circ}\text{C}$)	Surface	s(long, lat)	13.57	4.85	< 0.001
Turbidity (NTU)	Surface	s(year)	7.73	14.62	< 0.001
Turbidity (NTU)	Surface	s(long, lat)	7.04	1.34	< 0.001

Table 2-13. Proportion of null deviance explained by the explanatory variables for each model.

Measure	Deviance Explained
Chlorophyll <i>a</i> ($\mu\text{g/L}$)	46.7%
Dissolved oxygen (mg/L)	21.6%
PAR attenuation (K_d (m^{-1}))	63.3%
Salinity (psu)	37.9%
Temperature ($^{\circ}\text{C}$)	59.4%
Turbidity (NTU)	57.8%

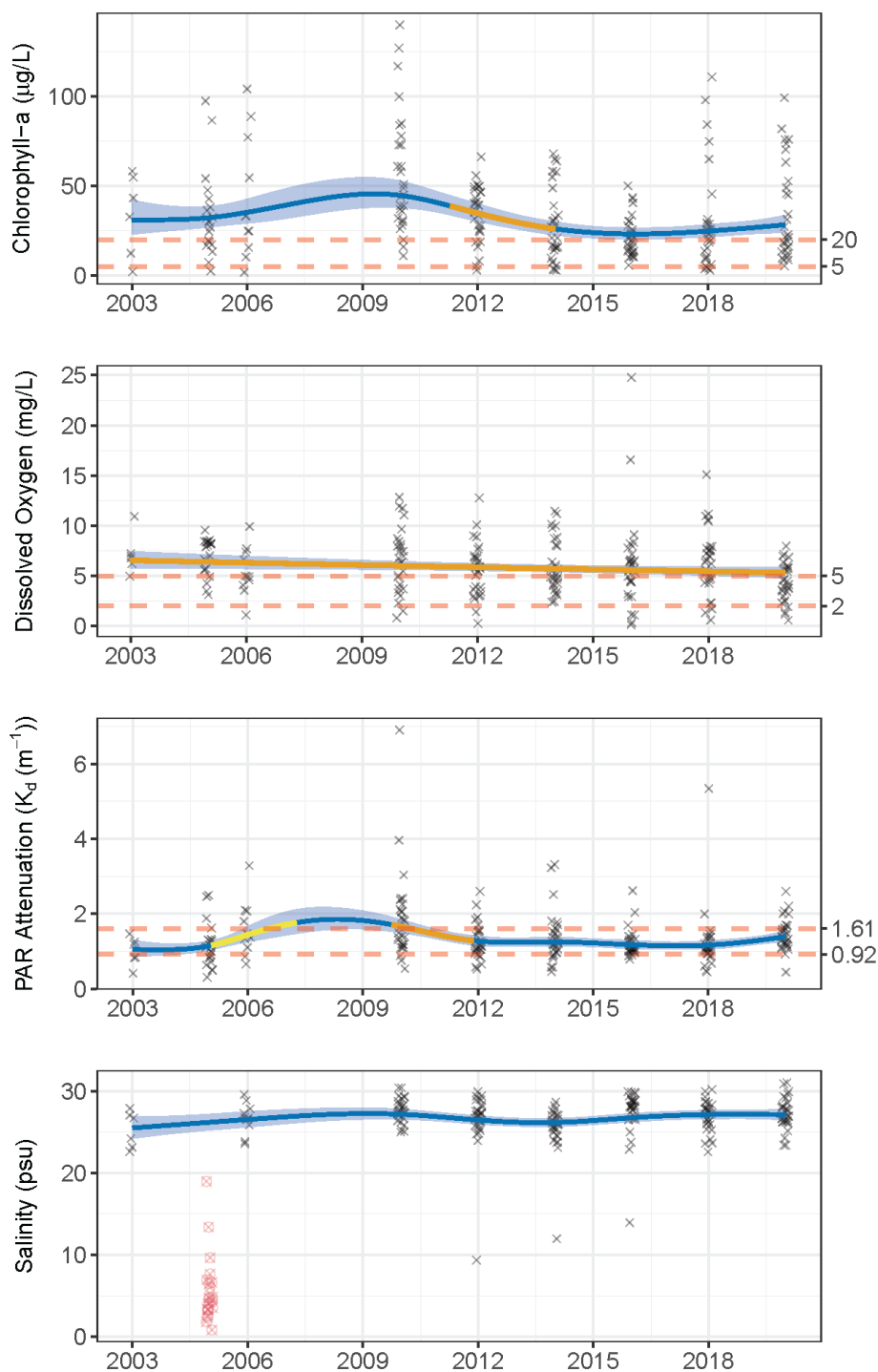


Figure 2-7. Trends in Gateway National Recreation Area water quality measures based on estimated generalized additive model smooths. The fitted smooth is plotted in dark blue with shaded 95% confidence interval. Periods of significant temporal change (increases/decreases), if present, are plotted in yellow/orange. Estuarine quality category thresholds (Table 2-1) are plotted as red dashed lines. Note: years that were excluded from model fitting are plotted in red.

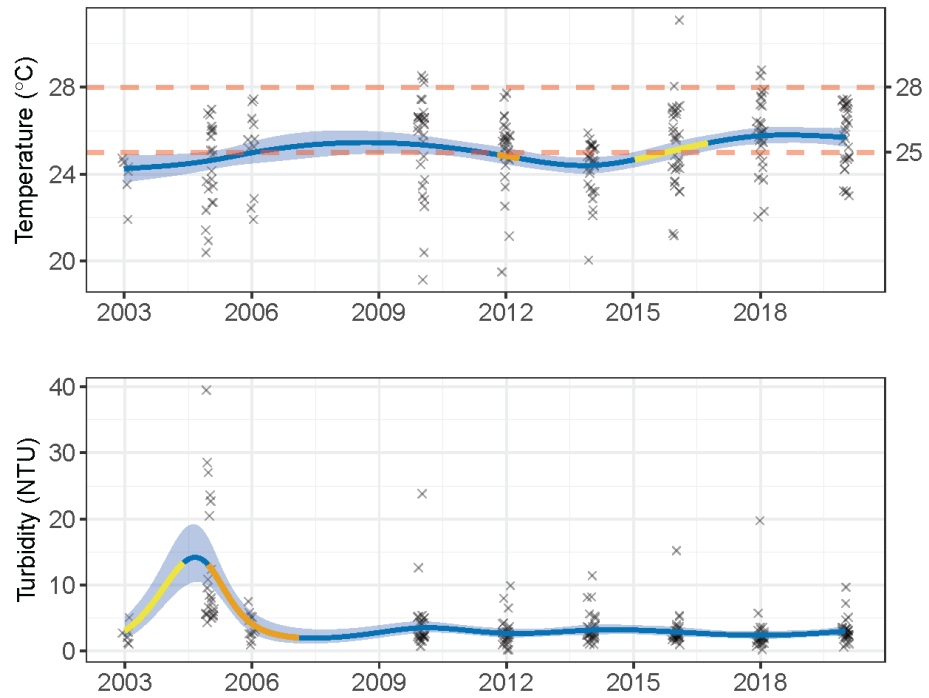


Figure 2-7 (continued). Trends in Gateway National Recreation Area water quality measures based on estimated generalized additive model smooths. The fitted smooth is plotted in dark blue with shaded 95% confidence interval. Periods of significant temporal change (increases/decreases), if present, are plotted in yellow/orange. Estuarine quality category thresholds (Table 2-1) are plotted as red dashed lines.

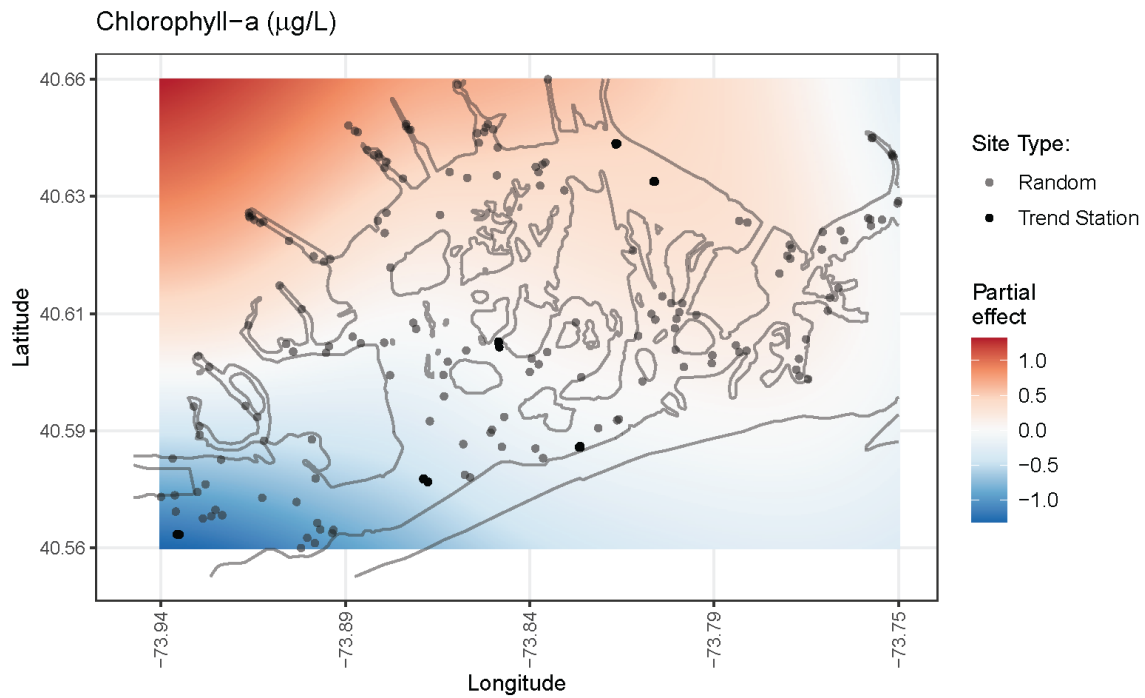


Figure 2-8. Mean surface for the generalized additive model latitude/longitude smooth of surface chlorophyll-a concentration within Gateway National Recreation Area during 2003-2018.

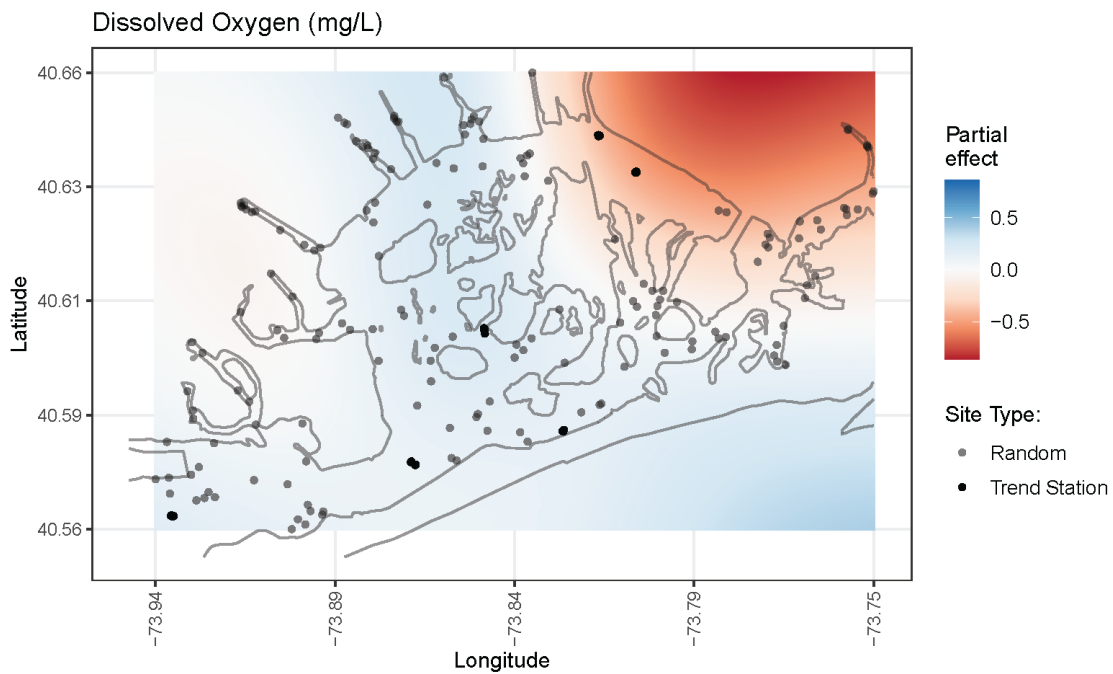


Figure 2-9. Mean surface for the generalized additive model latitude/longitude smooth of bottom dissolved oxygen concentration within Gateway National Recreation Area during 2003-2018.

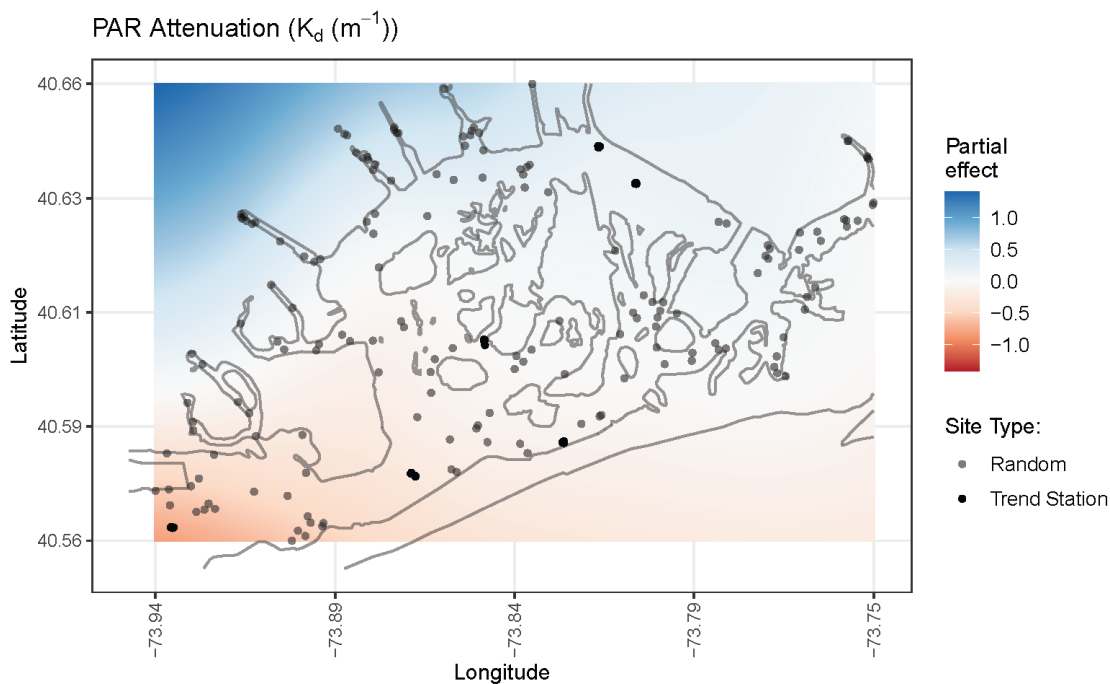


Figure 2-10. Mean surface for the generalized additive model latitude/longitude smooth of PAR attenuation within Gateway National Recreation Area during 2003-2018.

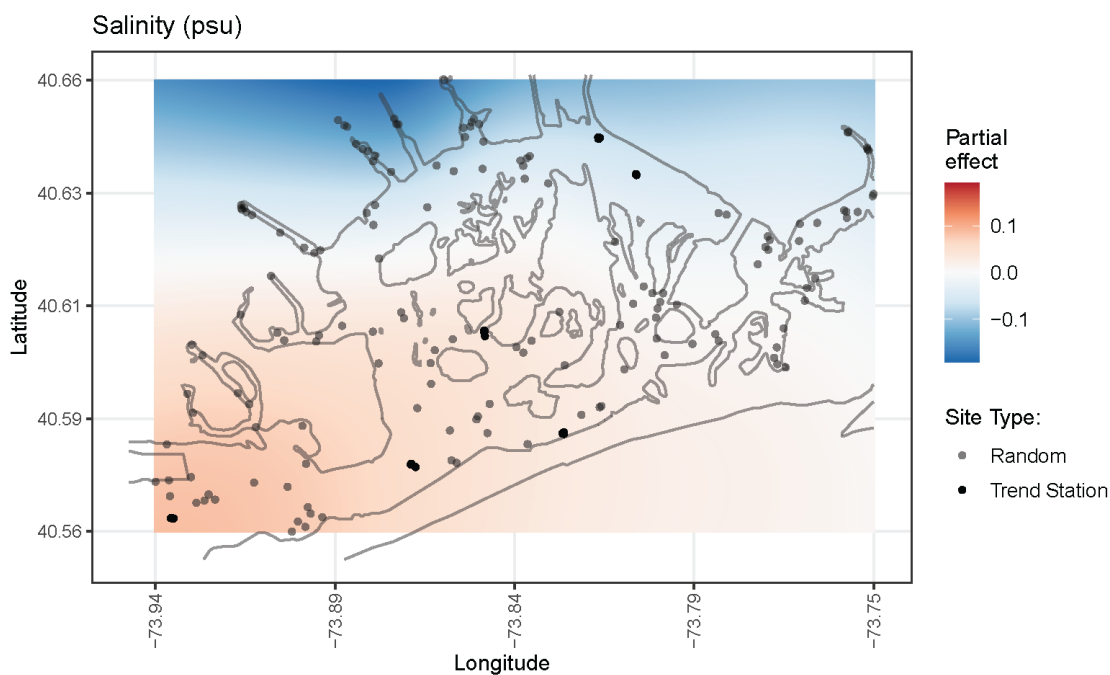


Figure 2-11. Mean surface for the generalized additive model latitude/longitude smooth of surface salinity within Gateway National Recreation Area during 2003-2018.

