

**Estuarine Nutrient Enrichment Monitoring:
Data Summary and Trend Analysis for
Colonial National Historical Park**

January 6, 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 Colonial National Historical Park

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 Colonial National Historical Park. Neptune and Company, Inc., Lakewood, Colorado.

Contents

	Page
Figures.....	5
Tables.....	9
Acronyms and Abbreviations.....	13
1. Background Information.....	14
2. Spatial-Station Water Quality Data.....	16
2.1 Methods.....	16
2.1.1 Area-Weighted Condition.....	16
2.1.2 General Linear Mixed Model (GLMM).....	17
2.1.3 Generalized Additive Model (GAM).....	17
2.1.4 Multivariate Mann-Kendall Trend Analysis.....	19
2.2 Tidal Creek Stratum Results.....	19
2.2.1 Area-Weighted Condition Results.....	19
2.2.2 GAM Results.....	21
2.2.3 Multivariate Mann-Kendall Trend Test Results.....	25
2.3. Thoroughfare Stratum Results.....	26
2.3.1 Area-Weighted Condition Results.....	26
2.3.2 GAM Results.....	28
2.3.3 Multivariate Mann-Kendall Trend Test Results.....	32
2.4 All Strata Results.....	33
2.4.1 Area-Weighted Condition Results.....	33
2.4.2 GLMM Results.....	35
2.4.3 GAM Results.....	48
2.4.4 Multivariate Mann-Kendall Trend Test Results.....	58
3. Trend-Station Water Quality Data.....	60
3.1 Methods.....	60
3.1.1 Generalized Additive Model (GAM).....	60

3.1.2 Multivariate Mann-Kendall Trend Analysis	60
3.2 Summary Statistics	60
3.3 GAM Results	66
3.3.1 Chlorophyll-a (near surface).....	66
3.3.2 Dissolved Oxygen (near bottom).....	66
3.3.3 PAR Attenuation K_d (SAV standard).....	67
3.3.4 Salinity (near surface).....	67
3.3.5 Temperature (near surface).....	67
3.3.6 Turbidity (near surface).....	67
3.4 Multivariate Mann-Kendall Trend Test Results.....	74
4. Logging-Station Continuous Water Quality Data.....	75
4.1 Continuous Record Plots	75
4.2 Methods	75
4.2.1 Generalized Additive Model (GAM)	75
4.2.2 Additional Data Summaries	76
4.3 Thoroughfare Stratum Results.....	76
4.3.1 Chlorophyll-a.....	76
4.3.2 Dissolved Oxygen	76
4.3.3 PAR Attenuation K_d	76
4.3.4 pH	76
4.3.5 Salinity.....	76
4.3.6 Temperature.....	76
4.3.7 Turbidity	77
5. Power Simulations	88
5.1 Methods	88
5.2 Simulation Results.....	90
5.2.1 Tidal Creek Stratum	90
5.2.2 Thoroughfare Stratum	96

5.2.3 All Strata.....	102
5.3 Discussion.....	107
5.3.1 Key Question #1	107
Literature Cited	110
Appendix A.....	112

Figures

	Page
Figure 1-1. Locations of Northeast Coastal and Barrier Inventory & Monitoring (I&M) Network parks monitored for estuarine nutrient enrichment.	14
Figure 2-1. Trends in Colonial National Historical Park Tidal Creek water quality measures based on estimated generalized additive model smooths.	24
Figure 2-1 (continued). Trends in Colonial National Historical Park Tidal Creek water quality measures based on estimated generalized additive model smooths.	25
Figure 2-2. Trends in Colonial National Historical Park Thoroughfare water quality measures based on estimated generalized additive model smooths.	31
Figure 2-2 (continued). Trends in Colonial National Historical Park Thoroughfare water quality measures based on estimated generalized additive model smooths.	32
Figure 2-3. (Top): Boxplot of log surface chlorophyll- <i>a</i> concentration by year within Colonial National Historical Park (all strata). Estuarine quality category thresholds (Table 2-1) plotted as red dashed lines. (Bottom): Compact display of the multiple comparison test results.	37
Figure 2-4. (Top): Boxplot of log bottom dissolved oxygen concentration by year within Colonial National Historical Park (all strata). Estuarine quality category thresholds (Table 2-1) plotted as red dashed lines. (Bottom): Compact display of the multiple comparison test results.	39
Figure 2-5. (Top): Boxplot of log PAR attenuation by year within Colonial National Historical Park (all strata). Estuarine quality category thresholds (Table 2-1) plotted as red dashed lines. (Bottom): Compact display of the multiple comparison test results.	41
Figure 2-6. (Top): Boxplot of log surface salinity by year within Colonial National Historical Park (all strata). (Bottom): Compact display of the multiple comparison test results.	43
Figure 2-7. (Top): Boxplot of log surface temperature by year within Colonial National Historical Park (all strata). Estuarine quality category thresholds (Table 2-1) plotted as red dashed lines. (Bottom): Compact display of the multiple comparison test results.	45
Figure 2-8. (Top): Boxplot of log surface turbidity by year within Colonial National Historical Park (all strata). (Bottom): Compact display of the multiple comparison test results.	47
Figure 2-9. Trends in Colonial National Historical Park (all strata) water quality measures based on estimated generalized additive model smooths.	51

Figure 2-9 (continued). Trends in Colonial National Historical Park (all strata) water quality measures based on estimated generalized additive model smooths.....	52
Figure 2-10. Mean surface for the generalized additive model latitude/longitude smooth of surface chlorophyll- <i>a</i> concentration within Colonial National Historical Park (top = all strata; bottom = Thoroughfare stratum inset) during 2008-2020.....	53
Figure 2-11. Mean surface for the generalized additive model latitude/longitude smooth of bottom dissolved oxygen concentration within Colonial National Historical Park (top = all strata; bottom = Thoroughfare stratum inset) during 2008-2020.....	54
Figure 2-12. Mean surface for the generalized additive model latitude/longitude smooth of PAR attenuation within Colonial National Historical Park (top = all strata; bottom = Thoroughfare stratum inset) during 2008-2020.	55
Figure 2-13. Mean surface for the generalized additive model latitude/longitude smooth of surface salinity within Colonial National Historical Park (top = all strata; bottom = Thoroughfare stratum inset) during 2008-2020.	56
Figure 2-14. Mean surface for the generalized additive model latitude/longitude smooth of surface temperature within Colonial National Historical Park (top = all strata; bottom = Thoroughfare stratum inset) during 2008-2020.	57
Figure 2-15. Mean surface for the generalized additive model latitude/longitude smooth of surface turbidity within Colonial National Historical Park (top = all strata; bottom = Thoroughfare stratum inset) during 2008-2020.	58
Figure 3-1. Mean surface chlorophyll- <i>a</i> concentration ($\mu\text{g/L}$) trend at each Colonial National Historical Park trend station based on estimated generalized additive model smooths	68
Figure 3-2. Mean bottom dissolved oxygen concentration (mg/L) trend at each Colonial National Historical Park trend station based on estimated generalized additive model smooths	69
Figure 3-3. Mean PAR attenuation (K_d (m^{-1})) Mean PAR attenuation trend at each Colonial National Historical Park trend station based on estimated generalized additive model smooths.	70
Figure 3-4. Mean surface salinity (psu) trend at each Colonial National Historical Park trend station based on estimated generalized additive model smooths.....	71
Figure 3-5. Mean surface temperature ($^{\circ}\text{C}$) trend at each Colonial National Historical Park trend station based on estimated generalized additive model smooths	72
Figure 3-6. Mean surface turbidity (NTU) trend at each Colonial National Historical Park trend station based on estimated generalized additive model smooths	73