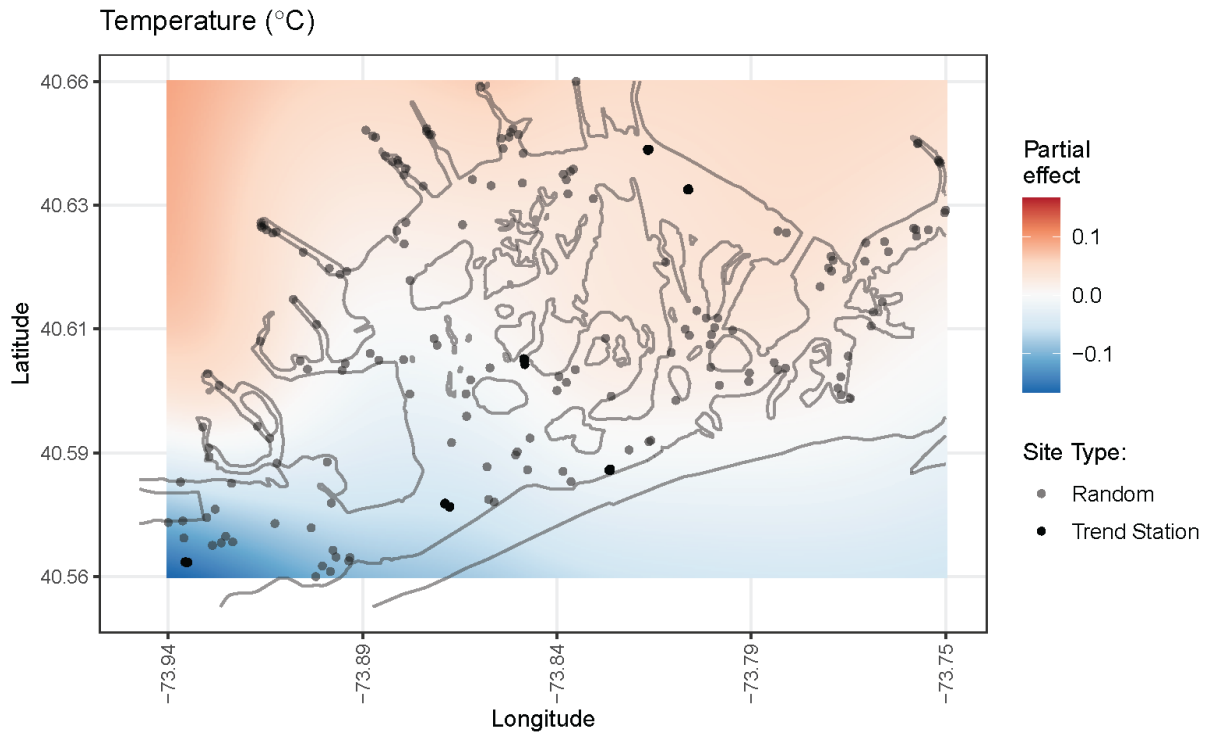


Figure 2-12. Mean surface for the generalized additive model latitude/longitude smooth of surface temperature within Gateway National Recreation Area during 2003-2018.

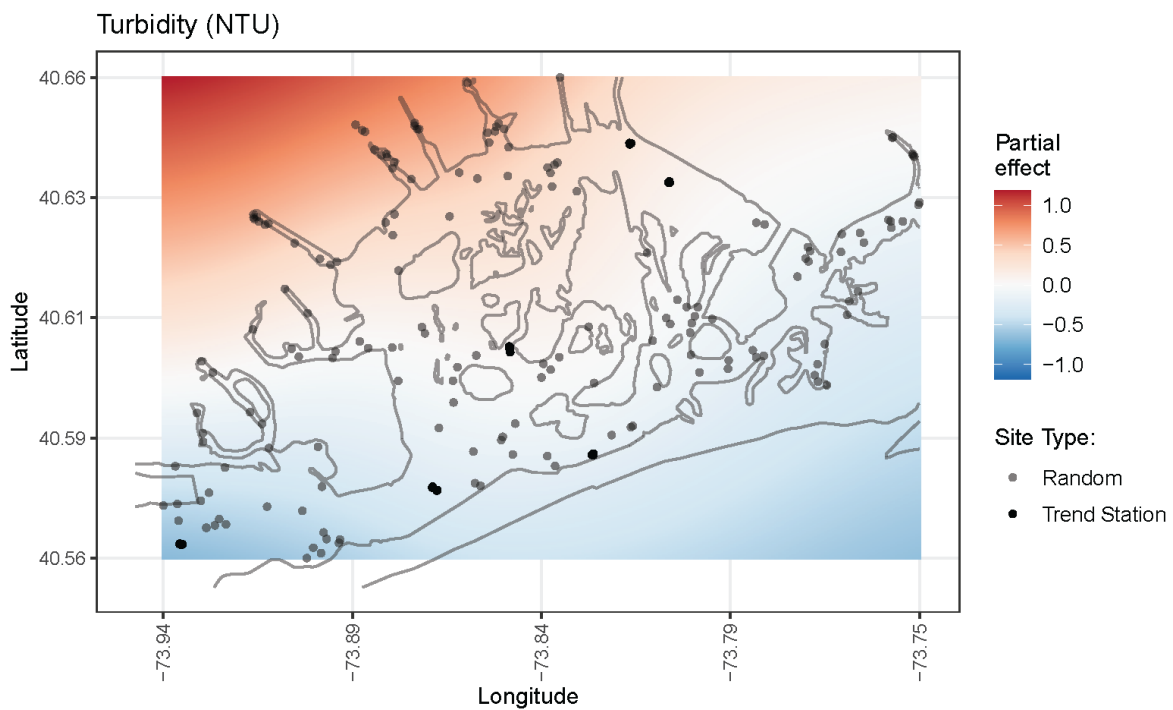


Figure 2-13. Mean surface for the generalized additive model latitude/longitude smooth of surface turbidity within Gateway National Recreation Area during 2003-2018.

2.5 Multivariate Mann-Kendall Trend Test Results

The multivariate Mann-Kendall test results indicate there is statistical evidence of decreasing monotonic trends in bottom dissolved oxygen (p -value = 0.025) and surface turbidity (p -value = 0.022) during the index periods of 2003 to 2020 (Table 2-14).

Table 2-14. Summary of multivariate Mann-Kendall test for monotonic trend for each measure: H_0 : no monotonic trend versus H_1 : monotonic trend is present. A positive (negative) value of the Mann-Kendall test statistic (Z_{MK}) indicates that measurements tend to increase (decrease) with time.

Measure	Depth	Z_{MK}	p -value
Chlorophyll a ($\mu\text{g/L}$)	Surface	-1.55	0.122
Dissolved oxygen (mg/L)	Bottom	-2.24	0.025
PAR attenuation (K_d (m^{-1}))	Water Column	0.07	0.947
Salinity (psu)	Surface	0.17	0.865
Temperature ($^{\circ}\text{C}$)	Surface	1.23	0.220
Turbidity (NTU)	Surface	-2.29	0.022

3. Trend-Station Water Quality Data

3.1 Methods

3.1.1 Generalized Additive Model (GAM)

Generalized additive models (GAMs) are used to estimate trends in water quality measures for each trend station within an estuary. For each measure, we model the mean response as a function of time. Mathematically, the model can be written as:

$$\begin{aligned}\text{response}_i &\sim \text{Gamma}(\mu_i, \phi) \\ \log(\mu_i) &= \beta_0 + f_{\text{station}_i}(\text{year}_i),\end{aligned}$$

where $f_{\text{station}_i}(\text{year})$ is the TRPS smooth of *year* for the *i*th trend station. The year smoother is known as a factor-smoother and assumes that all factors (i.e., trend stations) have the same smoothness, but may differ in shape over time. The rank of the basis expansion was set equal to the number of years of data. The relatively short time series within each trend station precluded any modeling of temporal autocorrelation.

The first derivative is approximated along the length of the *year* smooth using the finite difference method. Periods of significant temporal change are identified as intervals of time wherein the 95% simultaneous confidence interval of the first derivative (i.e., slope of tangent line) does not contain the value of zero.

Analyses were carried out using the R Statistical Computing Environment (R Core Team 2020). Each model is fit using the *mgcv* package (Wood 2011) and results were summarized/visualized using the *gratia* package (Simpson 2020).

3.1.2 Multivariate Mann-Kendall Trend Analysis

The multivariate Mann-Kendall test (*trend* R package; (Pohlert 2020)) is used to determine whether there exists a monotonically increasing or decreasing trend in values for all trend station sites within the estuary over time (Libiseller and Grimvall 2002). Positive (negative) values of the test statistic Z_{MK} indicates that measurements tend to increase (decrease) with time.

3.2 Summary Statistics

Yearly summary statistics for each trend station were computed for surface chlorophyll-*a* concentration (Table 3-1), bottom dissolved oxygen (Table 3-2), PAR attenuation (Table 3-3), surface salinity (Table 3-4), surface temperature (Table 3-5), and surface turbidity (Table 3-6).

Table 3-1. Yearly summary statistics for surface chlorophyll-a concentration ($\mu\text{g/L}$) at Gateway National Recreation Area trend stations. [SD = standard deviation]

Station	Statistic	Sample Year								
		2003	2005	2006	2010	2012	2014	2016	2018	2020
1	Mean	3.96	2.30	5.23	19.00	2.82	3.99	13.59	3.99	11.44
1	SD	3.03	0.72	4.82	10.23	1.16	2.01	4.74	0.51	5.07
1	n	6	4	2	4	4	4	4	4	4
5	Mean	14.06	12.14	11.05	41.75	15.99	11.30	19.00	9.09	19.77
5	SD	5.51	16.22	1.10	23.70	9.39	8.45	7.18	6.48	10.24
5	n	7	4	2	4	4	4	4	4	4
7	Mean	28.83	28.29	27.32	31.75	29.14	14.59	20.07	19.17	21.19
7	SD	14.29	19.86	8.45	8.06	7.56	11.50	9.73	3.43	5.63
7	n	7	4	2	4	4	4	4	4	4
16	Mean	44.49	40.16	31.77	50.50	40.52	30.88	18.52	22.92	28.36
16	SD	21.59	18.55	32.39	17.79	6.46	16.66	5.89	9.45	6.34
16	n	6	4	2	4	4	4	4	4	4
24	Mean	49.67	32.58	69.23	66.50	40.01	30.41	26.25	35.99	46.85
24	SD	22.90	16.94	49.51	27.62	8.71	20.64	10.90	24.41	21.66
24	n	7	4	2	4	4	4	4	4	4
29	Mean	63.87	44.48	54.69	64.33	36.67	37.84	27.00	40.98	39.39
29	SD	46.53	24.06	31.81	31.50	12.46	29.29	4.86	33.81	26.91
29	n	7	4	2	3	4	4	4	4	4

Table 3-2. Yearly summary statistics for bottom dissolved oxygen concentration (mg/L) at Gateway National Recreation Area trend stations. [SD = standard deviation]

Station	Statistic	Sample Year								
		2003	2005	2006	2010	2012	2014	2016	2018	2020
1	Mean	9.14	6.11	5.83	7.21	5.52	5.87	6.69	6.25	6.76
1	SD	2.25	0.53	1.24	1.31	0.60	0.62	0.59	0.71	0.78
1	n	6	5	4	4	4	4	4	4	4
5	Mean	9.78	5.81	6.03	6.57	5.84	5.68	6.41	6.36	6.44
5	SD	3.39	1.04	0.85	0.84	0.33	0.73	0.39	0.24	0.66
5	n	7	5	5	4	4	4	4	4	4
7	Mean	10.47	6.02	5.85	5.66	6.04	5.31	5.59	6.40	5.14
7	SD	4.52	1.53	1.84	2.06	1.07	0.66	0.65	0.84	0.46
7	n	7	5	4	4	4	4	4	4	4
16	Mean	13.10	8.32	7.32	8.31	7.92	7.72	6.76	8.20	5.16
16	SD	4.32	0.36	2.53	3.10	0.87	0.54	0.77	0.97	0.99
16	n	7	5	4	4	4	4	4	4	4
24	Mean	8.02	3.76	1.74	1.86	2.19	3.65	3.15	3.42	2.19
24	SD	3.13	1.09	0.64	1.39	1.58	0.78	0.29	1.67	0.14
24	n	7	5	4	4	4	4	4	4	4
29	Mean	8.31	4.66	3.18	2.15	4.48	3.78	3.96	3.44	2.99
29	SD	3.80	0.78	1.43	1.38	1.80	1.24	1.03	0.85	0.86
29	n	7	5	4	3	4	4	4	4	4

Table 3-3. Yearly summary statistics for PAR attenuation (K_d (m^{-1})) at Gateway National Recreation Area trend stations. [SD = standard deviation]

Station	Statistic	Sample Year								
		2003	2005	2006	2010	2012	2014	2016	2018	2020
1	Mean	0.46	0.32	0.69	0.81	0.50	0.58	0.81	0.56	0.85
1	SD	0.11	0.16	0.26	0.23	0.04	0.12	0.11	0.17	0.33
1	n	6	4	4	3	4	4	4	4	4
5	Mean	0.74	0.70	1.06	1.09	0.83	0.86	0.94	0.95	1.12
5	SD	0.20	0.11	0.53	0.26	0.14	0.25	0.16	0.20	0.35
5	n	7	5	5	4	4	4	4	4	4
7	Mean	0.97	0.85	1.19	1.15	1.08	1.02	1.01	1.02	1.36
7	SD	0.15	0.28	0.54	0.11	0.14	0.22	0.05	0.08	0.27
7	n	7	5	4	4	4	4	4	4	4
16	Mean	1.23	1.11	1.68	2.13	2.08	1.65	1.12	1.23	1.43
16	SD	0.46	0.35	0.72	0.47	1.30	0.53	0.13	0.35	0.15
16	n	6	4	4	4	4	4	4	4	4
24	Mean	1.12	1.07	3.76	1.80	1.28	1.28	1.25	1.21	1.63
24	SD	0.19	0.33	3.70	0.58	0.20	0.21	0.21	0.10	0.33
24	n	7	5	4	4	4	4	4	4	4
29	Mean	1.25	1.27	1.36	2.00	1.32	1.56	1.20	1.13	1.68
29	SD	0.29	0.23	0.64	0.57	0.23	0.66	0.19	0.24	0.54
29	n	7	5	5	2	4	4	4	4	4

Table 3-4. Yearly summary statistics for surface salinity (psu) at Gateway National Recreation Area trend stations. [SD = standard deviation]

Station	Statistic	Sample Year								
		2003	2005	2006	2010	2012	2014	2016	2018	2020
1	Mean	26.09	7.89	28.59	29.57	29.91	28.88	30.11	29.82	30.44
1	SD	3.54	2.69	1.30	1.56	0.41	1.57	0.53	0.52	0.50
1	n	6	5	4	4	4	4	4	4	4
5	Mean	26.52	8.27	28.34	28.87	28.41	27.88	29.30	28.93	29.24
5	SD	1.71	4.78	0.97	1.32	0.95	1.30	0.75	0.57	0.74
5	n	7	5	5	4	4	4	4	4	4
7	Mean	24.96	8.78	27.13	28.29	27.48	27.30	28.74	27.75	28.16
7	SD	1.99	4.96	0.41	1.09	1.23	1.31	0.47	0.75	0.43
7	n	7	5	4	4	4	4	4	4	4
16	Mean	24.68	7.69	27.16	28.45	27.41	26.90	28.89	28.14	27.95
16	SD	2.03	4.15	0.43	0.91	0.95	0.12	0.20	0.82	0.16
16	n	7	5	4	4	4	4	4	4	4
24	Mean	23.62	10.17	24.58	26.52	25.37	25.11	26.96	26.09	24.97
24	SD	1.45	5.51	1.51	0.54	1.06	0.51	0.50	1.03	1.66
24	n	7	5	4	4	4	4	4	4	4
29	Mean	22.60	8.00	24.46	26.06	25.30	24.58	26.36	24.91	25.65
29	SD	1.34	4.82	0.97	1.04	0.56	0.97	0.91	1.81	1.01
29	n	7	5	4	3	4	4	4	4	4

Table 3-5. Yearly summary statistics for surface temperature (°C) at Gateway National Recreation Area trend stations. [SD = standard deviation]

Station	Statistic	Sample Year								
		2003	2005	2006	2010	2012	2014	2016	2018	2020
1	Mean	18.42	21.72	22.61	22.09	21.91	20.18	22.48	22.54	23.14
1	SD	3.46	0.91	1.50	2.25	1.74	2.10	1.80	1.13	0.30
1	n	6	5	4	4	4	4	4	4	4
5	Mean	19.68	22.35	23.96	23.94	23.83	22.00	24.08	23.94	24.33
5	SD	3.36	1.67	1.90	1.52	1.08	2.04	1.83	0.43	0.46
5	n	7	5	5	4	4	4	4	4	4
7	Mean	20.87	23.78	24.75	24.80	24.62	22.76	24.68	25.28	25.31
7	SD	3.67	2.15	1.89	1.41	0.79	1.37	0.99	0.56	0.55
7	n	7	5	4	4	4	4	4	4	4
16	Mean	21.58	23.57	25.81	26.12	25.25	23.97	25.31	25.25	25.45
16	SD	3.50	1.95	3.13	1.22	1.02	0.56	1.48	0.66	0.90
16	n	7	5	4	4	4	4	4	4	4
24	Mean	21.65	24.38	26.36	26.02	25.84	24.96	26.53	26.16	26.39
24	SD	3.53	2.11	2.81	0.81	0.85	0.17	1.51	0.73	1.07
24	n	7	5	4	4	4	4	4	4	4
29	Mean	21.65	24.42	26.09	26.04	26.02	25.19	26.48	26.53	26.33
29	SD	3.54	1.45	2.75	1.06	1.00	0.66	1.29	1.05	1.20
29	n	7	5	4	3	4	4	4	4	4

Table 3-6. Yearly summary statistics for surface turbidity (NTU) at Gateway National Recreation Area trend stations. [SD = standard deviation]

Station	Statistic	Sample Year								
		2003	2005	2006	2010	2012	2014	2016	2018	2020
1	Mean	1.15	8.88	1.27	1.77	0.80	1.23	2.30	1.03	1.47
1	SD	0.99	4.21	0.56	0.81	0.50	0.26	0.51	0.85	0.69
1	n	5	5	4	4	4	4	4	4	4
5	Mean	2.25	20.02	1.98	2.02	1.77	2.10	2.12	1.76	2.63
5	SD	1.55	15.99	0.71	0.59	0.88	0.52	0.46	0.40	1.64
5	n	6	5	5	4	4	4	4	4	4
7	Mean	2.56	61.77	2.08	2.98	2.35	2.62	2.55	1.84	3.01
7	SD	1.76	119.52	1.31	1.35	0.70	0.57	0.85	0.19	0.36
7	n	6	5	4	4	4	4	4	4	4
16	Mean	4.38	12.13	4.10	4.80	3.67	3.28	2.88	2.46	3.46
16	SD	1.66	7.70	1.35	0.82	0.53	0.88	0.64	0.84	0.37
16	n	6	5	4	4	4	4	4	4	4
24	Mean	3.37	8.29	3.15	3.25	2.62	3.25	2.65	1.70	2.51
24	SD	1.35	4.02	1.65	1.31	0.61	1.34	0.87	0.49	1.03
24	n	6	5	4	4	4	4	4	4	4
29	Mean	6.05	11.46	3.10	3.60	3.78	4.78	2.92	2.34	3.46
29	SD	4.23	7.92	1.12	1.87	1.13	3.25	0.36	1.09	1.98
29	n	6	5	4	3	4	4	4	4	4

3.3 GAM Results

The generalized additive model summaries indicate that there exists statistically significant temporal trends (assuming a significance level of 0.05) at the trend stations in all water quality measures (Table 3-7). Notable observations from the fitted smooths for each measure are as follows:

3.3.1 Chlorophyll-*a* (near surface)

Temporal trends in chlorophyll-*a* are estimated to be decreasing at trend stations 7, 16, 24, 29; approximately constant at trend station 5; and increasing at trend station 1 (Figure 3-1). No periods of significant temporal change were identified likely due to high within-year variability.