Figure 4-1. Continuous dissolved oxygen readings from sonde deployed at the Colonial National Historical Park logging station in 2008.....	75
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 Colonial National Historical Park Thoroughfare logging station based on generalized additive model smooths.	79
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 Colonial National Historical Park Thoroughfare logging station.....	80
Figure 4-4. Intra-annual trends in surface salinity (top), surface temperature (middle), and surface turbidity (bottom) at Colonial National Historical Park Thoroughfare logging station based on generalized additive model smooths.	81
Figure 4-5. Distribution of index period time series values for surface salinity (top), surface temperature (middle), and surface turbidity (bottom) at the Colonial National Historical Park Thoroughfare logging station.....	82
Figure 4-6. Intra-annual trends in pH (top) and depth (bottom) at the Colonial National Historical Park Thoroughfare logging station based on generalized additive model smooths.	83
Figure 4-7. Distribution of index period time series values for pH (top) and depth (bottom) at the Colonial National Historical Park Thoroughfare logging station.	84
Figure 4-8. Total hours per day that surface chlorophyll- <i>a</i> ($\mu\text{g/L}$) concentration was observed within each condition category at the Colonial National Historical Park Thoroughfare logging station.....	85
Figure 4-9. Total hours per day that bottom dissolved oxygen concentration (mg/L) was observed within each condition category at the Colonial National Historical Park Thoroughfare logging station.....	86
Figure 4-10. Total hours per day that PAR attenuation ($K_d \text{ (m}^{-1}\text{)}$) was observed within each condition category at the Colonial National Historical Park Thoroughfare logging station.....	87
Figure 5-1. Process model realizations (black dots) representing samples from 30 sites within a hypothetical estuary.	89
Figure 5-2. Estimated power to detect a linear trend in mean surface chlorophyll- <i>a</i> concentration within the Tidal Creek stratum of Colonial National Historical Park.	90
Figure 5-3. Estimated power to detect a linear trend in mean bottom dissolved oxygen concentration within the Tidal Creek stratum of Colonial National Historical Park.	92

Figure 5-4. Estimated power to detect a linear trend in mean PAR attenuation within the Tidal Creek stratum of Colonial National Historical Park.....	94
Figure 5-5. Estimated power to detect a linear trend in mean surface chlorophyll-a concentration within the Thoroughfare stratum of Colonial National Historical Park.....	96
Figure 5-6. Estimated power to detect a linear trend in mean bottom dissolved oxygen concentration within the Thoroughfare stratum of Colonial National Historical Park.....	98
Figure 5-7. Estimated power to detect a linear trend in mean PAR attenuation within the Thoroughfare stratum of Colonial National Historical Park.....	100
Figure 5-8. Estimated power to detect a linear trend in mean surface chlorophyll-a concentration within all strata of Colonial National Historical Park.	102
Figure 5-9. Estimated power to detect a linear trend in mean bottom dissolved oxygen concentration within all strata of Colonial National Historical Park.	104
Figure 5-10. Estimated power to detect a linear trend in mean PAR attenuation within all strata of Colonial National Historical Park.....	106
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).	113

Tables

	Page
Table 1-1. Sampling years for water quality monitoring within Colonial National Historical Park.	15
Table 2-1. Threshold values associated with each measure quality category.	16
Table 2-2. Area-weighted average water quality condition at Tidal Creek stratum of Colonial National Historical Park during the index periods of 2008-2020.	20
Table 2-3. Proportion of monitored estuarine resources falling into each quality category at Tidal Creek stratum of Colonial National Historical Park during the index periods of 2008-2020.	21
Table 2-4. Sampling start and end dates for each index period within Colonial National Historical Park Tidal Creek stratum.....	21
Table 2-5. Approximate significance of the generalized additive model smooth terms for each Colonial National Historical Park Tidal Creek model.....	23
Table 2-6. Proportion of null deviance explained by the explanatory variables for each Colonial National Historical Park Tidal Creek model.	23
Table 2-7. Summary of multivariate Mann-Kendall test for monotonic trend for each Colonial National Historical Park Tidal Creek measure.....	25
Table 2-8. Area-weighted average water quality condition at Thoroughfare stratum of Colonial National Historical Park during the index periods of 2008-2020.	27
Table 2-9. Proportion of monitored estuarine resources falling into each quality category at Thoroughfare stratum of Colonial National Historical Park during the index periods of 2008-2020.	28
Table 2-10. Sampling start and end dates for each index period within Colonial National Historical Park Thoroughfare stratum.....	28
Table 2-11. Approximate significance of the generalized additive model smooth terms for each Colonial National Historical Park Thoroughfare model	30
Table 2-12. Proportion of null deviance explained by the explanatory variables for each Colonial National Historical Park Thoroughfare model.	30
Table 2-13. Summary of multivariate Mann-Kendall test for monotonic trend for each Colonial National Historical Park Thoroughfare measure.....	32
Table 2-14. Area-weighted average water quality condition at Colonial National Historical Park (allstrata) during the index periods of 2008-2020.....	34