3.3.2 Dissolved Oxygen (near bottom)

Decreases in bottom dissolved oxygen were observed at each trend station from 2003 to 2006, with trend stations 24 and 29 exhibiting statistically significant decreases (Figure 3-2). The yearly average concentrations have been approximately constant since 2006.

3.3.3 PAR Attenuation K_d (SAV standard)

Temporal trends are relatively linear in nature except for trend stations 16 and 24 (Figure 3-3). No periods of significant temporal change were identified and years 2005 and 2006 tend to exhibit the highest within-year variability for each trend station.

3.3.4 Salinity (near surface)

Temporal trends are consistent across all trend stations with salinity being relatively constant since 2010 (Figure 3-4). A significant increase in salinity was identified around 2006 and particularly low salinity was observed during 2005.

3.3.5 Temperature (near surface)

Surface temperature has been approximately constant for all stations since 2006 (Figure 3-5). No periods of significant temporal change were identified.

3.3.6 Turbidity (near surface)

A significant increase in surface turbidity was observed in 2005 for trend stations 1, 5, and 7 (Figure 3-6). Similar to the other measures, turbidity has remained approximately constant since 2006.

Table 3-7. Approximate significance of the generalized additive model smooth terms for each trend station model. Each respective F -statistic and p -value corresponds to the test of the hypotheses: $H_0: f(x) = 0$ versus $H_1: f(x) \neq 0$. Note: p -values are computed without considering uncertainty in the smoothing parameter estimates and may need to be interpreted with caution when near the α level. [edf = effective degrees of freedom; df_{ref} = reference degrees of freedom]

Measure	Depth	Model Term	edf(df_{ref})	F -value	p -value
Chlorophyll a ($\mu\text{g/L}$)	Surface	s(year, hex)	12.97 (53)	4.48	< 0.001
Dissolved oxygen (mg/L)	Bottom	s(year, hex)	28.81 (53)	9.28	< 0.001
PAR attenuation (K_d (m^{-1}))	Water Column	s(year, hex)	23.82 (53)	3.80	< 0.001
Salinity (psu)	Surface	s(year, hex)	18.92 (47)	8.59	< 0.001
Temperature ($^{\circ}\text{C}$)	Surface	s(year, hex)	23.21 (53)	3.68	< 0.001
Turbidity (NTU)	Surface	s(year, hex)	47.96 (53)	5.13	< 0.001

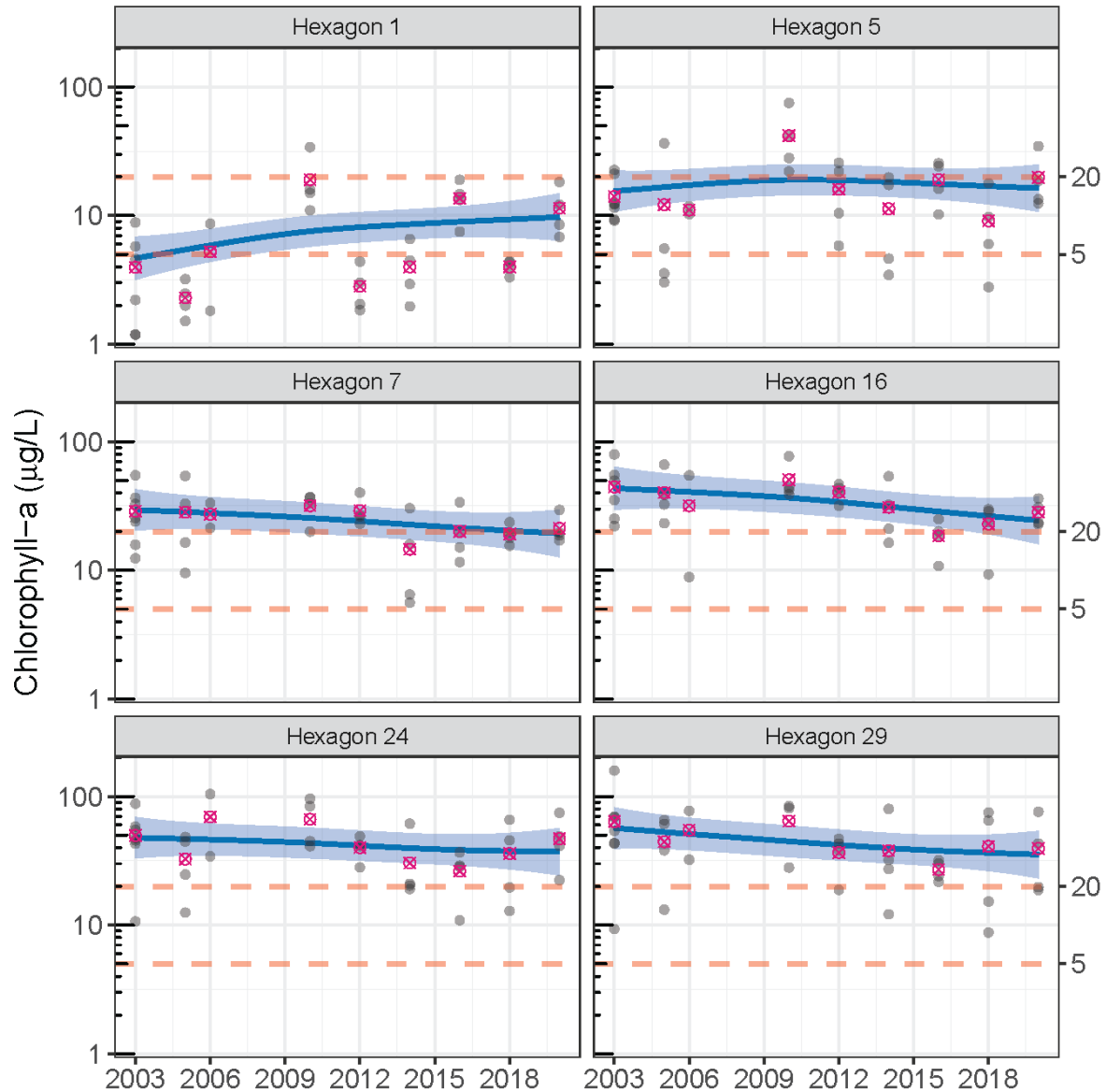


Figure 3-1. Mean surface chlorophyll-a concentration ($\mu\text{g/L}$) trend at each Gateway National Recreation Area trend station based on estimated generalized additive model smooths. The fitted smooth is plotted in dark blue with shaded 95% confidence interval. Periods of significant temporal change (increases/decreases), if present, are plotted in yellow/orange. Within each year, individual measurements and the average are plotted in gray and magenta, respectively. Estuarine quality category thresholds (Table 2-1) are plotted as red dashed lines.

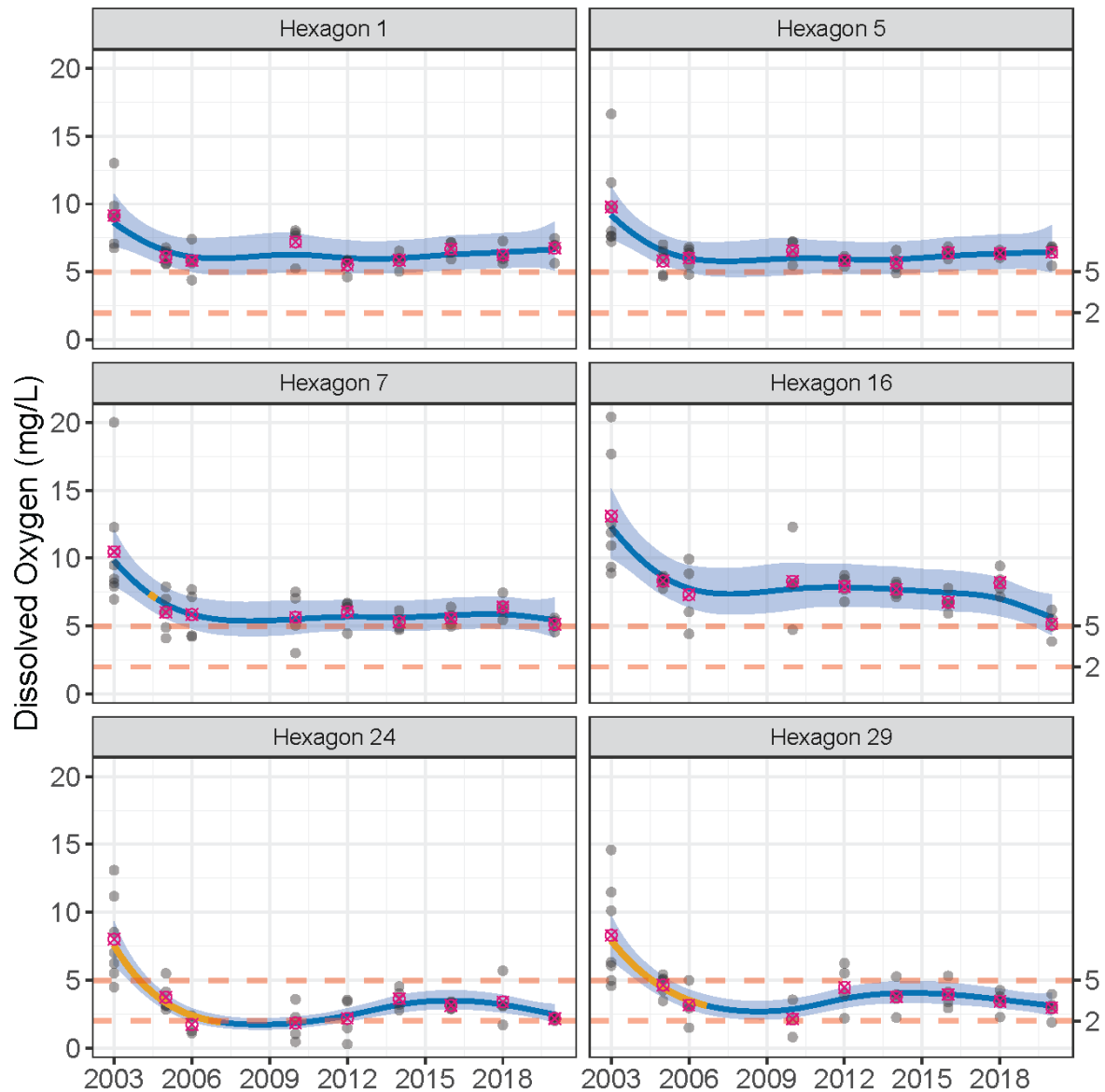


Figure 3-2. Mean bottom dissolved oxygen concentration (mg/L) trend at each Gateway National Recreation Area trend station based on estimated generalized additive model smooths. The fitted smooth is plotted in dark blue with shaded 95% confidence interval. Periods of significant temporal change (increases/decreases), if present, are plotted in yellow/orange. Within each year, individual measurements and the average are plotted in gray and magenta, respectively. Estuarine quality category thresholds (Table 2-1) are plotted as red dashed lines.

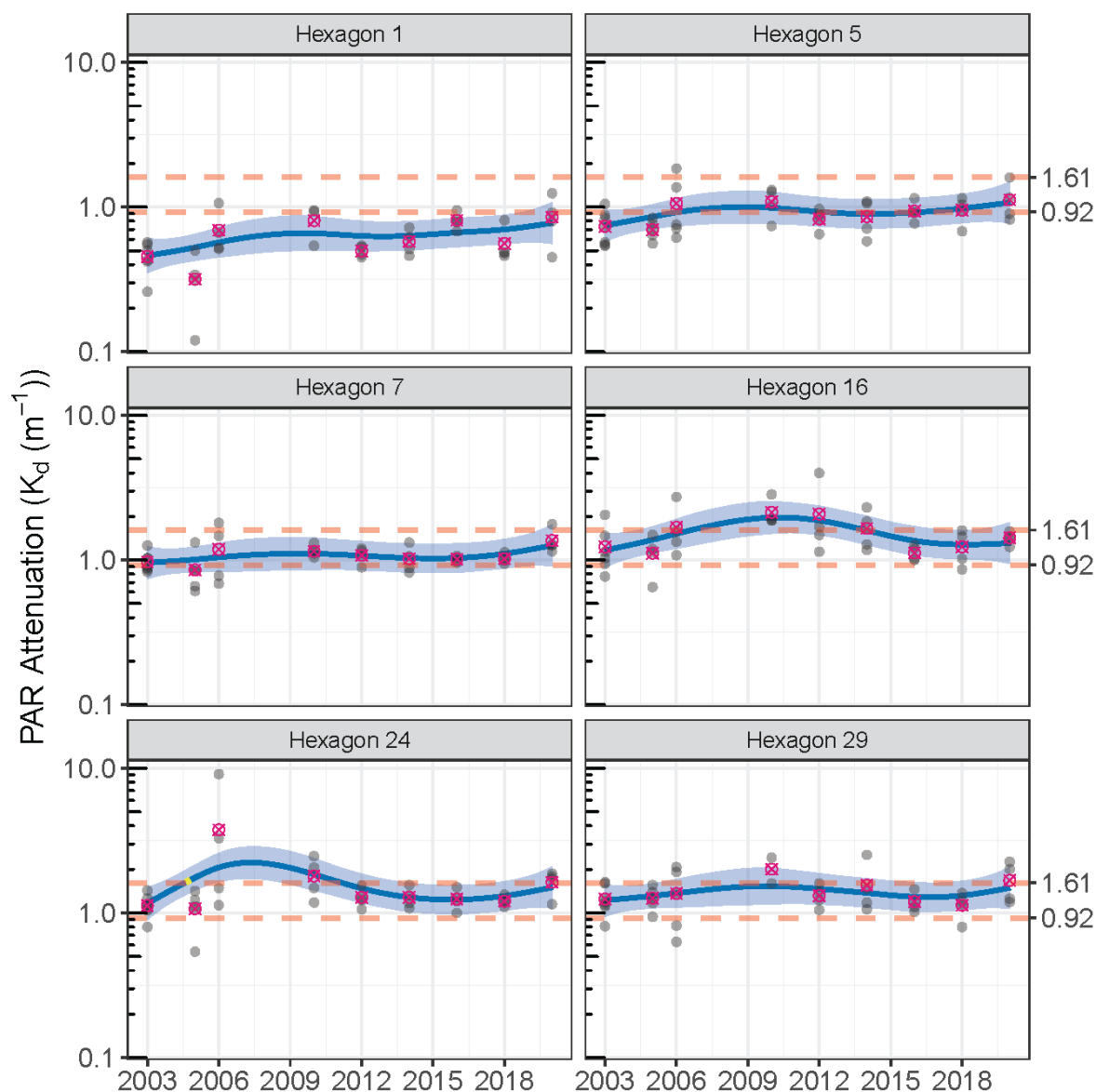


Figure 3-3. Mean PAR attenuation (K_d (m^{-1})) trend at each Gateway National Recreation Area trend station based on estimated generalized additive model smooths. The fitted smooth is plotted in dark blue with shaded 95% confidence interval. Periods of significant temporal change (increases/decreases), if present, are plotted in yellow/orange. Within each year, individual measurements and the average are plotted in gray and magenta, respectively. Estuarine quality category thresholds (Table 2-1) are plotted as red dashed lines.

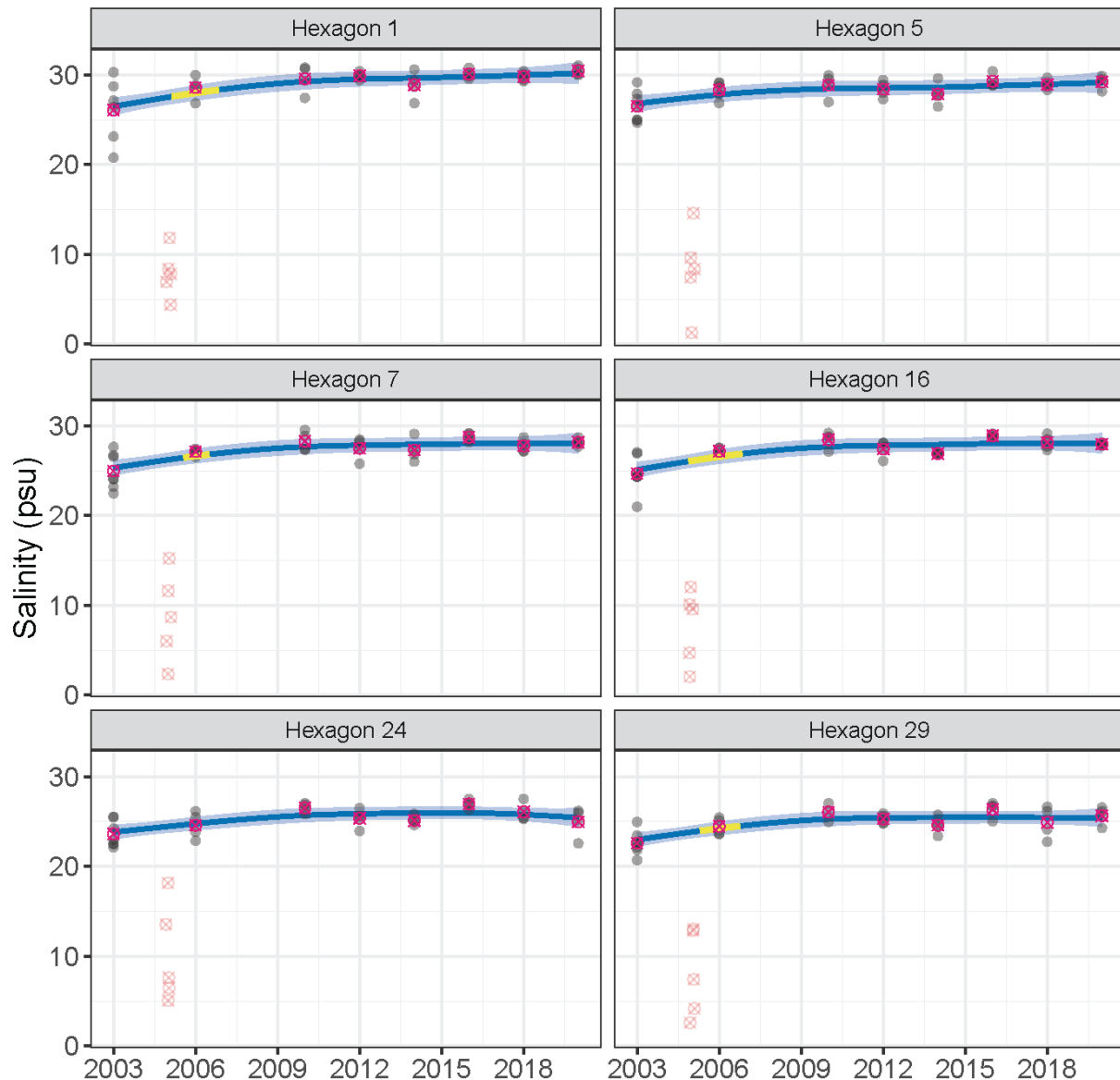


Figure 3-4. Mean surface salinity (psu) trend at each Gateway National Recreation Area trend station based on estimated generalized additive model smooths. The fitted smooth is plotted in dark blue with shaded 95% confidence interval. Periods of significant temporal change (increases/decreases), if present, are plotted in yellow/orange. Within each year, individual measurements and the average are plotted in gray and magenta, respectively. Note: observations that were excluded from model fitting are plotted in red.

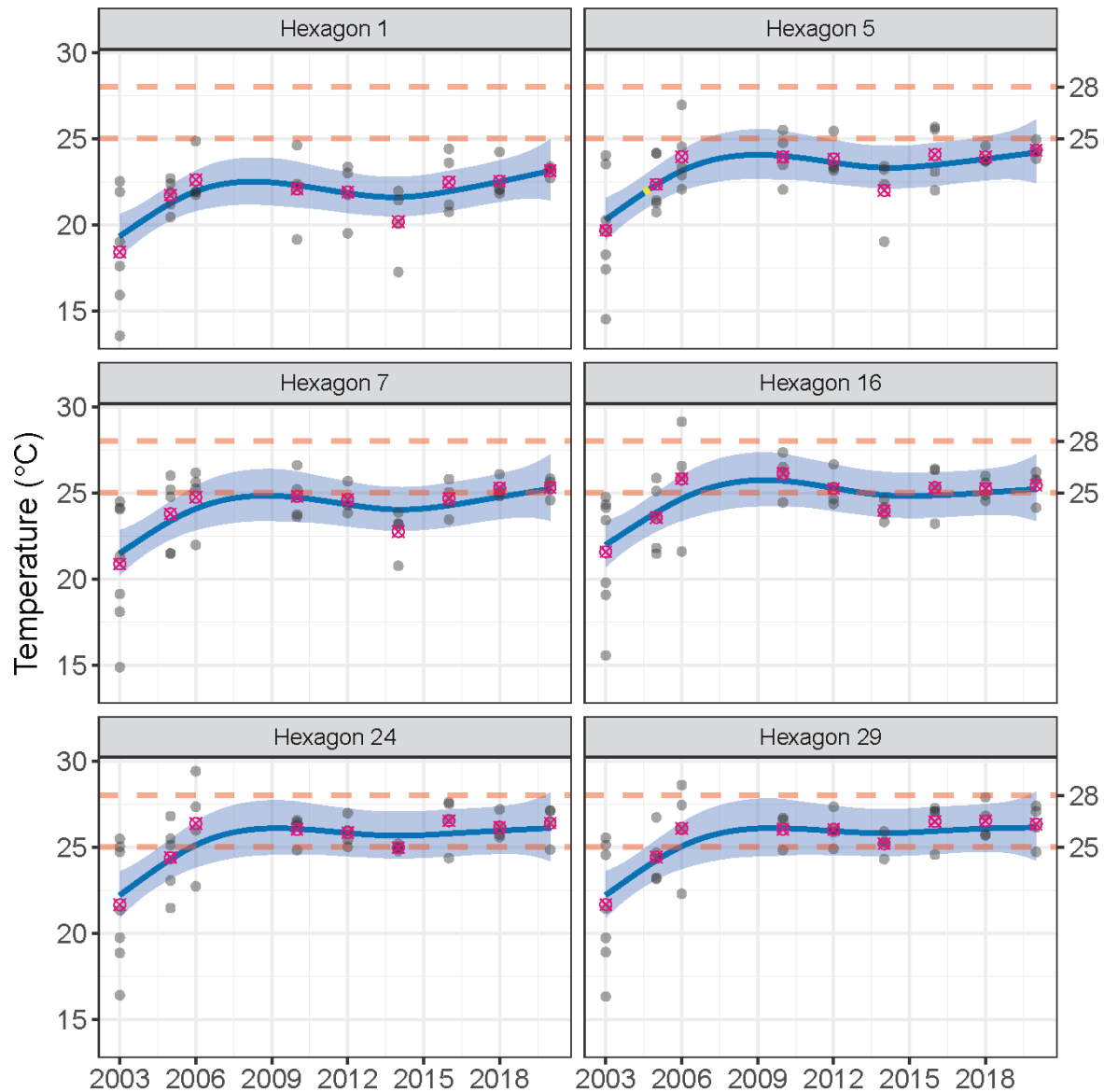


Figure 3-5. Mean surface temperature (°C) trend at each Gateway National Recreation Area trend station based on estimated generalized additive model smooths. The fitted smooth is plotted in dark blue with shaded 95% confidence interval. Periods of significant temporal change (increases/decreases), if present, are plotted in yellow/orange. Within each year, individual measurements and the average are plotted in gray and magenta, respectively. Estuarine quality category thresholds (Table 2-1) are plotted as red dashed lines.

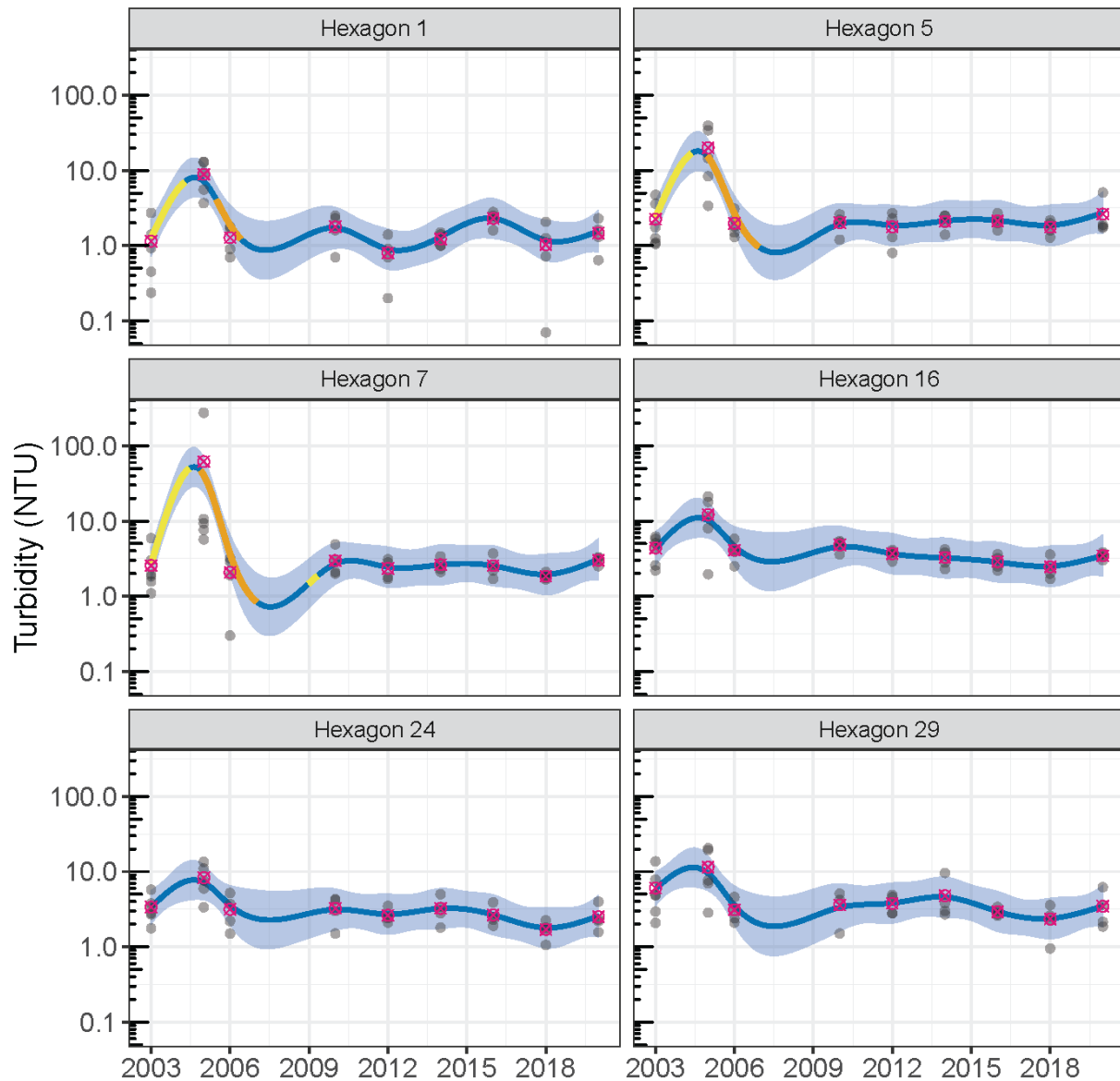


Figure 3-6. Mean surface turbidity (NTU) trend at each Gateway National Recreation Area trend station based on estimated generalized additive model smooths. The fitted smooth is plotted in dark blue with shaded 95% confidence interval. Periods of significant temporal change (increases/decreases), if present, are plotted in yellow/orange. Within each year, individual measurements and the average are plotted in gray and magenta, respectively. Note: the y-axis is transformed to the log₁₀ scale to better visualize the estimated trend.

3.4 Multivariate Mann-Kendall Trend Test Results

The multivariate Mann-Kendall test results indicate there is statistical evidence of an increasing monotonic trend in surface temperature (p -value = 0.021) and surface salinity (p -value = 0.033) during the index periods of 2003 to 2020 (Table 3-8).

Table 3-8. Summary of multivariate Mann-Kendall test for monotonic trend in each trend station measure: H_0 : no monotonic trend versus H_1 : monotonic trend is present. A positive (negative) value of the Mann-Kendall test statistic (Z_{MK}) indicates that measurements tend to increase (decrease) with time.

Measure	Depth	Z_{MK}	p -value
Chlorophyll <i>a</i> ($\mu\text{g/L}$)	Surface	-0.25	0.801
Dissolved oxygen (mg/L)	Bottom	-0.69	0.490
PAR attenuation (K_d (m^{-1}))	Water Column	1.52	0.128
Salinity (psu)	Surface	2.13	0.033
Temperature ($^{\circ}\text{C}$)	Surface	2.31	0.021
Turbidity (NTU)	Surface	-0.88	0.379

4. Logging-Station Continuous Water Quality Data

4.1 Continuous Record Plots

Plots of the entire record of continuous data collected at each logging-station are not included in this report due to the large volume of figures produced. We refer the reader to the folder /report/figures/ if inspection of the plots is desired. All plots follow the following filename format:

`ls_[park]_s[stratum]_[measure]_[year].pdf`

An example plot is provided in Figure 4-1.

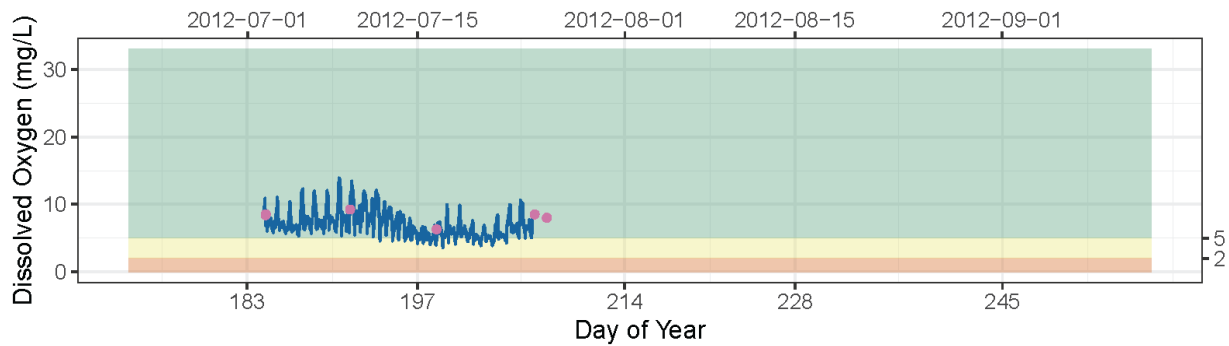


Figure 4-1. Continuous dissolved oxygen readings from sonde deployed at Gateway National Recreation Area logging station in 2012. Magenta points represent discrete grab samples collected every week for quality assurance. Estuarine quality category thresholds (Table 2-1) are delineated by the background color.

4.2 Methods

4.2.1 Generalized Additive Model (GAM)

Generalized additive models (GAMs) are used to characterize intra-annual trends in continuous logging station data. For each water quality measure, we model the mean response as a function of the day of year (doy) and the hour of day. Mathematically, the model can be written as:

$$\begin{aligned} \text{response}_i &\sim \text{Gamma}(\mu_i, \phi) \\ \log(\mu_i) &= \beta_0 + f_{\text{year}_i}(\text{doy}_i) + f(\text{hour}_i), \end{aligned}$$

where $f_{\text{year}_i}(\text{year})$ is a rank 20 TRPS smooth of doy for the i th year, $f(\text{hour})$ is a rank 24 cyclic cubic spline. The doy smoother is known as a factor-smoother and assumes that all factors (i.e., years) have the same smoothness, but may differ in shape over time. The smooth for hour is included to account for intra-day trends since many of the measures exhibit diel patterns (e.g., dissolved oxygen). Temporal autocorrelation was not modeled due to the high computational cost of doing so with 15k-45k observations and the inability to model such structure with the `bam()` function

in the *mgcv* R package. However, we do not believe the inability to model such structure impacts the fitted smooths included in the model.

Each model is fit using the *mgcv* package (Wood 2011) and results were summarized/visualized using the *gratia* package (Simpson 2020).

4.2.2 Additional Data Summaries

In addition to utilizing GAMs to visualize intra-annual trends, the following additional data summaries for each year's index period are included: (1) frequency distribution, and (2) number of hours per day measure was observed within each estuarine quality category (Table 2-1).

4.3 Results

Notable observations from the fitted smooths and data summaries for each measure are as follows:

4.3.1 Chlorophyll-*a*

Intra-annual mean trends in chlorophyll-*a* concentration are highly non-linear for the years 2011 and 2013 (Figure 4-2). The associated distribution of measurements for 2011 and 2013 are also highly skewed (Figure 4-3); however, skewness in 2011 is primarily due to the time series not being stationary. Chlorophyll-*a* condition is generally poor with occasional fair conditions (Figure 4-8). The highest mean chlorophyll-*a* concentration was observed in 2011 (107.27 $\mu\text{g/L}$) and the lowest (34.24 $\mu\text{g/L}$) in 2017 (Table 4-1).

4.3.2 Dissolved Oxygen

Intra-annual mean trends in dissolved oxygen exhibit a cyclic nature (Figure 4-2). Measurements in 2006 were much higher than all other years. Dissolved oxygen conditions are generally good with some days experiencing fair conditions for part of the day (Figure 4-9).

4.3.3 PAR Attenuation K_d

Mean trends tend to be more non-linear and variable during August compared July (Figure 4-2). The distribution of measurements in 2005 and 2020 are more right skewed compared to other years (Figure 4-3). PAR attenuation conditions during the day tend to be of either good or fair quality, with the former being more prevalent during July (Figure 4-10).

4.3.4 pH

Intra-annual mean trends for pH in 2018 appear similar to what was observed in 2020, although the period of record for 2020 is considerably shorter. Both years exhibited a non-linear intra-annual mean trend (Figure 4-6). Distributions of pH measurements are approximately symmetric (Figure 4-7).

4.3.5 Salinity

Intra-annual mean trends in salinity are non-linear and with some years exhibiting overall increasing or decreasing trends (Figure 4-4). Distributions of salinity measurements within each year tend to be left-skewed (Figure 4-5).

4.3.6 Temperature

Intra-annual mean trends in water temperature are highly non-linear and the mean tends to increase with increasing day of the year (Figure 4-4). At least half of the distribution of water temperature measurements have been above 25°C since 2016 (Figure 4-5).

4.3.7 Turbidity

Intra-annual mean trends in turbidity are non-linear and tend to be increasing during the month of July (Figure 4-4). The range of turbidity measurements was greatest during 2005, with observations spanning approximately 0 to 100 NTU (Figure 4-5). Measurements 2003 and 2014 also exhibiting larger than usual ranges.