Table 2-15. Proportion of monitored estuarine resources falling into each quality category at Colonial National Historical Park (all strata) during the index periods of 2008-2020.	35
Table 2-16. Analysis of variance for the general linear mixed model fit to log-transformed water quality measurements collected within Colonial National Historical Park (all strata).....	36
Table 2-17. Summary of the multiple comparison test results for all possible yearly contrasts of mean log surface chlorophyll- <i>a</i> concentration within Colonial National Historical Park (all strata)	38
Table 2-18. Summary of the multiple comparison test results for all possible yearly contrasts of mean log bottom dissolved oxygen concentration within Colonial National Historical Park (all strata)	40
Table 2-19. Summary of the multiple comparison test results for all possible yearly contrasts of mean log PAR attenuation within Colonial National Historical Park (all strata).....	42
Table 2-20. Summary of the multiple comparison test results for all possible yearly contrasts of mean log surface salinity within Colonial National Historical Park (all strata).....	44
Table 2-21. Summary of the multiple comparison test results for all possible yearly contrasts of mean log surface temperature within Colonial National Historical Park (all strata).....	46
Table 2-22. Summary of the multiple comparison test results for all possible yearly contrasts of mean log surface turbidity within Colonial National Historical Park (all strata).....	48
Table 2-23. Approximate significance of the generalized additive model smooth terms for each Colonial National Historical Park (all strata) model.....	50
Table 2-24. Proportion of null deviance explained by the explanatory variables for each Colonial National Historical Park (all strata) model.....	50
Table 2-25. Summary of multivariate Mann-Kendall test for monotonic trend for each Colonial National Historical Park (all strata) measure.	59
Table 3-1. Yearly summary statistics for surface chlorophyll- <i>a</i> concentration ($\mu\text{g/L}$) at Colonial National Historical Park trend stations.....	61
Table 3-2. Yearly summary statistics for bottom dissolved oxygen concentration (mg/L) at Colonial National Historical Park trend stations.....	62
Table 3-3. Yearly summary statistics for PAR attenuation (K_d (m^{-1})) at Colonial National Historical Park trend stations	63

Table 3-4. Yearly summary statistics for surface salinity (psu) at Colonial National Historical Park trend stations	64
Table 3-5. Yearly summary statistics for surface temperature (°C) at Colonial National Historical Park trend stations	65
Table 3-6. Yearly summary statistics for surface turbidity (NTU) at Colonial National Historical Park trend stations	66
Table 3-7. Approximate significance of the generalized additive model smooth terms for each Colonial National Historical Park trend station model.....	67
Table 3-8. Summary of multivariate Mann-Kendall test for monotonic trend in each measure for Colonial National Historical Park trend stations.....	74
Table 4-1. Yearly summary statistics for measures collected at the Colonial National Historical Park Thoroughfare logging station.....	78
Table 5-1. Process model dispersion (ϕ) used to simulate trends in each water quality measure at Colonial National Historical Park.....	89
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 the Tidal Creek stratum of Colonial National Historical Park.....	91
Table 5-3. Estimated power to detect a linear trend in mean bottom dissolved oxygen concentration large enough to cross both quality thresholds within the Tidal Creek stratum of Colonial National Historical Park.....	93
Table 5-4. Estimated power to detect a linear trend in mean PAR attenuation large enough to cross both quality thresholds within the Tidal Creek stratum of Colonial National Historical Park.....	95
Table 5-5. Estimated power to detect a linear trend in mean surface chlorophyll- <i>a</i> concentration large enough to cross both quality thresholds within the Thoroughfare stratum of Colonial National Historical Park.....	97
Table 5-6. Estimated power to detect a linear trend in mean bottom dissolved oxygen concentration large enough to cross both quality thresholds within the Thoroughfare stratum of Colonial National Historical Park.....	99
Table 5-7. Estimated power to detect a linear trend in mean PAR attenuation large enough to cross both quality thresholds within the Thoroughfare stratum of Colonial National Historical Park.....	101
Table 5-8. Estimated power to detect a linear trend in mean surface chlorophyll- <i>a</i> concentration large enough to cross both quality thresholds within all stratum of Colonial National Historical Park.....	103

Table 5-9. Estimated power to detect a linear trend in mean bottom dissolved oxygen concentration large enough to cross both quality thresholds within all strata of Colonial National Historical Park.....	105
Table 5-10. Estimated power to detect a linear trend in mean PAR attenuation large enough to cross both quality thresholds within all strata of Colonial National Historical Park.	107

Acronyms and Abbreviations

COLO	Colonial National Historical Park
ENE	Estuarine nutrient enrichment
GAM	Generalized additive model
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 Colonial National Historical Park to date (Table 1-1).