Table 4-1. Yearly summary statistics for measures collected at the Gateway National Recreation Area logging station. [SD = standard deviation]

Measure	Statistic	2003	2005	2006	2010	2012	2014	2016	2018	2020
Chlorophyll <i>a</i> (µg/L)	Mean	5.87	79.66	55.53	65.87	26.22	18.51	16.09	25.91	41.41
Chlorophyll <i>a</i> (µg/L)	SD	5.10	49.72	24.44	39.95	16.71	10.44	6.60	14.56	16.71
Chlorophyll <i>a</i> (µg/L)	n	3012	4654	643	1932	2104	2770	2512	3435	1429
Depth (m)	Mean	NA	NA	3.02	1.87	1.34	2.53	2.43	2.31	2.56
Depth (m)	SD	NA	NA	0.71	0.54	0.58	0.58	0.60	0.56	0.53
Depth (m)	n	0	0	647	1935	2120	2782	5466	3438	1429
dissolved oxygen (mg/L)	Mean	6.57	8.04	27.36	8.21	7.10	6.25	6.32	6.55	6.33
dissolved oxygen (mg/L)	SD	2.21	1.35	1.58	2.30	1.86	1.21	0.87	1.34	1.42
dissolved oxygen (mg/L)	n	3004	877	621	1935	2120	2781	2515	3438	1429
PAR attenuation (K_d (m ⁻¹))	Mean	1.18	1.08	0.62	0.85	0.81	0.83	0.75	0.63	1.20
PAR attenuation (K_d (m ⁻¹))	SD	0.46	0.53	0.29	0.30	0.33	0.32	0.29	0.24	0.40
PAR attenuation (K_d (m ⁻¹))	n	495	736	145	513	462	675	1299	798	705
pH	Mean	NA	NA	NA	NA	NA	NA	NA	7.85	7.88
pH	SD	NA	NA	NA	NA	NA	NA	NA	0.13	0.11
pH	n	0	0	0	0	0	0	0	3438	1429
Salinity (psu)	Mean	26.87	27.25	28.56	29.19	27.75	27.47	29.35	27.87	28.25
Salinity (psu)	SD	0.67	0.42	0.54	0.41	0.48	0.54	0.48	0.56	0.50
Salinity (psu)	n	3008	2168	622	1935	2120	2782	2515	3430	1429
Temperature (°C)	Mean	24.09	25.61	24.30	23.94	24.53	23.22	24.82	25.12	25.34
Temperature (°C)	SD	0.97	0.75	0.64	1.17	1.07	0.89	1.41	1.09	0.89
Temperature (°C)	n	3013	2168	649	1935	2120	2782	5466	3438	2768
Turbidity (NTU)	Mean	2.69	9.48	2.86	2.97	3.02	2.00	3.08	4.16	5.33
Turbidity (NTU)	SD	7.32	15.76	1.60	0.81	2.99	1.35	0.79	1.48	1.60
Turbidity (NTU)	n	3007	4665	645	1908	2108	2743	2501	3431	1427

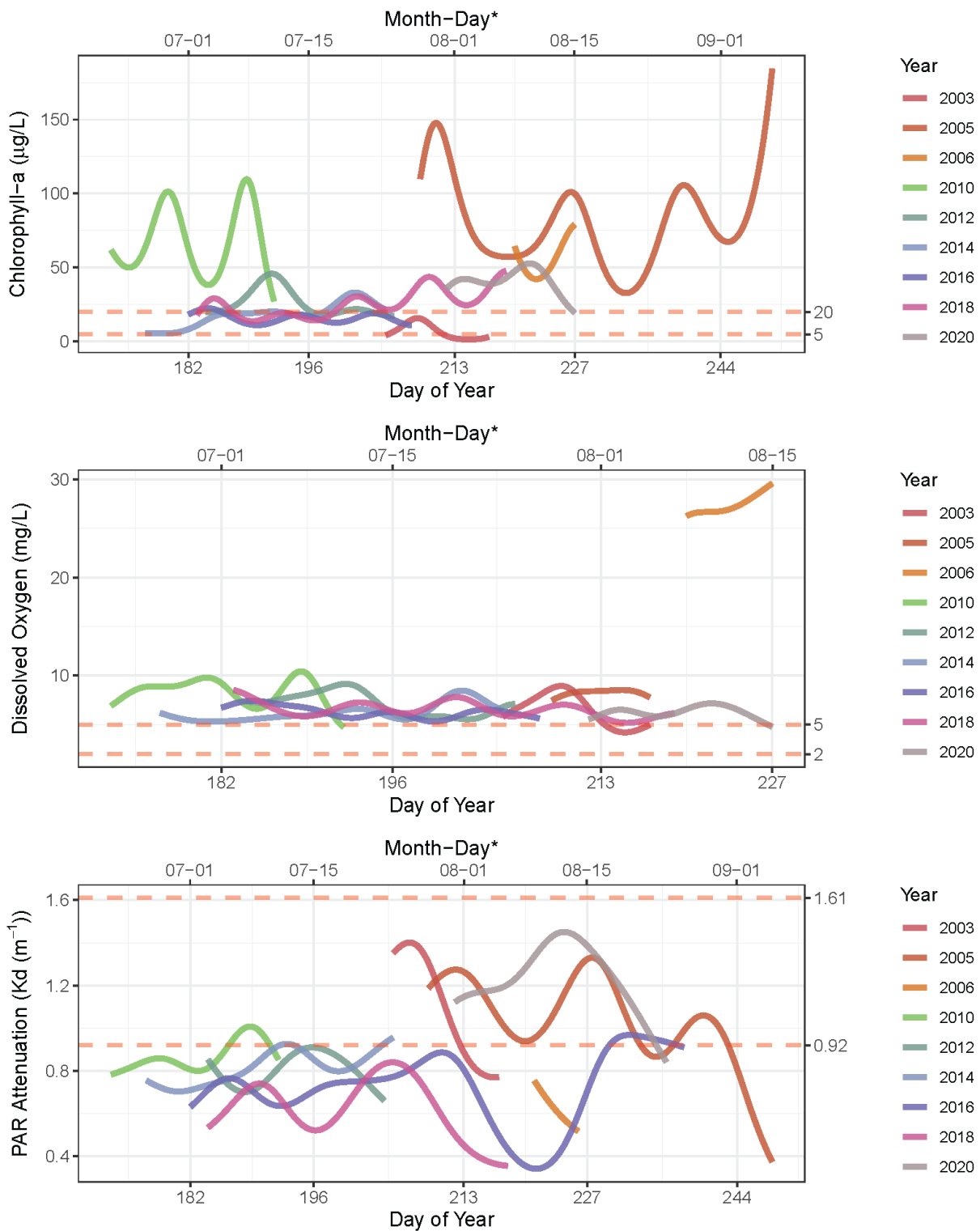


Figure 4-2. Intra-annual trends in surface chlorophyll-*a* concentration (top), bottom dissolved oxygen concentration (middle), and PAR attenuation (bottom) at the Gateway National Recreation Area logging station based on estimated generalized additive model smooths. Estuarine quality category thresholds (Table 2-1) plotted as red dashed lines.

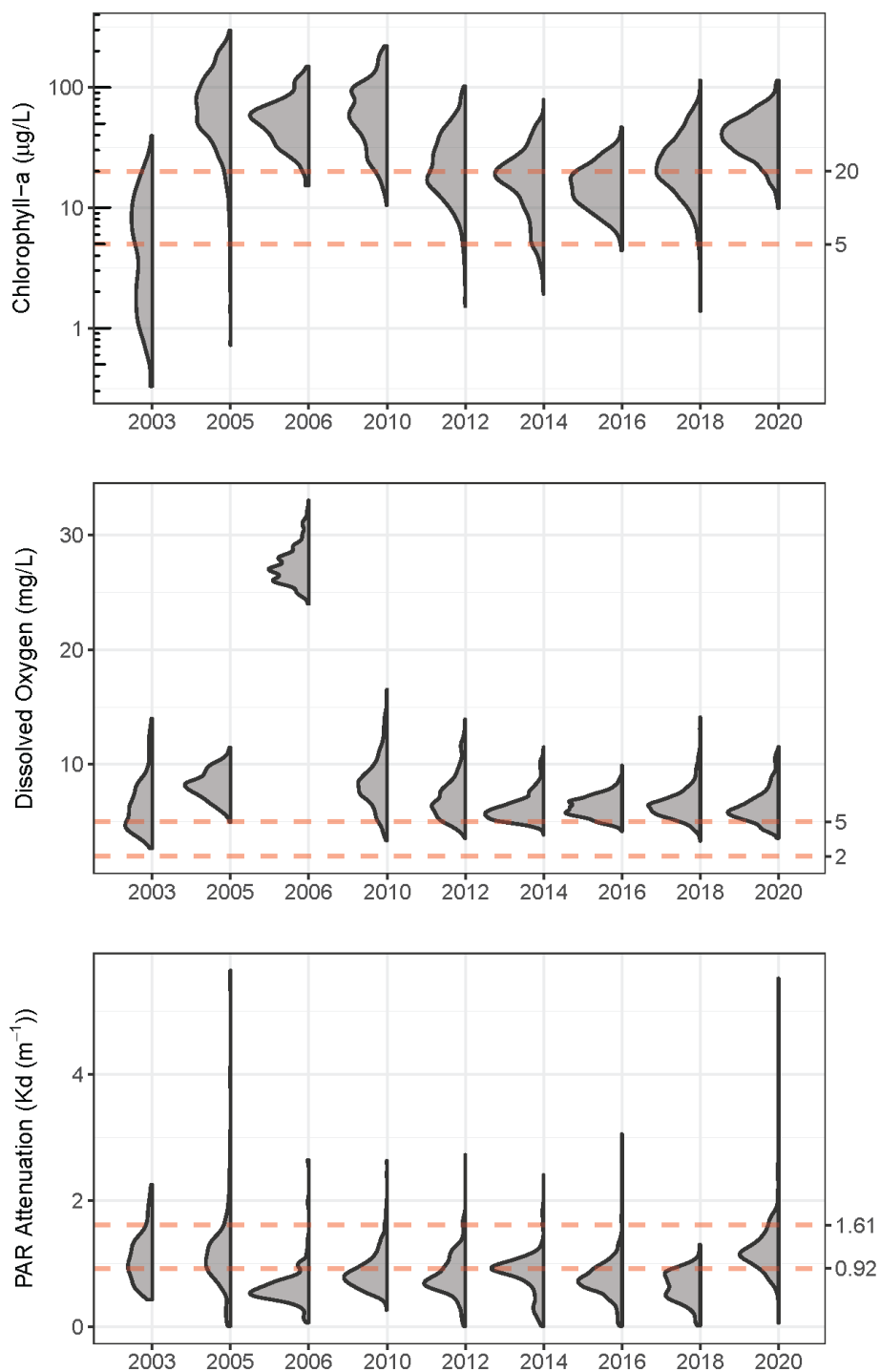


Figure 4-3. Distribution of index period time series values for surface chlorophyll-a concentration (top), bottom dissolved oxygen concentration (middle), and PAR attenuation (bottom) at the Gateway National Recreation Area logging station. Estuarine quality category thresholds (Table 2-1) plotted as red dashed lines. Note: the surface chlorophyll-a concentration y-axis is transformed to the $\log_{10}$ scale to better visualize the estimated distributions.

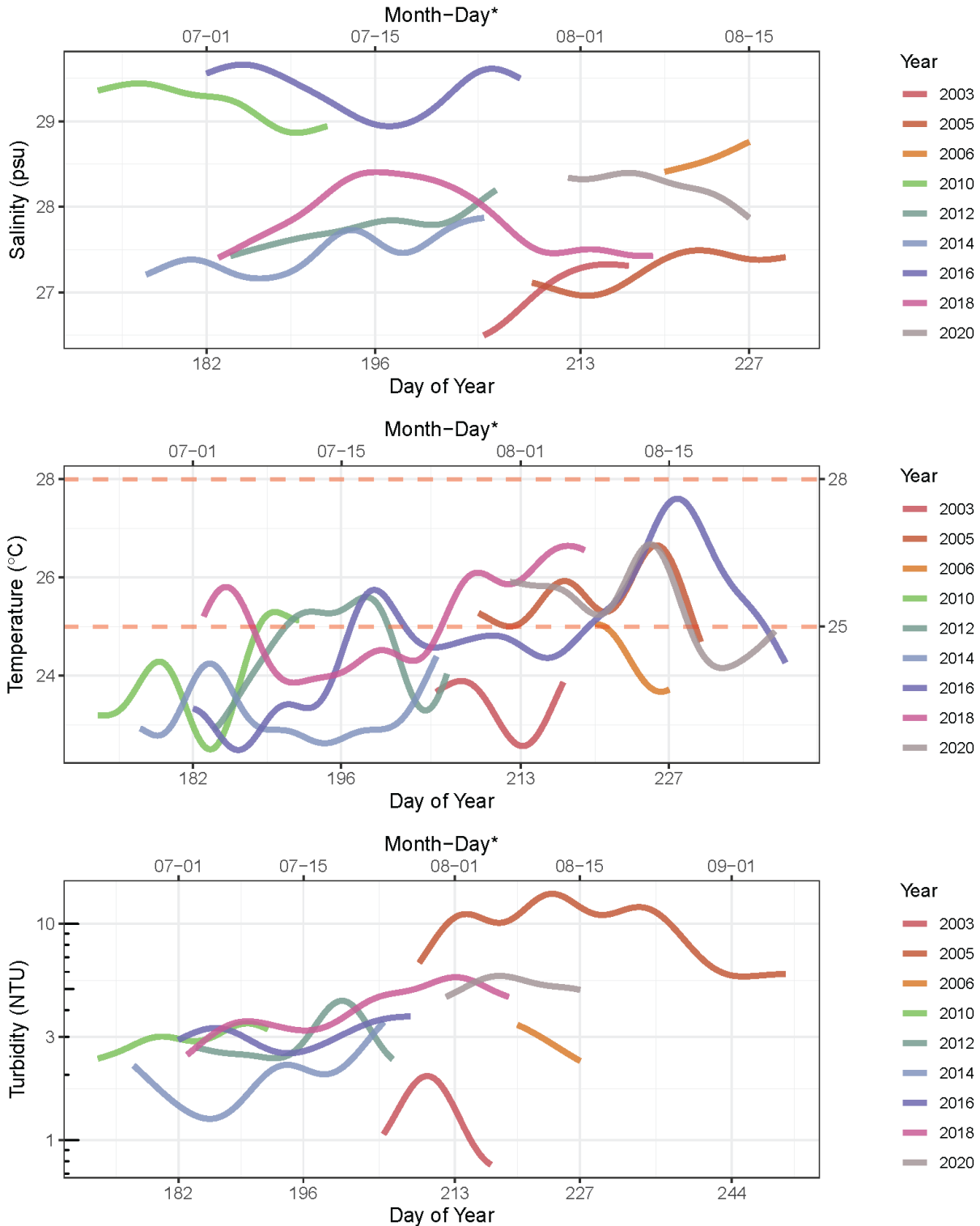


Figure 4-4. Intra-annual trends in surface salinity (top), surface temperature (middle), and surface turbidity (bottom) at the Gateway National Recreation Area logging station based on estimated generalized additive model smooths. Estuarine quality category thresholds (Table 2-1) plotted as red dashed lines. Note: the surface turbidity y-axis is transformed to the log₁₀ scale to better visualize the estimated trends.

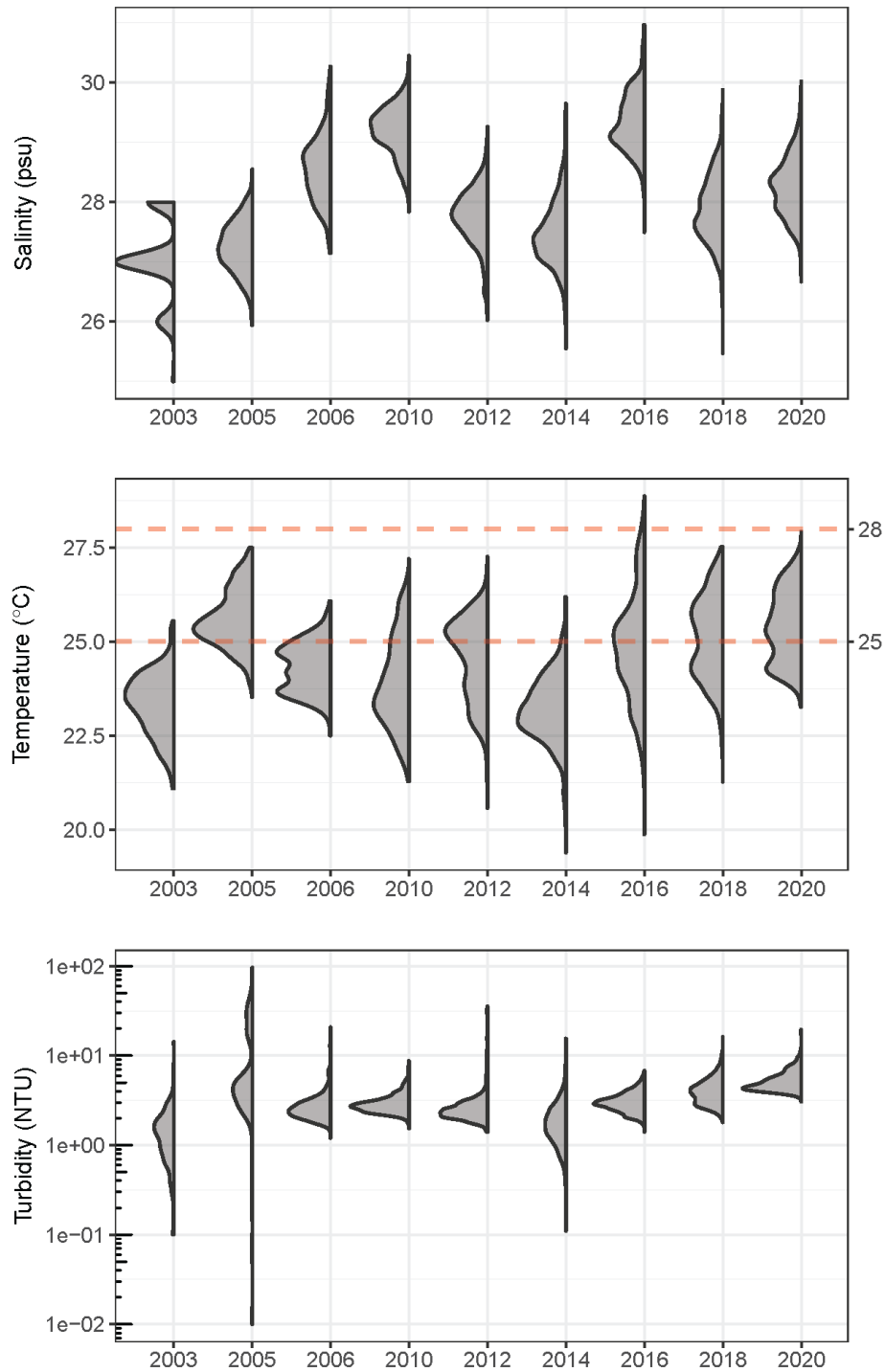


Figure 4-5. Distribution of index period time series values for surface salinity (top), surface temperature (middle), and surface turbidity (bottom) at the Gateway National Recreation Area logging station. Estuarine quality category thresholds (Table 2-1) plotted as red dashed lines. Note: the surface turbidity y-axis is transformed to the log₁₀ scale to better visualize the estimated distributions.

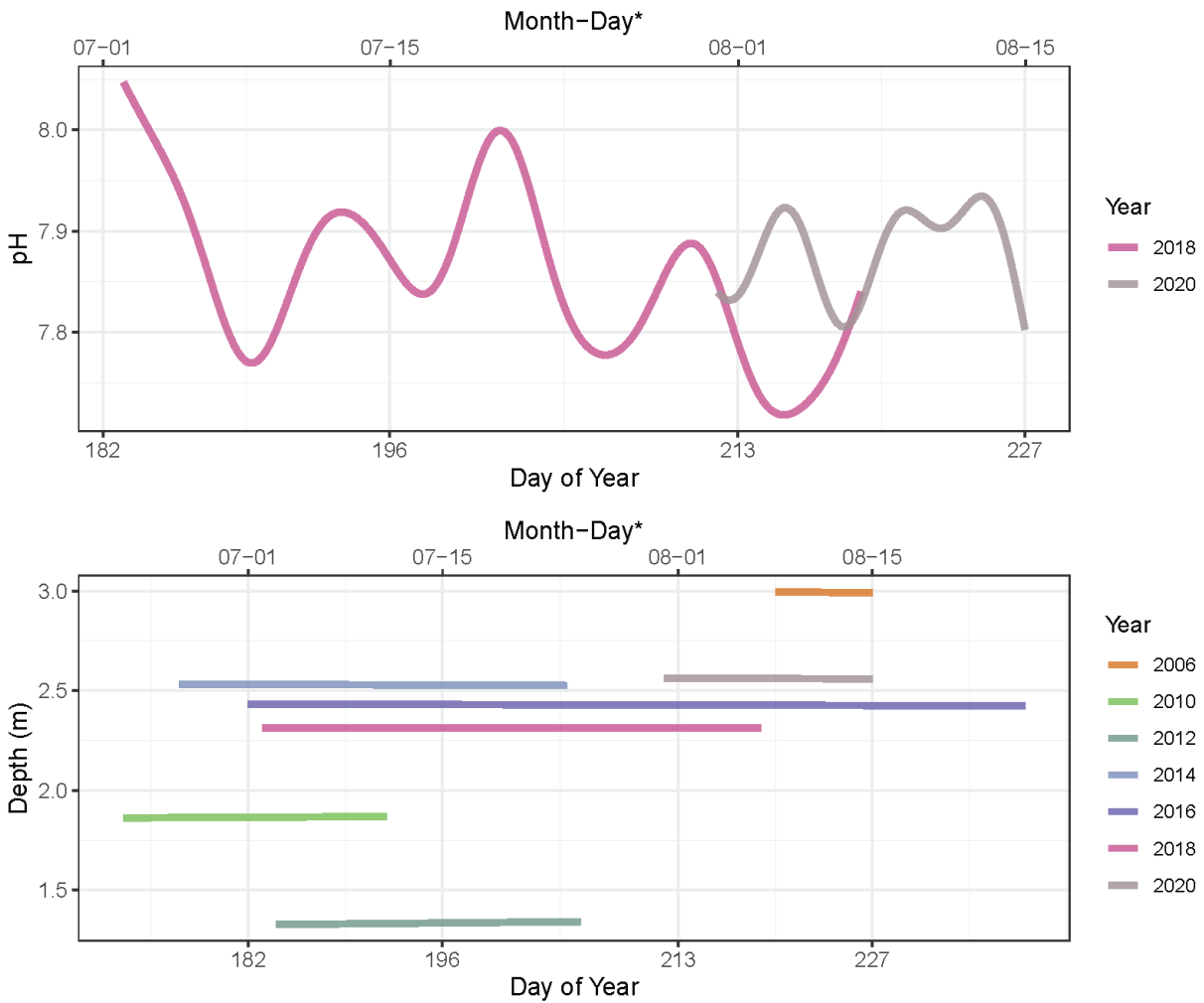


Figure 4-6. Intra-annual trends in pH (top) and depth (bottom) at the Gateway National Recreation Area logging station based on estimated generalized additive model smooths.

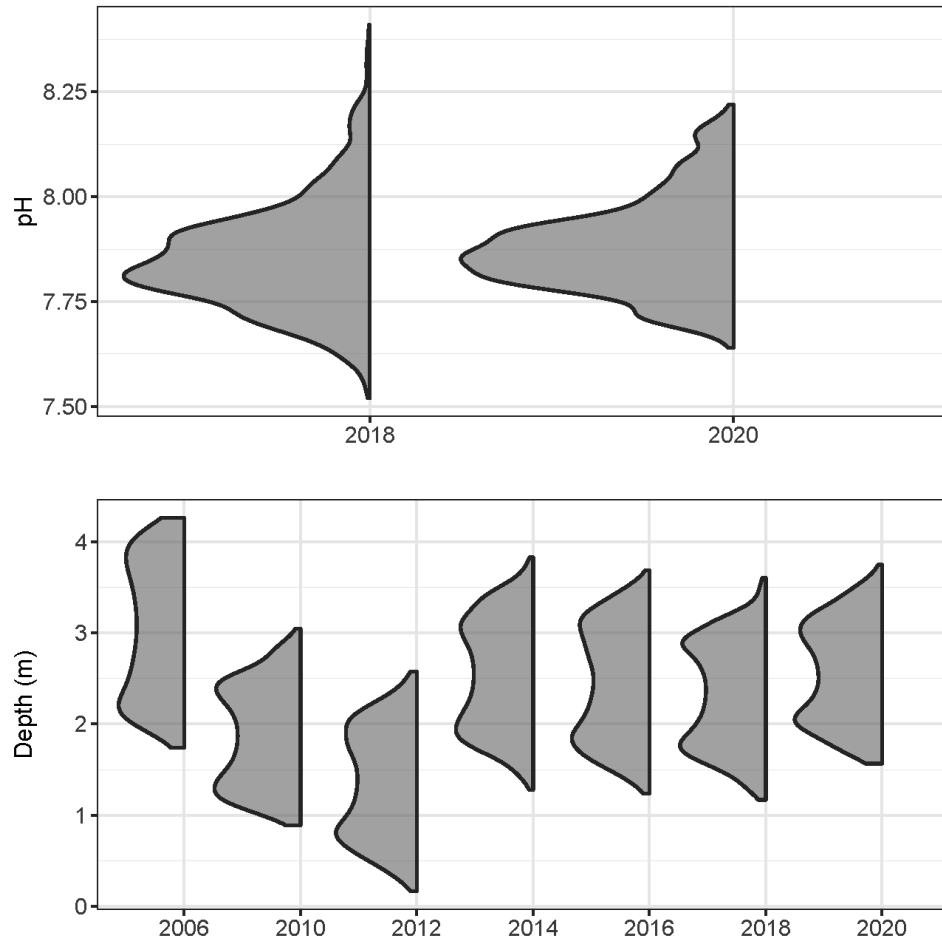


Figure 4-7. Distribution of index period time series values for pH (top) and depth (bottom) at the Gateway National Recreation Area logging station.

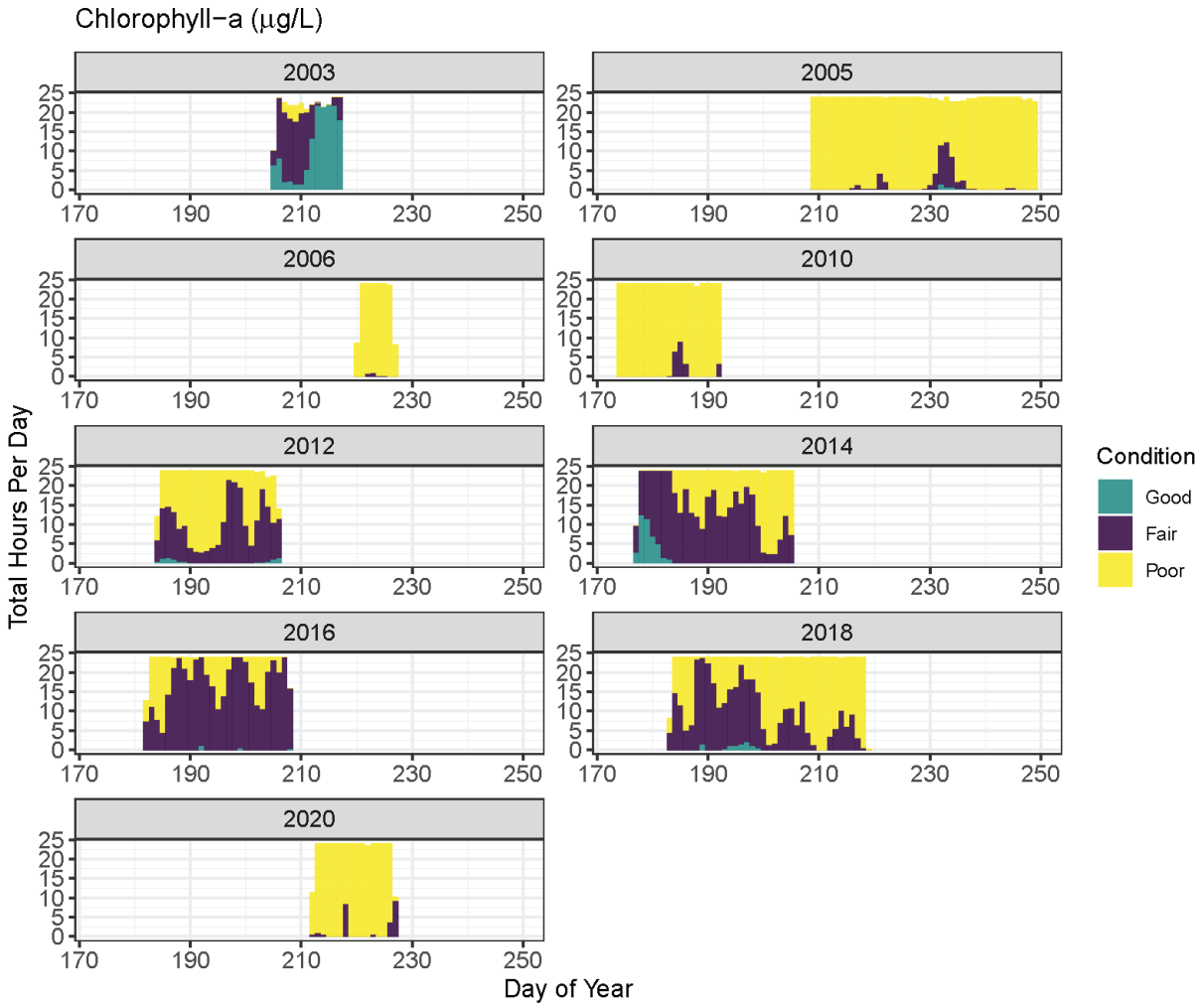


Figure 4-8. Total hours per day that surface chlorophyll-a concentration was observed within each condition category at the Gateway National Recreation Area logging station.

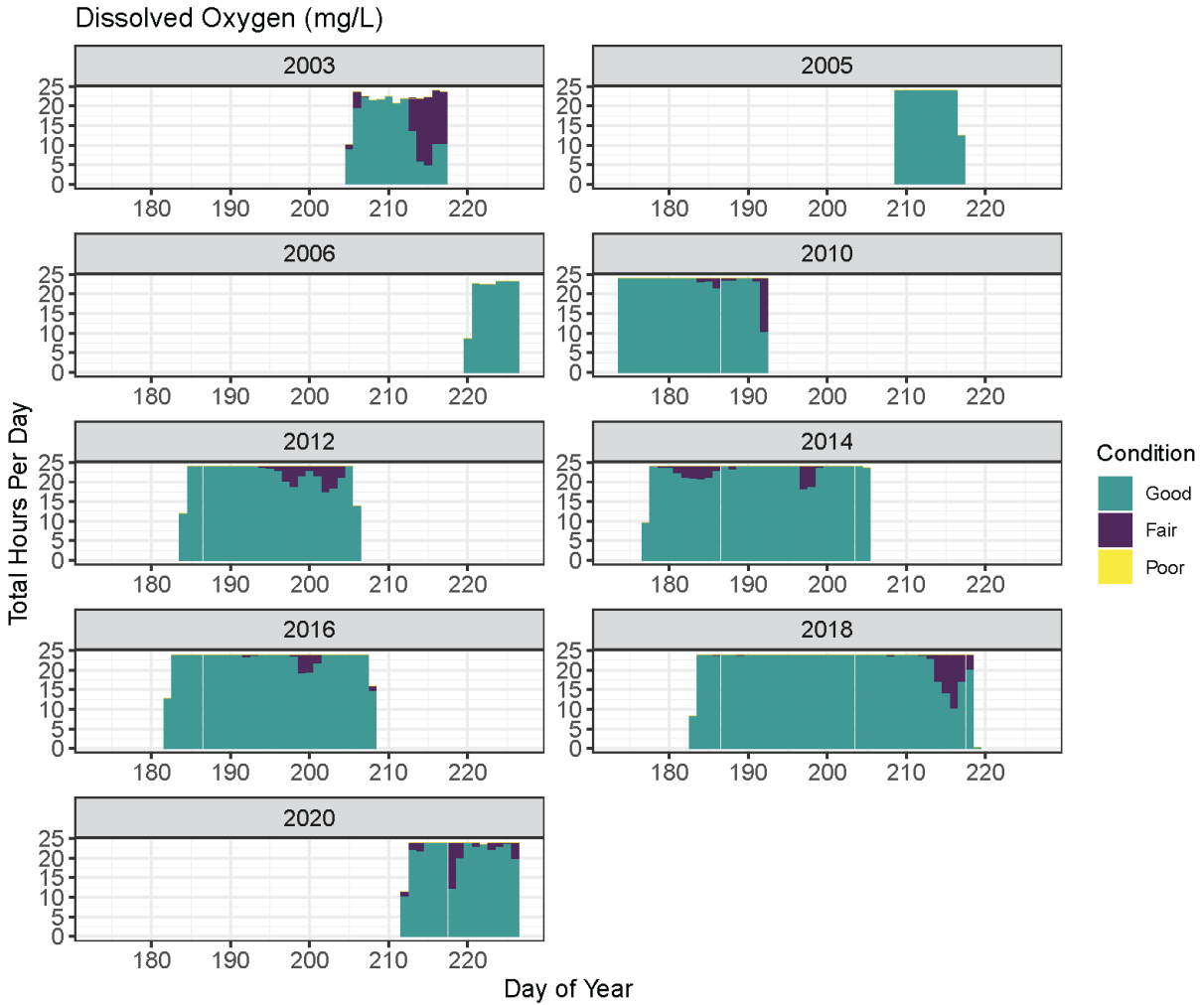


Figure 4-9. Total hours per day that bottom dissolved oxygen concentration was observed within each condition category at the Gateway National Recreation Area logging station.

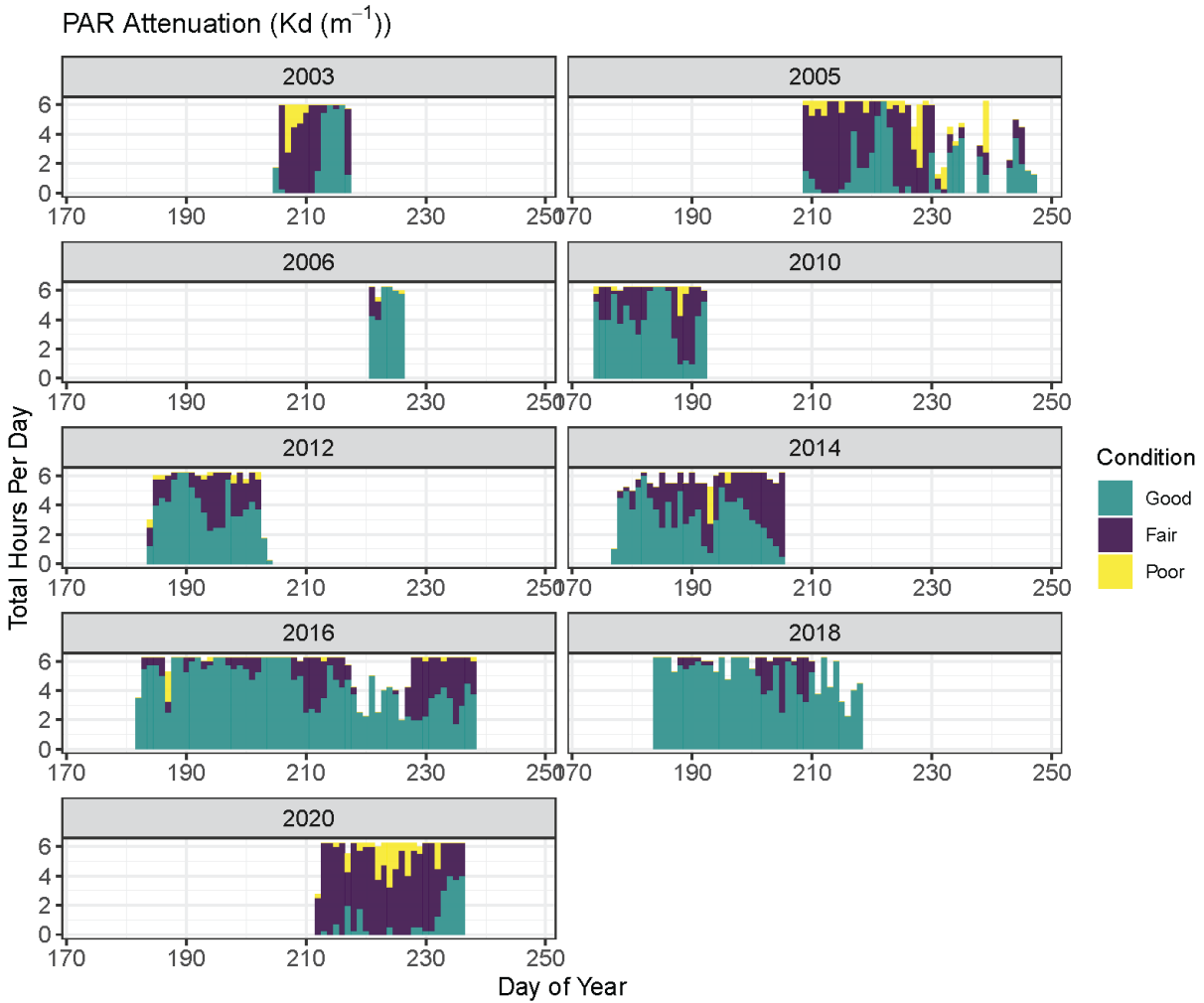


Figure 4-10. Total hours per day that PAR attenuation was observed within each condition category at the Gateway National Recreation Area logging station.

5. Power Simulations

5.1 Methods

Power analyses are a function of the following five parameters: (1) power: the probability of rejecting the null hypothesis when it is false; (2) sample size; (3) type I error rate (alpha level): the probability of rejecting the null hypothesis given that it is true; (4) magnitude of the trend coefficient; and (5) variance of the measure. Increasing the value of the sample size, type I error rate, or trend coefficient, while holding the remaining parameters constant, will result in increased power. Decreasing the variance, while holding the remaining parameters constant, will also result in increased power. If given values for four of the parameters, it is possible to solve for the remaining parameter value.

The empirical power to detect a linear trend in the mean over time is estimated using Monte Carlo simulation following the method detailed in Gelman and Hill (2006). Power simulations are performed using the spatial-station monitoring data for the following measures: surface chlorophyll-*a* concentration, bottom dissolved oxygen concentration, and PAR attenuation.