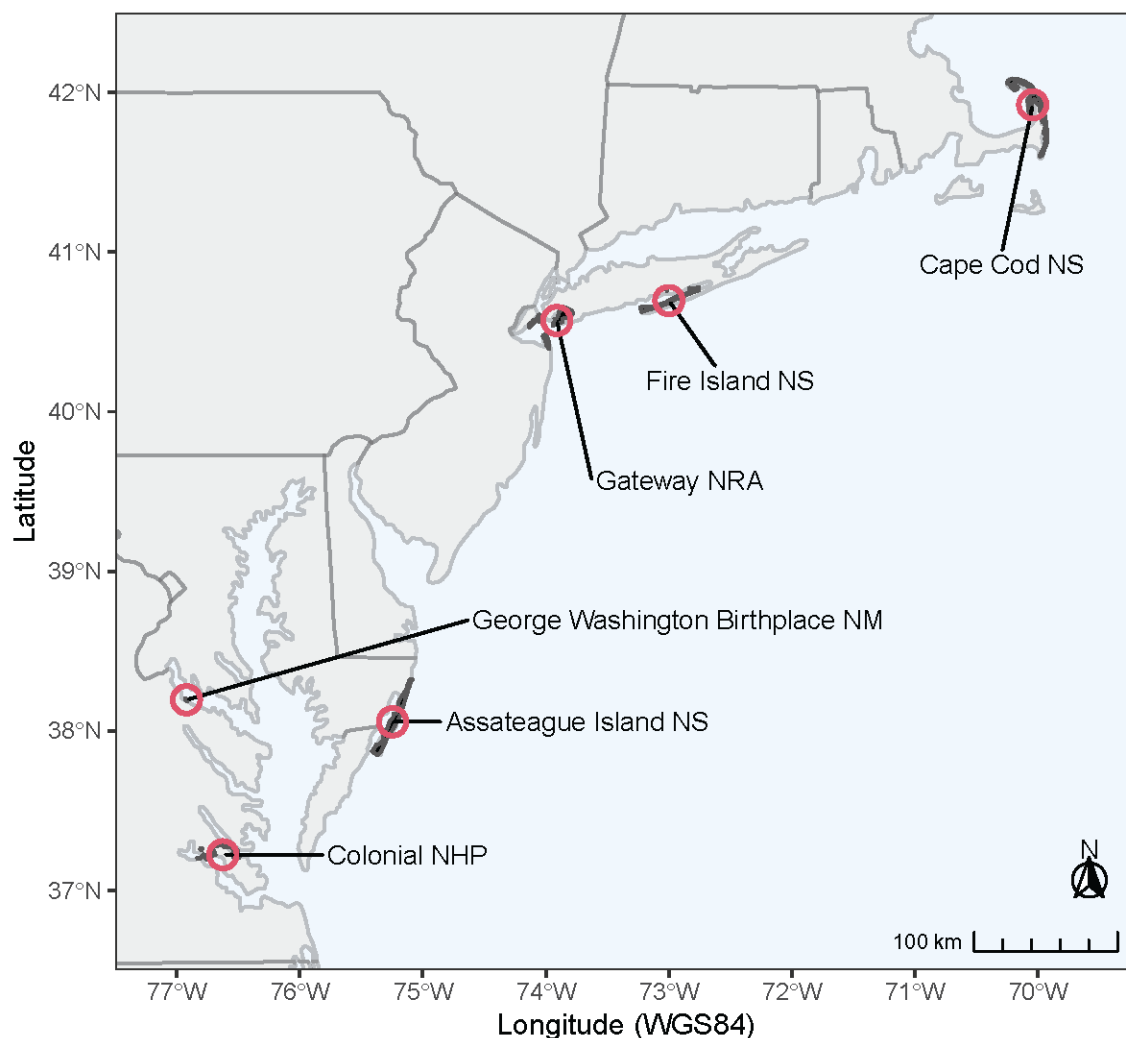


Figure 1-1. Locations of Northeast Coastal and Barrier Inventory & Monitoring (I&M) Network parks monitored for estuarine nutrient enrichment.

Table 1-1. Sampling years for water quality monitoring within Colonial National Historical Park.

Year	Water Quality
2003	sampled
2004	not sampled
2005	not sampled
2006	not sampled
2007	not sampled
2008	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

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 µg/L	5-20 µg/L	> 20 µg/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)	< 28 °C	28-33 °C	> 33 °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 2012 and 2014 (Figure 2-3). In this matrix, the values for $\bar{Z}_t$ are provided on the diagonal and the differences ($\bar{Z}_i - \bar{Z}_j, i \neq j$) 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 COLO during 2012 is $\exp(0.185) \approx 1.20$ times the median during 2014 (95% CI equal to $(1.04, 1.39)^2$). Stated another way: with 95% confidence, it is estimated that the median PAR attenuation in COLO during 2012 is between 1.04 and 1.39 times the median in 2014.

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

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

² The confidence intervals given in Table 2-19 are on the log scale. Therefore, we need to take the exponential of the upper and lower bound to put it on the response scale.

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 response (as opposed to the geometric mean³). 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 specification⁴) 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).

³ 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 $(2 \cdot 5 \cdot 7)^{1/3} = 70^{1/3} \approx 4.12$.

⁴ 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 monotonically⁵ increasing or decreasing trend in values for all sites within the estuary or stratum 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 2-25, we'd conclude at the 5% significance level that surface salinity measurements have been decreasing over time ($Z_{MK} = -2.35$; p -value = 0.019).

2.2 Tidal Creek Stratum Results

2.2.1 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:

Chlorophyll-a (near surface)

Area-weighted average near surface chlorophyll-*a* ranged from 18.09 to 30.90 $\mu\text{g/L}$. Poor area-weighted average condition ($> 20 \mu\text{g/L}$) was observed from 2008 to 2012 and also in 2020. The proportion of the Tidal Creek stratum categorized as fair condition ranged from 11.5 to 78.5%. More than 80% of the Tidal Creek stratum was in poor condition during 2010 and 2012.

Dissolved Oxygen (near bottom)

Area-weighted average near bottom dissolved oxygen concentration was good (i.e., $> 5 \text{ mg/L}$) for every index period and ranged from 6.08 to 7.38 mg/L . The majority of area ($> 90\%$) within the Tidal Creek stratum was characterized as in good condition. For each index period, no area in the Tidal Creek stratum was characterized as having poor dissolved oxygen condition (i.e., $< 2 \text{ mg/L}$).

PAR Attenuation K_d (SAV standard)

Area-weighted average PAR attenuation was poor for every index period (i.e., $K_d > 1.61 \text{ m}^{-1}$) and ranged from 2.68 to 4.16 m^{-1} (Table 2-2). The entirety of the Tidal Creek stratum was characterized as being in poor quality condition (Table 2-3).

pH

Area-weighted average pH ranged from 7.54 to 7.75. Measurements of pH began in 2018.

Salinity

Area-weighted average salinity ranged from 0.87 (surface depth in 2018) to 9.15 psu (mid-depth in 2010).

⁵ A monotonically increasing function $f(x)$ is a function that is never decreasing for all values of x (e.g., $f(x) = 2x$, $f(x) = \log(x)$, $f(x) = e^x$, etc.), and vice versa for monotonically decreasing functions. Note that these functions can be either linear or non-linear.

Temperature (near surface)

Area-weighted average near surface water temperature ranged from 27.69 to 30.29 °C and exceeded 28 °C during five index periods. Excluding 2008 and 2014 (which were the coolest), temperatures in the 28-33 °C range were observed in at least 68% of the Tidal Creek stratum (highest years were 2012 and 2016 with 99% and 100%, respectively).

Turbidity

Area-weighted average turbidity ranged from 14.01 (surface in 2018) to 50 NTU (bottom depth in 2012).

Table 2-2. Area-weighted average water quality condition at Tidal Creek stratum of Colonial National Historical Park during the index periods of 2008-2020.

Measure	Depth	Sampling Year						
		2008	2010	2012	2014	2016	2018	2020
Chlorophyll <i>a</i> (µg/L)	Surface	22.92	28.17	30.90	19.59	19.50	18.09	24.73
Dissolved oxygen (% saturation)	Surface	87.50	90.20	97.22	96.41	92.00	99.20	103.67
Dissolved oxygen (% saturation)	Mid-Depth	85.40	85.43	95.00	96.99	89.62	97.42	94.31
Dissolved oxygen (% saturation)	Bottom	82.56	83.19	88.90	94.99	89.45	93.37	94.60
Dissolved oxygen (mg/L)	Surface	6.56	6.58	7.18	7.47	6.86	7.65	7.74
Dissolved oxygen (mg/L)	Mid-Depth	6.39	6.25	7.03	7.50	6.68	7.54	7.12
Dissolved oxygen (mg/L)	Bottom	6.21	6.08	6.64	7.38	6.67	7.26	7.13
PAR attenuation (K_d (m ⁻¹))	Water Column	3.64	3.52	4.16	3.65	3.02	2.68	3.72
pH	Surface	NA	NA	NA	NA	NA	7.66	7.75
pH	Mid-Depth	NA	NA	NA	NA	NA	7.62	7.60
pH	Bottom	NA	NA	NA	NA	NA	7.54	7.59
Salinity (psu)	Surface	8.46	8.76	3.48	1.99	3.79	0.87	1.35
Salinity (psu)	Mid-Depth	8.04	9.15	3.62	2.19	3.76	0.99	1.98
Salinity (psu)	Bottom	8.33	8.91	3.68	2.05	3.98	0.92	1.43
Temperature (°C)	Surface	27.69	29.28	30.17	27.92	29.62	28.46	30.29
Temperature (°C)	Mid-Depth	27.94	28.96	29.99	27.98	29.60	28.27	29.41
Temperature (°C)	Bottom	27.55	29.06	29.52	27.73	29.49	27.99	29.67
Turbidity (NTU)	Surface	27.67	27.69	30.07	29.04	20.42	14.01	22.69
Turbidity (NTU)	Mid-Depth	31.55	31.16	34.83	31.29	26.85	16.73	24.07
Turbidity (NTU)	Bottom	42.55	43.65	50.00	34.42	33.33	20.74	33.88