The simulation framework⁴ consists of two steps: (1) use a process model to generate a time series with a true underlying trend with dispersion (variance) approximately equal to what has been observed in past monitoring data; (2) use estimation model to estimate the trend. The systematic component of the process model is defined mathematically as:

$$\log(\mu_i) = \beta_0 + \beta_1 \cdot \text{year}_i + f(\text{longitude}_i, \text{latitude}_i),$$

where β_1 is the linear trend and $f(\text{longitude}_i, \text{latitude}_i)$ is the predicted spatial effect estimated from the models fit in Section 2. Simulated locations ($i = 1, \dots, 30$) within each tessellated hexagon were randomly sampled with replacement from all previous locations within the hexagon. The random component of the simulated trend consists of response values being drawn from:

$$\text{response}_i \sim \text{Gamma}(\mu_i, \phi),$$

where the dispersion parameter ϕ equals to the estimated dispersion (i.e., residual variation) from the corresponding model fit in Section 2 (Table 5-1).

Trends were simulated for a period of both 10 and 20 years. The sampling interval was simulated to occur every 1-3 years for the 10-year simulation and every 1-5 years for the 20-year simulation. The magnitude of the linear trend (β_1) was varied so as to sufficiently characterize the estimated power curve. It is possible that past estimates of variability may be too conservative and not representative

⁴ An example of a similar framework applied to monitoring brown bear density in southwestern Alaska is detailed in Reynolds, Thompson, and Russell (2011).

of future variability due to changing climate. Therefore, additional scenarios wherein the dispersion was doubled were also included in the simulations.

To estimate the empirical power for a given simulation scenario, one thousand hypothetical time series were simulated according to the process model defined above. The estimation model consisted of a GAM with year as a parametric term and a Gaussian process smooth of spatial location within the estuary. Empirical power is equal to the proportion of the thousand regressions that rejected the null hypothesis $H_0: \beta_1 = 0$ in favor of $H_1: \beta_1 \neq 0$, assuming a type I error rate equal to 0.05.

Note that while the simulated mean trend is linear on the link function scale, back-transforming to the original scale results in a non-linear mean trend (Figure 5-1). Interpretation of the β_1 parameter is as follows: for every 1-year increase in time, the mean of the response values increases by a factor of $\exp(\beta_1)$. For example, $\beta_1 = 0.05$ in Figure 5-1 which corresponds to an approximate 5.1% increase in the mean each year ($\exp(0.05) \approx 1.051$).

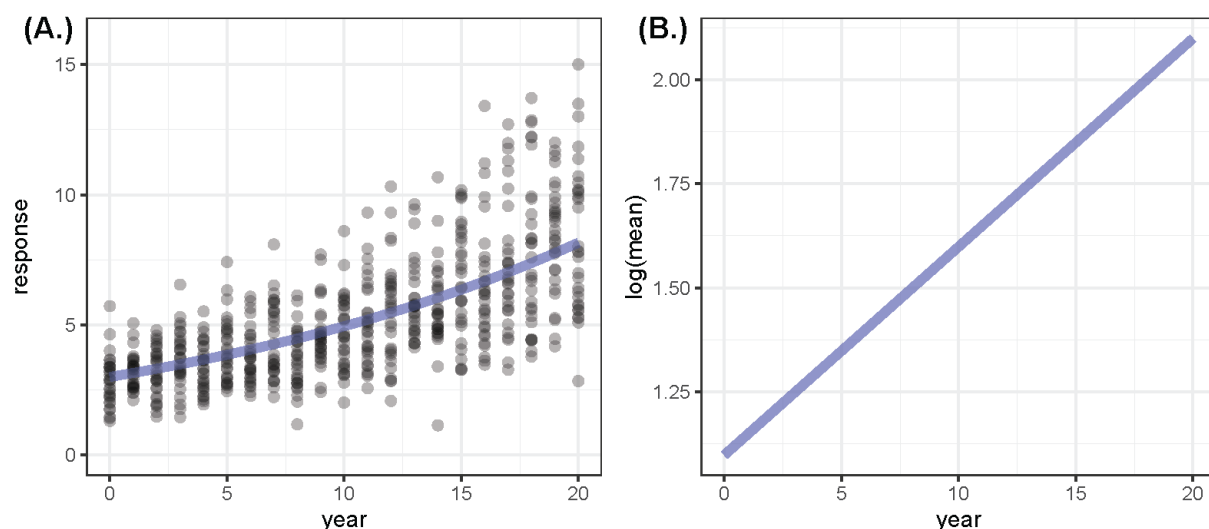


Figure 5-1. (A.) Process model realizations (black dots) representing samples from 30 sites within a hypothetical estuary. The i th hexagon's response value is simulated from $\text{Gamma}(\mu_i, 0.1)$, where $\ln(\mu_i) = 3 + 0.05 \cdot \text{year}_i$. The blue curve represents the true underlying mean trend. (B.) True underlying trend on the link function scale.

Table 5-1. Process model dispersion (ϕ) used to simulate trends in each water quality measure.

Measure	ϕ
Chlorophyll <i>a</i> ($\mu\text{g/L}$)	0.307
Dissolved oxygen (mg/L)	0.201
PAR attenuation (K_d (m^{-1}))	0.102

5.2 Simulation Results

Estimated power for detecting a linear trend is summarized in both graphical and tabular format for surface chlorophyll-*a* concentration (Figure 5-2; Table 5-2), bottom dissolved oxygen concentration (Figure 5-3; Table 5-3), and PAR attenuation (Figure 5-4; Table 5-4).

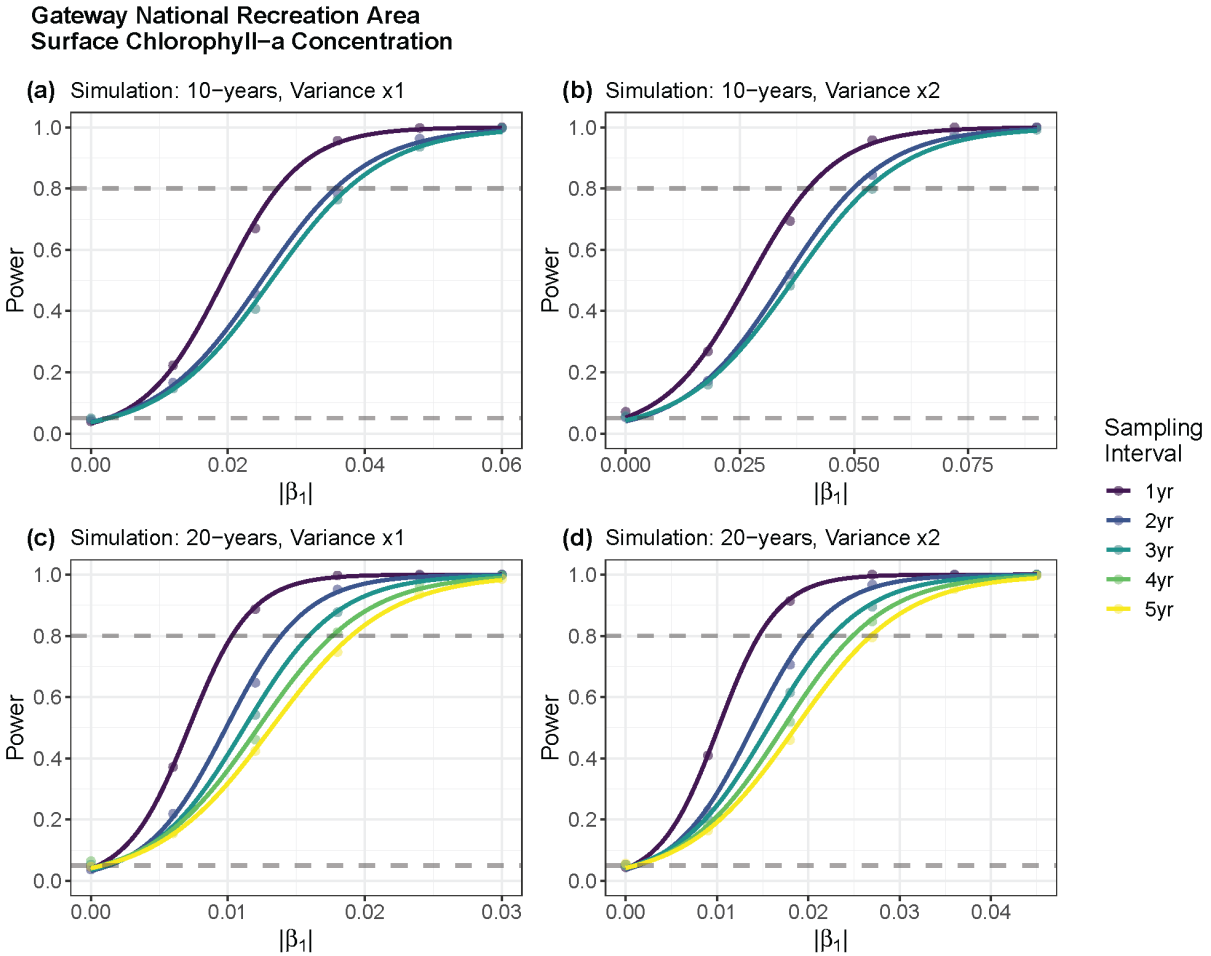


Figure 5-2. Estimated power to detect a linear trend in mean surface chlorophyll-*a* concentration within Gateway National Recreation Area. Power estimates vary based on the assumed sampling interval (see legend) and duration: (a) 10-year period, (c) 20-year period. Power estimates in (b) and (d) correspond to simulations wherein the process model variance has been doubled. Note: the x-axis may not be consistent across simulations because the linear trend (β_1) within each simulation was varied so as to sufficiently characterize the estimated power curve.

Table 5-2. Estimated power to detect a linear trend in mean surface chlorophyll-*a* concentration large enough to cross both quality thresholds within Gateway National Recreation Area. The β_1 column value corresponds to the magnitude of trend that would result in the mean equaling the stated threshold value, assuming the stated duration and process variance multiplier. For reference, the most current estimate of mean surface chlorophyll-*a* concentration from the generalized additive model fit in Section 2.4 is provided in the table footnotes.

Period	Variance	Threshold	β_1	Power by Sampling Interval				
				1yr	2yr	3yr	4yr	5yr
10 yrs	×1	5	−0.151	1.00	1.00	1.00	NA	NA
10 yrs	×1	20	−0.013	0.24	0.17	0.15	NA	NA
10 yrs	×2	5	−0.151	1.00	1.00	1.00	NA	NA
10 yrs	×2	20	−0.013	0.18	0.12	0.12	NA	NA
20 yrs	×1	5	−0.076	1.00	1.00	1.00	1.00	1.00
20 yrs	×1	20	−0.006	0.41	0.23	0.19	0.18	0.16
20 yrs	×2	5	−0.076	1.00	1.00	1.00	1.00	1.00
20 yrs	×2	20	−0.006	0.23	0.14	0.14	0.12	0.11

**Gateway National Recreation Area
Bottom Dissolved Oxygen Concentration**

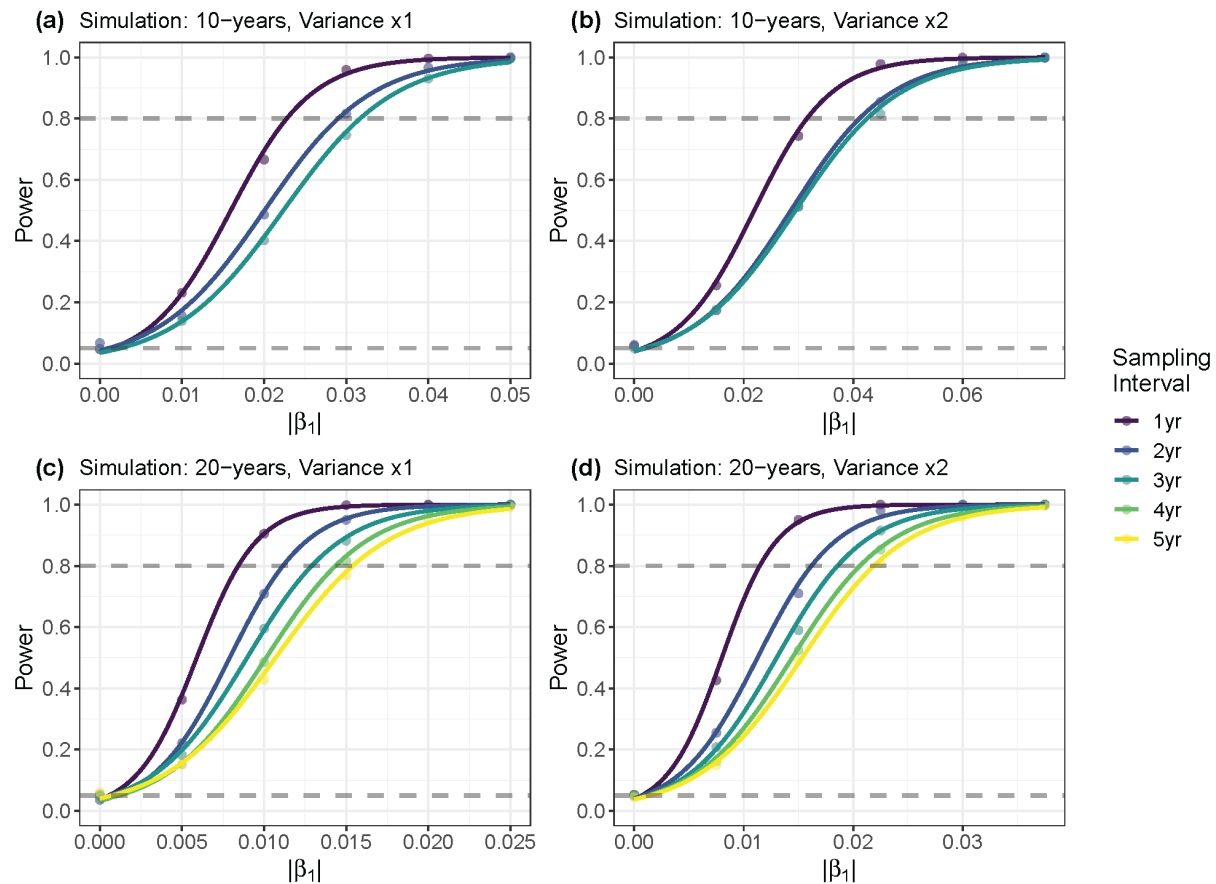


Figure 5-3. Estimated power to detect a linear trend in mean bottom dissolved oxygen concentration within Gateway National Recreation Area. Power estimates vary based on the assumed sampling interval (see legend) and duration: (a) 10-year period, (c) 20-year period. Power estimates in (b) and (d) correspond to simulations wherein the process model variance has been doubled. Note: the x-axis may not be consistent across simulations because the linear trend (β_1) within each simulation was varied so as to sufficiently characterize the estimated power curve.

Table 5-3. Estimated power to detect a linear trend in mean bottom dissolved oxygen concentration large enough to cross both quality thresholds within Gateway National Recreation Area. The β_1 column value corresponds to the magnitude of trend that would result in the mean equaling the stated threshold value, assuming the stated duration and process variance multiplier. For reference, the most current estimate of mean bottom dissolved oxygen concentration from the generalized additive model fit in Section 2.4 is provided in the table footnotes

Period	Variance	Threshold	β_1	Power by Sampling Interval				
				1yr	2yr	3yr	4yr	5yr
10 yrs	×1	2	−0.110	1.00	1.00	1.00	NA	NA
10 yrs	×1	5	−0.019	0.64	0.45	0.37	NA	NA
10 yrs	×2	2	−0.110	1.00	1.00	1.00	NA	NA
10 yrs	×2	5	−0.019	0.39	0.25	0.24	NA	NA
20 yrs	×1	2	−0.055	1.00	1.00	1.00	1.00	1.00
20 yrs	×1	5	−0.009	0.87	0.65	0.54	0.44	0.40
20 yrs	×2	2	−0.055	1.00	1.00	1.00	1.00	1.00
20 yrs	×2	5	−0.009	0.64	0.37	0.29	0.24	0.22

Gateway National Recreation Area PAR Attenuation

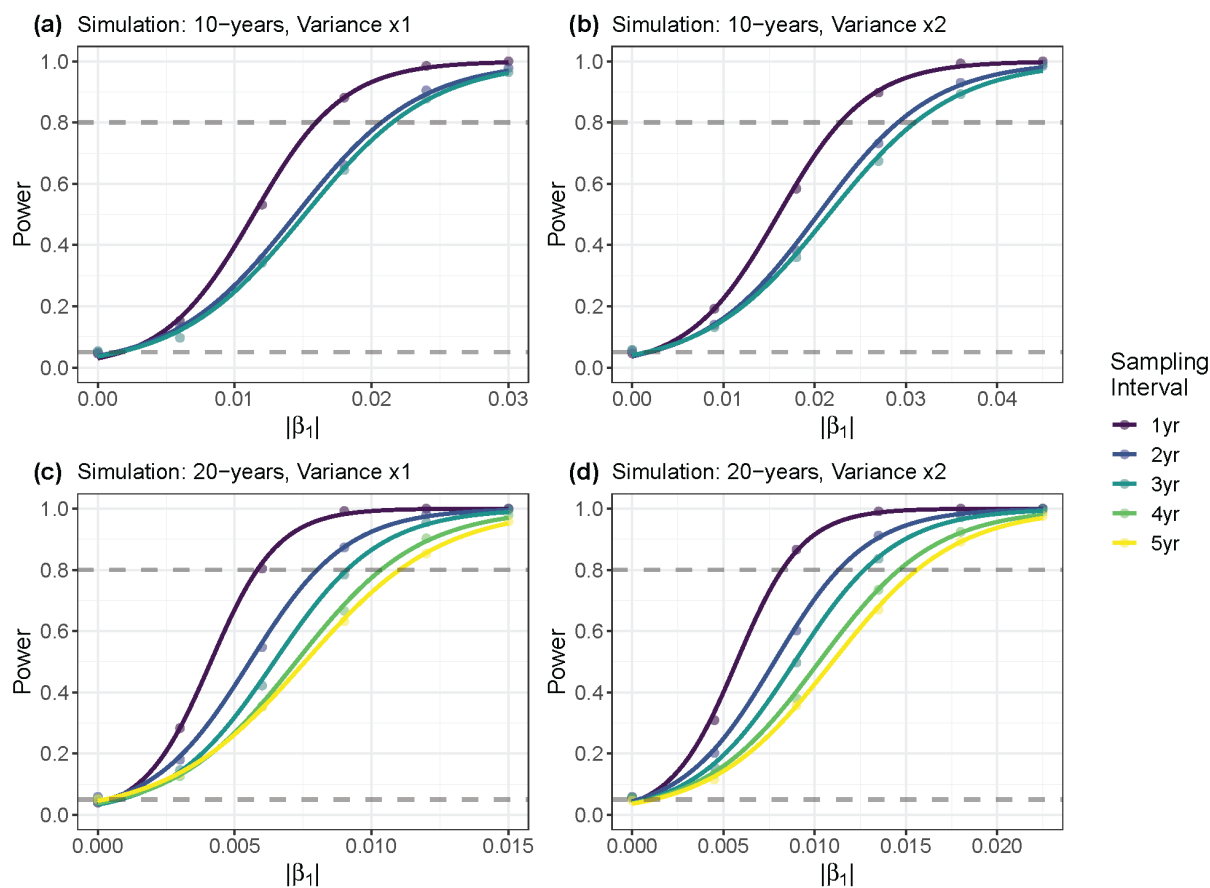


Figure 5-4. Estimated power to detect a linear trend in mean PAR attenuation within Gateway National Recreation Area. Power estimates vary based on the assumed sampling interval (see legend) and duration: (a) 10-year period, (c) 20-year period. Power estimates in (b) and (d) correspond to simulations wherein the process model variance has been doubled. Note: the x-axis may not be consistent across simulations because the linear trend (β_1) within each simulation was varied so as to sufficiently characterize the estimated power curve.

Table 5-4. Estimated power to detect a linear trend in mean PAR attenuation large enough to cross both quality thresholds within Gateway National Recreation Area. The β_1 column value corresponds to the magnitude of trend that would result in the mean equaling the stated threshold value, assuming the stated duration and process variance multiplier. For reference, the most current estimate of mean PAR attenuation from the generalized additive model fit in Section 2.4 is provided in the table footnotes.

Period	Variance	Threshold	β_1	Power by Sampling Interval				
				1yr	2yr	3yr	4yr	5yr
10 yrs	×1	0.92	−0.020	0.94	0.79	0.76	NA	NA
10 yrs	×1	1.61	0.036	1.00	0.99	0.99	NA	NA
10 yrs	×2	0.92	−0.020	0.71	0.50	0.46	NA	NA
10 yrs	×2	1.61	0.036	0.98	0.92	0.89	NA	NA
20 yrs	×1	0.92	−0.010	0.99	0.93	0.88	0.79	0.74
20 yrs	×1	1.61	0.018	1.00	1.00	1.00	0.99	0.98
20 yrs	×2	0.92	−0.010	0.92	0.72	0.62	0.50	0.44
20 yrs	×2	1.61	0.018	1.00	0.98	0.96	0.92	0.88

5.3 Discussion

5.3.1 Key Question #1

Is the frequency of sampling appropriate or are we over or under sampling?

Power analyses were performed to assess the adequacy of the sampling interval for detecting estuary-wide trends in surface chlorophyll-*a* concentration, bottom dissolved oxygen concentration, and PAR attenuation.

Prior to determining whether decreasing the sampling frequency has merit, we must first consider the following: (1) what magnitude of change over a period of 10 or 20 years is ecologically meaningful and/or large enough to warrant corrective action by managers? and (2) how important are shorter-term fluctuations/trends in the decision-making by managers and/or for understanding the current state of the system? Regarding item (1), it is important to remember that although the β_1 values are relatively small when interpreting change on a yearly basis, these changes are compounded over time. For example, a β_1 value of 0.02 corresponds to a yearly multiplicative increase of approximately 2% (i.e., $\exp(0.02) \approx 1.02$) in the mean of the measure. However, over a period of 10 and 20 years, this corresponds to an increase of approximately 22.1% and 49.2%, respectively. Regarding item (2), it is important to remember that the power to detect trend is estimated assuming a true mean trend that is both constant and linear through time for a period of 10 or 20 years (see Figure 5-1). Accordingly, all power estimates provided in this report do not reflect the ability to detect shorter-term trends or fluctuations.

Based on discussions with NCBN scientists, a measure's change from one condition category to another is ecologically meaningful and could potentially warrant corrective action by park managers. Accordingly, the magnitude of trend (β_1) that would result in a measure changing to another condition category after a period of 10 or 20 years was computed using the established condition thresholds and the most recent year's estimate of the measure's mean. Power for each of these scenarios was then estimated by identifying where the β_1 values intersect the power curve. Estimated

power to detect changes from one threshold to another have been summarized in tabular format following each power curve figure.

In general, power to detect a change to the condition category will be impacted by:

1. How close the current estimate is to the next threshold - the closer it is, the lower the power to detect the change will be (all else equal).
2. The residual (i.e., process) variance - the larger it is, the lower the power to detect the change will be (all else equal).
3. The time frame over which changes need to be detected - the larger it is, the higher the power to detect the change will be (all else equal).

For example, consider the power summary for surface chlorophyll-*a* concentration (Table 5-2). From the table's footnotes, the estimated mean chlorophyll-*a* concentration in 2018 was 22.69 $\mu\text{g/L}$. Looking across all simulation scenarios, we see that there is low power to detect changes from the poor to fair category (i.e., 20 $\mu\text{g/L}$ threshold) and high power to detect changes from poor to good category (i.e., 5 $\mu\text{g/L}$ threshold), regardless of sampling interval. The recommended sampling interval in these simulation scenarios is dependent on which change is more relevant from a management perspective. If changes to the next condition category are of interest (i.e., poor to fair), then the recommendation for each scenario would be to sample every year. Since power to detect such a change is estimated to be below 80%, from a practical standpoint we'd likely instead conclude that the sampling interval should not be decreased rather than recommend increasing the interval from every 2-years to every year.

Because measures are sampled concurrently, power to detect changes in all three measures should be considered. Applying the process in the preceding paragraph to each measure, we'd arrive at the following table of results:

Measure (current condition)	Time Period + Process Variance Multiplier			
	10 yrs + $\times 1$	10 yrs + $\times 2$	20 yrs + $\times 1$	20 yrs + $\times 2$
Surface chlorophyll- <i>a</i> (poor)	1 yr	1 yr	1 yr	1 yr
Bottom dissolved oxygen (good)	1 yr	1 yr	1 yr	1 yr
PAR attenuation (fair)	3 yr	3 yr	5 yr	5 yr
Recommendation*	1 yr	1 yr	1 yr	1 yr

*Based on minimum sampling interval within given scenario with at least 80% power.

Assuming that each measure is equally important to detect changes in, the recommendation under each scenario (i.e., column) would be to sample every year, or as frequently as resources permit. Again, these results correspond to detecting changes from the 'good' to 'fair' category for bottom dissolved oxygen, from 'poor' to 'fair' in surface chlorophyll-*a*, and from 'fair' to 'poor' in PAR attenuation.

The ENE protocol is applicable on a network-wide scale and it is logistically undesirable to vary the sampling interval for each individual park. Given that the recommendation for at least one park/estuary is to maintain or increase the sampling interval, the network-wide recommendation would be to not decrease the current sampling interval and, if resources permit, increase the sampling interval.

Literature Cited

- Caldwell, James M., Matthew E. Nixon, Hilary A. Neckles, and Penelope S Pooler. 2015. Estuarine Water Quality in Parks of the Northeast Coastal and Barrier Network: Vital Signs Estuarine Nutrient-Enrichment Monitoring, 2006-11. National Park Service.
- Gelman, Andrew, and Jennifer Hill. 2006. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge university press.
- Greve, Tina Maria, Jens Borum, and Ole Pedersen. 2003. Meristematic Oxygen Variability in Eelgrass (*Zostera Marina*). *Limnology and Oceanography* 48 (1):210–16.
- Grizzle, Raymond E, and Richard A Lutz. 1905. Descriptions of Macroscopic Banding Patterns in Sectioned Polished Shells of *Mercenaria Mercenaria* from Southern New Jersey. *Journal of Shellfish Research*.
- Hefley, Trevor J., Kristin M. Broms, Brian M. Brost, Frances E. Buderman, Shannon L. Kay, Henry R. Scharf, John R. Tipton, Perry J. Williams, and Mevin B. Hooten. 2017. The Basis Function Approach for Modeling Autocorrelation in Ecological Data. *Ecology* 98(3):632–46.
- Kopp, Blaine S., Martha Nielsen, Dejan Glisic, and Hilary A. Neckles. 2009. Estuarine Water Quality in Parks of the Northeast Coastal and Barrier Network: Development and Early Implementation of Vital Signs Estuarine Nutrient-Enrichment Monitoring, 2003-06. Natural Resource Technical Report 266.
- Kopp, B. S., and H. A. Neckles. 2009. A Protocol for Monitoring Estuarine Nutrient Enrichment in Coastal Parks of the National Park Service Northeast Region. NPS/NCBN/NRR–2009/110. Fort Collins, CO: National Park Service.
- Lenth, Russell. 2020. Emmeans: Estimated Marginal Means, Aka Least-Squares Means. <https://CRAN.R-project.org/package=emmeans>.
- Libiseller, Claudia, and Anders Grimvall. 2002. Performance of Partial Mann–Kendall Tests for Trend Detection in the Presence of Covariates. *Environmetrics: The Official Journal of the International Environmetrics Society* 13(1):71–84.
- Pinheiro, Jose, Douglas Bates, Saikat DebRoy, Deepayan Sarkar, and R Core Team. 2020. nlme: Linear and Nonlinear Mixed Effects Models. <https://CRAN.R-project.org/package=nlme>.
- Pohlert, Thorsten. 2020. Trend: Non-Parametric Trend Tests and Change-Point Detection. <https://CRAN.R-project.org/package=trend>.
- Ramsey, Fred, and Daniel Schafer. 2012. The Statistical Sleuth: A Course in Methods of Data Analysis. Cengage Learning.

- R Core Team. 2020. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Reusch, Thorsten B. H., Anneli Ehlers, August Hämmerli, and Boris Worm. 2005. Ecosystem Recovery After Climatic Extremes Enhanced by Genotypic Diversity. *Proceedings of the National Academy of Sciences* 102 (8):2826–31.
- Reynolds, Joel H., William L. Thompson, and Brook Russell. 2011. Planning for Success: Identifying Effective and Efficient Survey Designs for Monitoring. *Biological Conservation* 144(5):1278–84.
- Roman, Charles T., Norbert Jaworski, Frederick T. Short, Stuart Findlay, and R. Scott Warren. 2000. Estuaries of the Northeastern United States: Habitat and Land Use Signatures. *Estuaries* 23(6):743–64.
- Shields, Erin C, Dave Parrish, and Kenneth Moore. 2019. Short-Term Temperature Stress Results in Seagrass Community Shift in a Temperate Estuary. *Estuaries and Coasts* 42(3):755–64.
- Simpson, Gavin L. 2018. Modelling Palaeoecological Time Series Using Generalised Additive Models. *Frontiers in Ecology and Evolution* 6:149.
- Simpson, Gavin L. 2020. Gratia: Graceful 'Ggplot'-Based Graphics and Other Functions for Gams Fitted Using 'Mgcv'. <https://CRAN.R-project.org/package=gratia>.
- Stevens, S., B. Milstead, M. Albert, and G. Entsminger. 2005. Northeast Coastal and Barrier Network Vital Signs Monitoring Plan. NPS/NER/NRTR–2005/025. Boston, MA: National Park Service.
- Wood, Simon N. 2003. Thin Plate Regression Splines. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 65(1):95–114.
- Wood, Simon N. 2011. Fast Stable Restricted Maximum Likelihood and Marginal Likelihood Estimation of Semiparametric Generalized Linear Models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 73(1):3–36.
- Wood, Simon N. 2017. *Generalized Additive Models: An Introduction with R*. CRC press.

Appendix A

Short Primer on Generalized Additive Models

First, consider the following multiple linear regression model (Equation 1):

$$\mathbf{y} = \beta_0 + \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \boldsymbol{\epsilon},$$

where $\mathbf{y}$ is a vector of response values, $\mathbf{x}_1$ and $\mathbf{x}_2$ are covariate vectors, β_0 , β_1 and β_2 are the model parameters to be estimated, and $\boldsymbol{\epsilon}$ is the vector of error terms such that $\boldsymbol{\epsilon} \sim N(0, \sigma^2)$. Generalized additive models (GAMs) extend the standard linear model by allowing the model to fit non-linear functions (smooshs) in place of the linear components $\beta_j \mathbf{x}_j$. The GAM version of Equation 1 is written mathematically as:

$$\mathbf{y} = \beta_0 + f_1(\mathbf{x}_1) + f_2(\mathbf{x}_2) + \boldsymbol{\epsilon},$$

where f_1 and f_2 are non-linear functions. The model is an additive model because a separate smooth is computed for each covariate and contributions are then added together. This is advantageous from an inferential perspective because it means we are able to examine the effect of each covariate on the response, while holding all of the other covariates fixed.

Each f_j is represented parametrically within the model as a sum of K basis functions⁵:

$$f_j(\mathbf{x}_j) = \sum_{k=1}^K \beta_{j,k} b_{j,k}(\mathbf{x}_j),$$

where $b_{j,k}$ is the k th basis function and $\beta_{j,k}$ are estimated basis coefficients. Although there are a multitude of basis functions that can be used, smoothing splines are the most common. For reference, the default smoothing spline in the `mgcv` package is the thin plate regression spline (TPRS; (Wood 2003)). The value of K is referred to as the rank of the basis expansion and represents an upper limit on the expected complexity of the true underlying trend in the data. Furthermore, the upper limit is constrained to be less than or equal to the number of data points. In practice, the actual value of K specified is not critical; instead it must simply be large enough to provide the smooth enough flexibility to approximate the true underlying function f . The `mgcv::gam.check()` function provides a test that can be used to determine whether the rank should be increased (at the expense of additional computation time).

In order to prevent overfitting to the data, the degree of wiggleness of a smooth is controlled through the use of a wiggleness penalty and smoothing parameter. The smoothness parameter (denoted by λ) controls the extent to which the penalty contributes to the likelihood function. Simpson (2018) recommends that λ be estimated via restricted maximum likelihood (REML) for best results. Since

⁵ See Hefley et al. (2017) for an introduction to basis functions with applications to ecology.

each smooth is being penalized/constrained, the complexity of the smooth is measured using effective degrees of freedom (edf). The edf can be a non-integer value and higher values indicate a more wiggly smooth.

To provide a more tangible example of a GAM and its associated basis expansion, we will simulate some data and fit a GAM to it. Simulated data are generated using the nonlinear model (Equation 2):

$$y_i = \exp(x_i \cdot \cos(4\pi \cdot x_i)) + \epsilon_i,$$

where $x_i \sim \text{Uniform}(0, 0.8)$, $\epsilon_i \sim \text{Normal}(0, 0.2)$, and $i = 1, \dots, 400$ (see Figure A-1)

Using the `mgcv` package, we can fit a GAM with a rank 6 TPRS basis and REML smoothness selection with the following code:

```
> gam(y ~ s(x, k = 6, bs = "tp"), family = "gaussian", method = "REML")
```

The fitted smooth $\hat{f}(x)$ is plotted in black in panel (b) of Figure A-1. We can see that it does a pretty good job of approximating the true underlying function $f(x)$ (red curve). Additionally, Figure A-1 (c) depicts the six basis functions that are added together to form $\hat{f}(x)$.

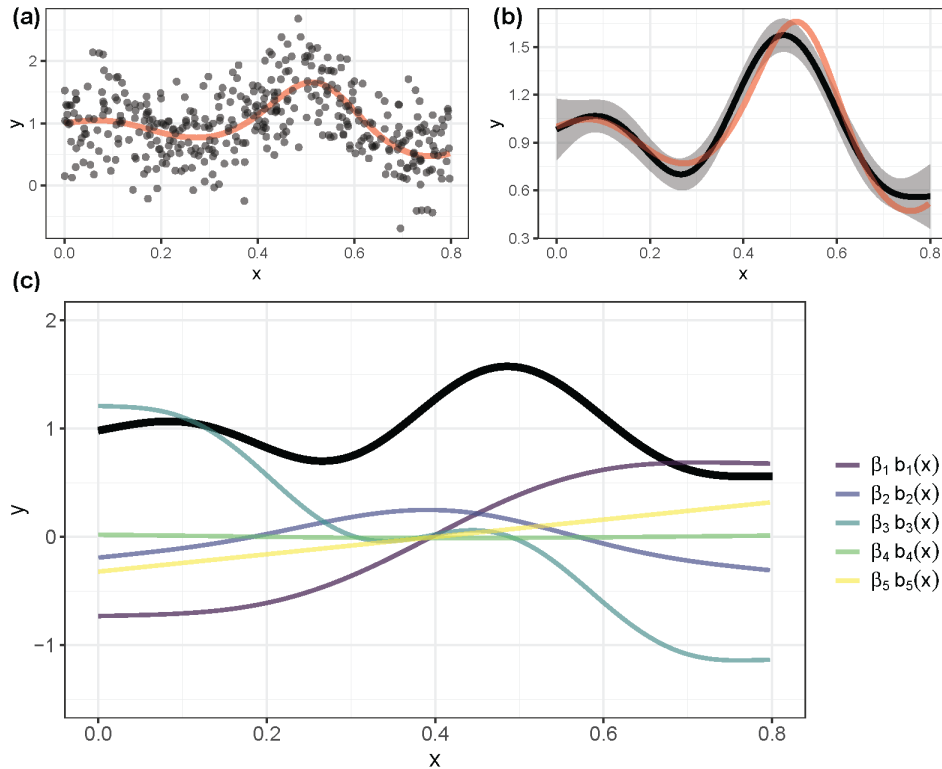


Figure A-1. (a) Data simulated from the nonlinear model given by Equation 2 and depicted in red. (b) Fitted generalized additive model smooth in black with corresponding 95% pointwise confidence interval in gray. True underlying curve provided in red for reference. (c) The six thin plate regression splines basis functions that make up the rank 6 basis expansion of $\hat{f}(x)$ (black curve).

MCT Family-Wise Error Rate

The type I error rate (alpha level; α) is the probability of rejecting the null hypothesis, given that it is true. For a single test, this is sometimes referred to as the experiment-wise or nominal error rate. In the case of a multiple comparison test, we are performing many tests simultaneously and collectively refer to them as a family of tests. Performing many tests, however, increases our chances of making a type I error. Specifically, the probability of making a type I error when performing m independent tests equals:

$$1 - (1 - \alpha)^m,$$

which we denote as the family-wise error rate. Based on this result, it is clear that as the family-wise error rate is much larger (inflated) compared to the nominal error rate. For example, Gateway National Recreation Area has 9 years of monitoring data and therefore has 36 potential pairwise comparisons. Assuming a nominal error rate of 0.05, we would expect at least one type I error to be committed for the set of 36 tests approximately 84% of the time. To assure that the family-wise error rate is at most the prescribed alpha level, several methods have been developed to adjust (inflate) the p -values from each individual test.