Table 2-3. Proportion of monitored estuarine resources falling into each quality category at Tidal Creek stratum of Colonial National Historical Park during the index periods of 2008-2020.

Measure	Condition	Sampling Year						
		2008	2010	2012	2014	2016	2018	2020
Bottom dissolved oxygen (mg/L)	Good	94.4	89.7	100.0	100.0	99.9	98.6	100.0
Bottom dissolved oxygen (mg/L)	Fair	5.6	10.3	0.0	0.0	0.1	1.4	0.0
Bottom dissolved oxygen (mg/L)	Poor	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Bottom dissolved oxygen (mg/L)	Missing	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Chlorophyll <i>a</i> (µg/L)	Good	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Chlorophyll <i>a</i> (µg/L)	Fair	52.5	17.0	11.5	66.1	57.7	78.5	33.7
Chlorophyll <i>a</i> (µg/L)	Poor	47.5	81.1	88.5	33.9	42.3	21.5	66.3
Chlorophyll <i>a</i> (µg/L)	Missing	0.0	1.9	0.0	0.0	0.0	0.0	0.0
PAR attenuation (K_d (m ⁻¹))	Good	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PAR attenuation (K_d (m ⁻¹))	Fair	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PAR attenuation (K_d (m ⁻¹))	Poor	100.0	100.0	100.0	100.0	100.0	100.0	100.0
PAR attenuation (K_d (m ⁻¹))	Missing	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Surface Temperature (°C)	<28	54.3	26.8	1.1	52.7	0.0	31.3	3.3
Surface Temperature (°C)	28-33	45.7	73.2	98.9	47.3	100.0	68.7	96.7
Surface Temperature (°C)	>33	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Surface Temperature (°C)	Missing	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 Colonial National Historical Park Tidal Creek stratum.

Stratum	Year	Start	End
Tidal Creek	2008	2008-08-07	2008-08-29
Tidal Creek	2010	2010-08-03	2010-08-25
Tidal Creek	2012	2012-07-05	2012-07-25
Tidal Creek	2014	2014-07-08	2014-07-29
Tidal Creek	2016	2016-08-08	2016-08-29
Tidal Creek	2018	2018-07-11	2018-07-31
Tidal Creek	2020	2020-08-05	2020-08-26

2.2.2 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 except surface chlorophyll-*a*, PAR attenuation, and surface turbidity, which each lacked a significant spatial trend (Table 2-5). Notable observations from the fitted smooths for each measure are as follows:

Chlorophyll-*a* (near surface)

Chlorophyll-*a* exhibited a non-linear temporal trend from 2008 to 2014 when there were two periods of statistically significant changes (Figure 2-1). Mean concentration has remained relatively constant in the Tidal Creek stratum since 2014, but within-year variability has been increasing. Based on the fitted model, both time and location account for only 26.1% of the variability in chlorophyll-*a* concentrations within the Tidal Creek stratum (Table 2-6).

Dissolved Oxygen (near bottom)

Bottom dissolved oxygen had a slightly increasing trend from 2008 to 2014, and has since remained relatively constant. Based on the fitted model, both time and location account for only 21.9% of the variability in bottom dissolved oxygen concentrations within the Tidal Creek stratum (Table 2-6).

PAR Attenuation K_d (SAV standard)

The temporal trend in PAR attenuation for 2014 to 2020 is slightly concave upward with two periods of statistically significant change (Figure 2-1). Based on the fitted model, both time and location account for only 23.1% of the variability in PAR attenuation within the Tidal Creek stratum (Table 2-6).

Salinity (near surface)

Surface salinity exhibited a highly non-linear temporal trend (edf = 5.82) with statistically decreasing trends during the period from 2010 to 2014 and 2016 to 2018 (Figure 2-1). There was also a statistically significant increase from 2014 to 2016. The fitted GAM had the best fit amongst the measures, with time and location accounting for 74.5% of the variation in salinity within the Tidal Creek stratum (Table 2-6).

Temperature (near surface)

Surface water temperature exhibited a highly non-linear temporal trend (edf = 5.74) with cyclic periods of statistically significant increases and decreases (Figure 2-1). Based on the fitted model, both time and location account for only 34.3% of the variability in water temperature within the Tidal Creek stratum (Table 2-6).

Turbidity (near surface)

The temporal trend in surface turbidity indicates a statistically significant decrease occurred from 2014 to 2018 (Figure 2-1). Time and location account for only 30.3% of the variation in turbidity within the Tidal Creek stratum (Table 2-6).

Table 2-5. Approximate significance of the generalized additive model smooth terms for each Colonial National Historical Park Tidal Creek 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(s)	Model Term	edf	F -value	p -value
Chlorophyll a ($\mu\text{g/L}$)	Surface	s(year)	5.19	9.76	< 0.001
Chlorophyll a ($\mu\text{g/L}$)	Surface	s(long, lat)	0.00	0.00	0.721
Dissolved oxygen (mg/L)	Bottom	s(year)	3.58	5.09	< 0.001
Dissolved oxygen (mg/L)	Bottom	s(long, lat)	1.78	0.29	0.001
PAR attenuation (K_d (m^{-1}))	Water Column	s(year)	3.77	12.11	< 0.001
PAR attenuation (K_d (m^{-1}))	Water Column	s(long, lat)	0.00	0.00	0.744
Salinity (psu)	Surface	s(year)	5.82	124.18	< 0.001
Salinity (psu)	Surface	s(long, lat)	2.93	1.77	< 0.001
Temperature ($^{\circ}\text{C}$)	Surface	s(year)	5.74	15.00	< 0.001
Temperature ($^{\circ}\text{C}$)	Surface	s(long, lat)	1.07	0.18	0.002
Turbidity (NTU)	Surface	s(year)	4.88	11.94	< 0.001
Turbidity (NTU)	Surface	s(long, lat)	1.69	0.05	0.239

Table 2-6. Proportion of null deviance explained by the explanatory variables for each Colonial National Historical Park Tidal Creek model.

Measure	Deviance Explained
Chlorophyll a ($\mu\text{g/L}$)	26.1%
Dissolved oxygen (mg/L)	21.9%
PAR attenuation (K_d (m^{-1}))	23.1%
Salinity (psu)	74.5%
Temperature ($^{\circ}\text{C}$)	34.3%
Turbidity (NTU)	30.3%

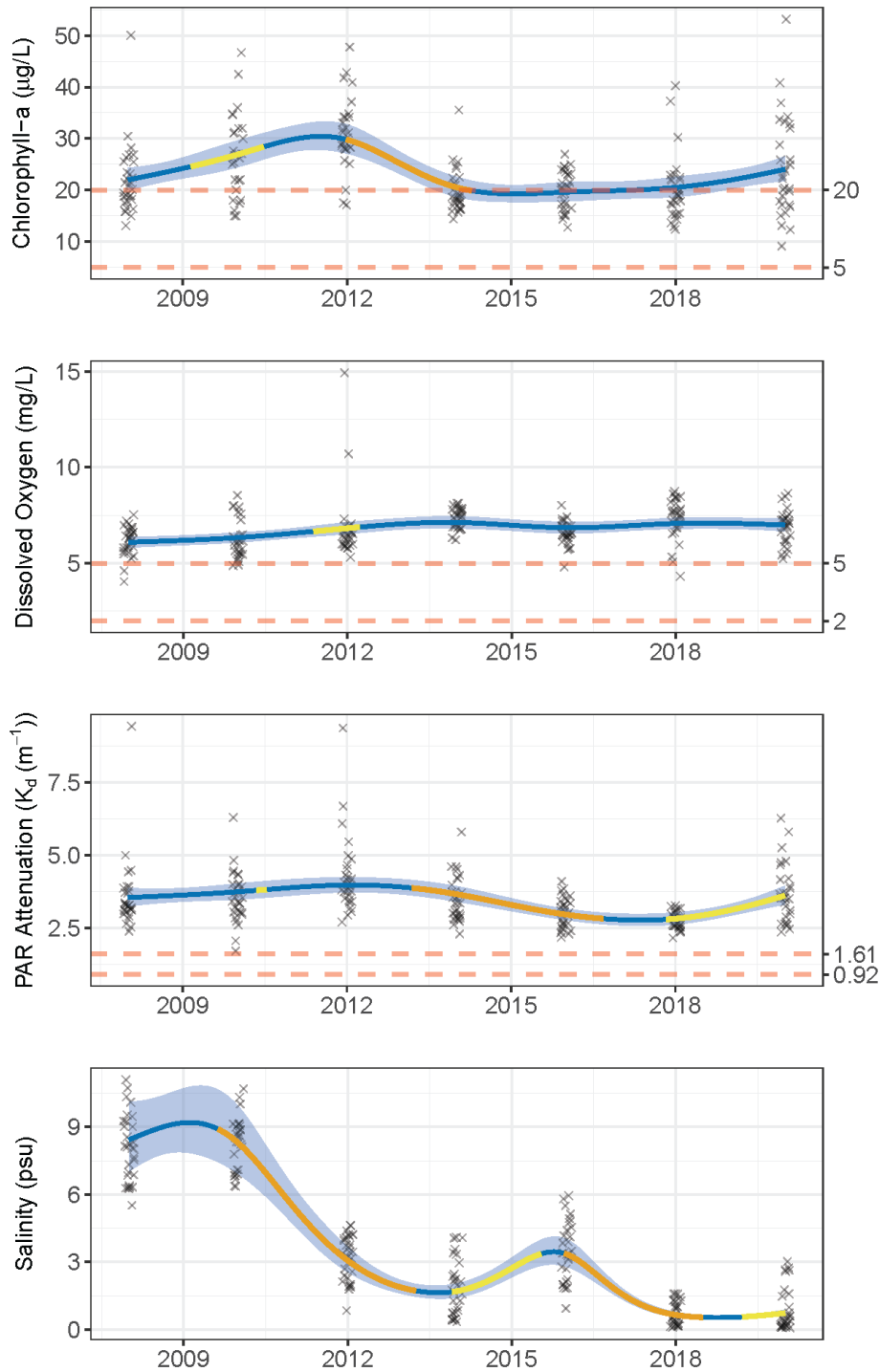


Figure 2-1. Trends in Colonial National Historical Park Tidal Creek water quality measures based on estimated generalized additive model smooths. The fitted smooth is plotted in dark blue with shaded 95% confidence interval. Periods of significant temporal change (increases/decreases), if present, are plotted in yellow/orange. Estuarine quality category thresholds (Table 2-1) plotted as red dashed lines. Note: in some instances, the y-axis is transformed to the $\log_{10}$ scale to better visualize the estimated trend.

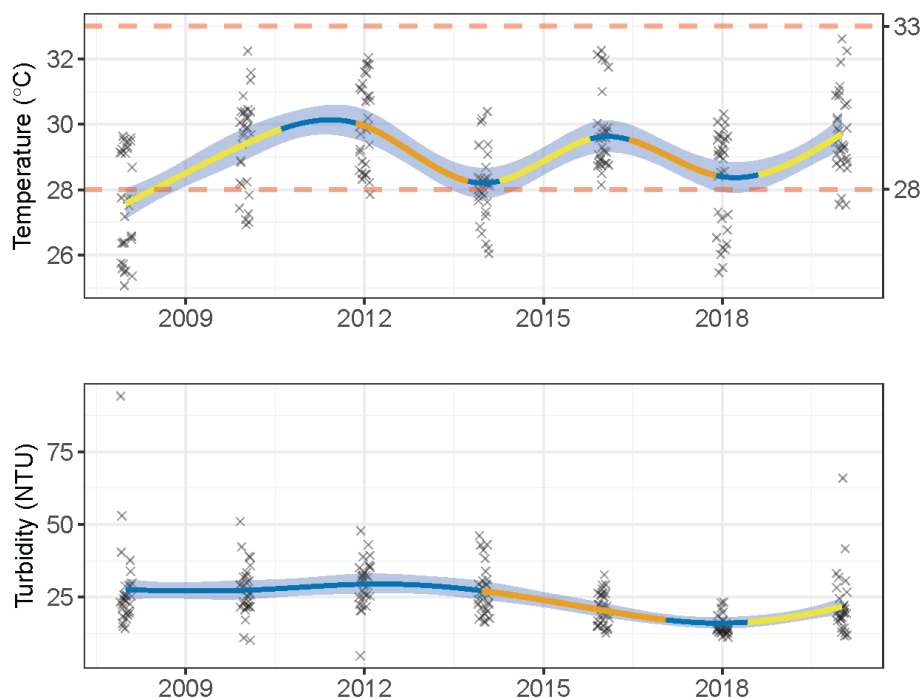


Figure 2-1 (continued). Trends in Colonial National Historical Park Tidal Creek water quality measures based on estimated generalized additive model smooths. The fitted smooth is plotted in dark blue with shaded 95% confidence interval. Periods of significant temporal change (increases/decreases), if present, are plotted in yellow/orange. Estuarine quality category thresholds (Table 2-1) plotted as red dashed lines. Note: in some instances, the y-axis is transformed to the $\log_{10}$ scale to better visualize the estimated trend.

2.2.3 Multivariate Mann-Kendall Trend Test Results

The multivariate Mann-Kendall test results indicate there is statistically significant increasing monotonic trend in bottom dissolved oxygen (p -value = 0.054) and decreasing monotonic trend in surface salinity (p -value = 0.018) within the Tidal Creek stratum during the index periods of 2008 to 2020 (Table 2-7).

Table 2-7. Summary of multivariate Mann-Kendall test for monotonic trend for each Colonial National Historical Park Tidal Creek 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	-0.35	0.727
Dissolved oxygen (mg/L)	Bottom	1.93	0.054
PAR attenuation (K_d (m^{-1}))	Water Column	-0.70	0.486
Salinity (psu)	Surface	-2.37	0.018
Temperature ($^{\circ}\text{C}$)	Surface	0.98	0.325
Turbidity (NTU)	Surface	-1.53	0.126

2.3. Thoroughfare Stratum Results

2.3.1 Area-Weighted Condition Results

Sampling start and end dates for each index period are provided in Table 2-10. Area-weighted condition summaries for each measure are as follows:

Chlorophyll-*a* (near surface)

Area-weighted average near surface chlorophyll-*a* ranged from 20.56 to 42.76 $\mu\text{g/L}$ and was poor (i.e., $> 20 \mu\text{g/L}$) for all of the index periods. The proportion of the Thoroughfare stratum categorized as fair condition ranged from 1.4 to 78.3%. More than 85% of the Thoroughfare stratum was in poor condition during 2008, 2010, 2012, and 2020. No proportion of the Thoroughfare stratum was observed in good condition.

Dissolved Oxygen (near bottom)

Area-weighted average near bottom dissolved oxygen concentration was good (i.e., $> 5 \text{ mg/L}$) for every index period and ranged from 5.63 to 7.12 mg/L . The majority of area ($> 85\%$) within the Thoroughfare stratum was characterized as in good condition. For each index period, no area within the Thoroughfare stratum was characterized as having poor dissolved oxygen condition (i.e., $< 2 \text{ mg/L}$).

PAR Attenuation K_d (SAV standard)

Area-weighted average PAR attenuation within the Thoroughfare stratum was poor for every index period (i.e., $K_d > 1.61 \text{ m}^{-1}$) and ranged from 2.93 to 4.46 m^{-1} (Table 2-8). The entirety of the Thoroughfare stratum was characterized as being of poor quality condition (Table 2-9).

pH

Area-weighted average pH ranged from 7.26 to 7.63. Measurements of pH began in 2018.

Salinity

Area-weighted average salinity ranged from 1.83 (mid-depth in 2018) to 11.07 psu (bottom depth in 2010).

Temperature (near surface)

Area-weighted average near surface water temperature within the Thoroughfare stratum ranged from 27.25 to 30.62 $^{\circ}\text{C}$, and exceeded 28 $^{\circ}\text{C}$ during five index periods. Excluding 2008 and 2014 (which were the coolest), temperatures in the 28-33 $^{\circ}\text{C}$ range were observed in nearly 80% of the Thoroughfare stratum (highest years were 2012 and 2016 with 100% and 92%, respectively).

Turbidity

Area-weighted average turbidity ranged from 18.83 (surface in 2020) to 53.27 NTU (bottom depth in 2012).

Table 2-8. Area-weighted average water quality condition at Thoroughfare stratum of Colonial National Historical Park during the index periods of 2008-2020.

Measure	Depth	Sampling Year						
		2008	2010	2012	2014	2016	2018	2020
Chlorophyll <i>a</i> ($\mu\text{g/L}$)	Surface	33.93	27.62	42.76	23.04	25.40	20.56	24.14
Dissolved oxygen (% saturation)	Surface	99.64	96.03	97.33	97.18	96.25	86.19	92.30
Dissolved oxygen (% saturation)	Mid-Depth	97.86	89.71	94.17	94.74	93.70	84.54	79.47
Dissolved oxygen (% saturation)	Bottom	94.26	77.13	86.10	92.81	90.37	74.24	77.08
Dissolved oxygen (mg/L)	Surface	7.46	6.86	7.10	7.41	6.96	6.65	6.85
Dissolved oxygen (mg/L)	Mid-Depth	7.34	6.46	6.91	7.26	6.78	6.55	5.94
Dissolved oxygen (mg/L)	Bottom	7.08	5.63	6.34	7.12	6.57	5.78	5.79
PAR attenuation (K_d (m^{-1}))	Water Column	3.35	2.93	4.46	3.26	2.95	3.26	3.43
pH	Surface	NA	NA	NA	NA	NA	7.44	7.63
pH	Mid-Depth	NA	NA	NA	NA	NA	7.35	7.43
pH	Bottom	NA	NA	NA	NA	NA	7.26	7.41
Salinity (psu)	Surface	10.24	10.79	6.30	5.12	6.12	2.27	2.86
Salinity (psu)	Mid-Depth	10.30	10.52	6.56	4.91	6.26	1.83	3.08
Salinity (psu)	Bottom	10.42	11.07	6.60	5.22	6.15	2.37	3.37
Temperature ($^{\circ}\text{C}$)	Surface	27.25	29.85	29.59	27.84	30.62	28.07	29.95
Temperature ($^{\circ}\text{C}$)	Mid-Depth	27.15	29.52	29.30	27.65	30.49	28.01	29.65
Temperature ($^{\circ}\text{C}$)	Bottom	27.05	28.87	29.04	27.54	30.18	27.59	29.26
Turbidity (NTU)	Surface	27.09	19.84	32.25	26.33	18.92	22.39	18.83
Turbidity (NTU)	Mid-Depth	30.41	24.64	36.58	30.15	21.86	28.59	21.19
Turbidity (NTU)	Bottom	34.00	38.49	53.27	39.31	27.01	25.98	30.79

Table 2-9. Proportion of monitored estuarine resources falling into each quality category at Thoroughfare stratum of Colonial National Historical Park during the index periods of 2008-2020.

Measure	Condition	Sampling Year						
		2008	2010	2012	2014	2016	2018	2020
Bottom dissolved oxygen (mg/L)	Good	87.5	90.6	91.7	100.0	100.0	85.2	92.9
Bottom dissolved oxygen (mg/L)	Fair	12.5	9.4	8.3	0.0	0.0	14.8	7.1
Bottom dissolved oxygen (mg/L)	Poor	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Bottom dissolved oxygen (mg/L)	Missing	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Chlorophyll <i>a</i> (µg/L)	Good	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Chlorophyll <i>a</i> (µg/L)	Fair	12.3	1.4	6.2	70.4	57.7	78.3	13.3
Chlorophyll <i>a</i> (µg/L)	Poor	87.7	98.6	93.8	29.6	42.3	21.7	86.7
Chlorophyll <i>a</i> (µg/L)	Missing	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PAR attenuation (K_d (m ⁻¹))	Good	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PAR attenuation (K_d (m ⁻¹))	Fair	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PAR attenuation (K_d (m ⁻¹))	Poor	100.0	100.0	100.0	100.0	100.0	100.0	100.0
PAR attenuation (K_d (m ⁻¹))	Missing	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Surface Temperature (°C)	<28	77.1	21.4	0.0	63.4	7.9	20.7	11.6
Surface Temperature (°C)	28-33	22.9	78.6	100.0	36.6	92.1	79.3	88.4
Surface Temperature (°C)	>33	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Surface Temperature (°C)	Missing	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table 2-10. Sampling start and end dates for each index period within Colonial National Historical Park Thoroughfare stratum.

Stratum	Year	Start	End
Thoroughfare	2008	2008-08-07	2008-08-29
Thoroughfare	2010	2010-08-03	2010-08-26
Thoroughfare	2012	2012-07-05	2012-07-24
Thoroughfare	2014	2014-07-07	2014-07-29
Thoroughfare	2016	2016-08-08	2016-08-29
Thoroughfare	2018	2018-07-11	2018-08-01
Thoroughfare	2020	2020-08-05	2020-08-27

2.3.2 GAM Results

The generalized additive model summaries indicate that there exists several statistically significant (assuming a 0.05 significance level) spatial and temporal trends in the water quality measures observed in the Thoroughfare stratum (Table 2-11). Notable observations from the fitted smooths for each measure are as follows:

Chlorophyll-*a* (near surface)

Chlorophyll-*a* exhibited a slightly concave down trend with the mean remaining relatively constant. There were no periods of statistically significant changes observed (Figure 2-2). Based on the fitted model, both time and location account for 57.3% of the variability in chlorophyll-*a* concentrations within the Thoroughfare stratum (Table 2-12).

Dissolved Oxygen (near bottom)

Bottom dissolved oxygen exhibited a slightly increasing trend from 2008 to 2014 and a slightly decreasing trend after 2014. There were no periods of statistically significant changes observed with the Thoroughfare stratum (Figure 2-2). The fitted model does very poorly explaining variation in dissolved oxygen concentrations within the stratum, with time and location accounting for only 12.2% of the variation (Table 2-12).

PAR Attenuation K_d (SAV standard)

PAR attenuation within the Thoroughfare stratum has been constant from 2008 to 2020 (Figure 2-2). Based on the fitted model, both time and location account for only 19% of the variability in PAR attenuation within the Thoroughfare stratum (Table 2-12).

Salinity (near surface)

Surface salinity exhibited a non-linear decreasing temporal trend with two periods of statistically significant decreases occurring from 2010 to 2012 and 2016 to 2018 (Figure 2-2). Within each year, there appears to be two distinct groups of salinity measurements. The fitted GAM had the best fit amongst the measures, with time and location accounting for 74.7% of the variation in salinity within the Thoroughfare stratum (Table 2-12).

Temperature (near surface)

Surface water temperature exhibited a non-linear temporal trend ($\text{edf} = 4.97$) with cyclic periods of increases and decreases (Figure 2-2 (continued)). Based on the fitted model, both time and location account for only 28.3% of the variability in water temperature within the Thoroughfare stratum (Table 2-12).

Turbidity (near surface)

Surface turbidity within the Thoroughfare stratum has been relatively constant from 2008 and 2020 with no trend (Figure 2-2 (continued)). Time and location account for only 15.1% of the variation in turbidity within the Thoroughfare stratum (Table 2-12).

Table 2-11. Approximate significance of the generalized additive model smooth terms for each Colonial National Historical Park Thoroughfare 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(s)	Model Term	edf	F -value	p -value
Chlorophyll a ($\mu\text{g/L}$)	Surface	s(year)	1.57	0.87	0.035
Chlorophyll a ($\mu\text{g/L}$)	Surface	s(long, lat)	9.40	1.76	< 0.001
Dissolved oxygen (mg/L)	Bottom	s(year)	1.69	0.93	0.036
Dissolved oxygen (mg/L)	Bottom	s(long, lat)	1.80	0.09	0.115
PAR attenuation (K_d (m^{-1}))	Water Column	s(year)	0.00	0.00	0.921
PAR attenuation (K_d (m^{-1}))	Water Column	s(long, lat)	4.89	0.35	0.018
Salinity (psu)	Surface	s(year)	3.86	23.63	< 0.001
Salinity (psu)	Surface	s(long, lat)	4.61	7.89	< 0.001
Temperature ($^{\circ}\text{C}$)	Surface	s(year)	4.97	4.87	< 0.001
Temperature ($^{\circ}\text{C}$)	Surface	s(long, lat)	0.00	0.00	0.635
Turbidity (NTU)	Surface	s(year)	0.51	0.12	0.235
Turbidity (NTU)	Surface	s(long, lat)	1.99	0.24	0.006

Table 2-12. Proportion of null deviance explained by the explanatory variables for each Colonial National Historical Park Thoroughfare model.

Measure	Deviance Explained
Chlorophyll a ($\mu\text{g/L}$)	57.3%
Dissolved oxygen (mg/L)	12.2%
PAR attenuation (K_d (m^{-1}))	19.0%
Salinity (psu)	74.7%
Temperature ($^{\circ}\text{C}$)	28.3%
Turbidity (NTU)	15.1%

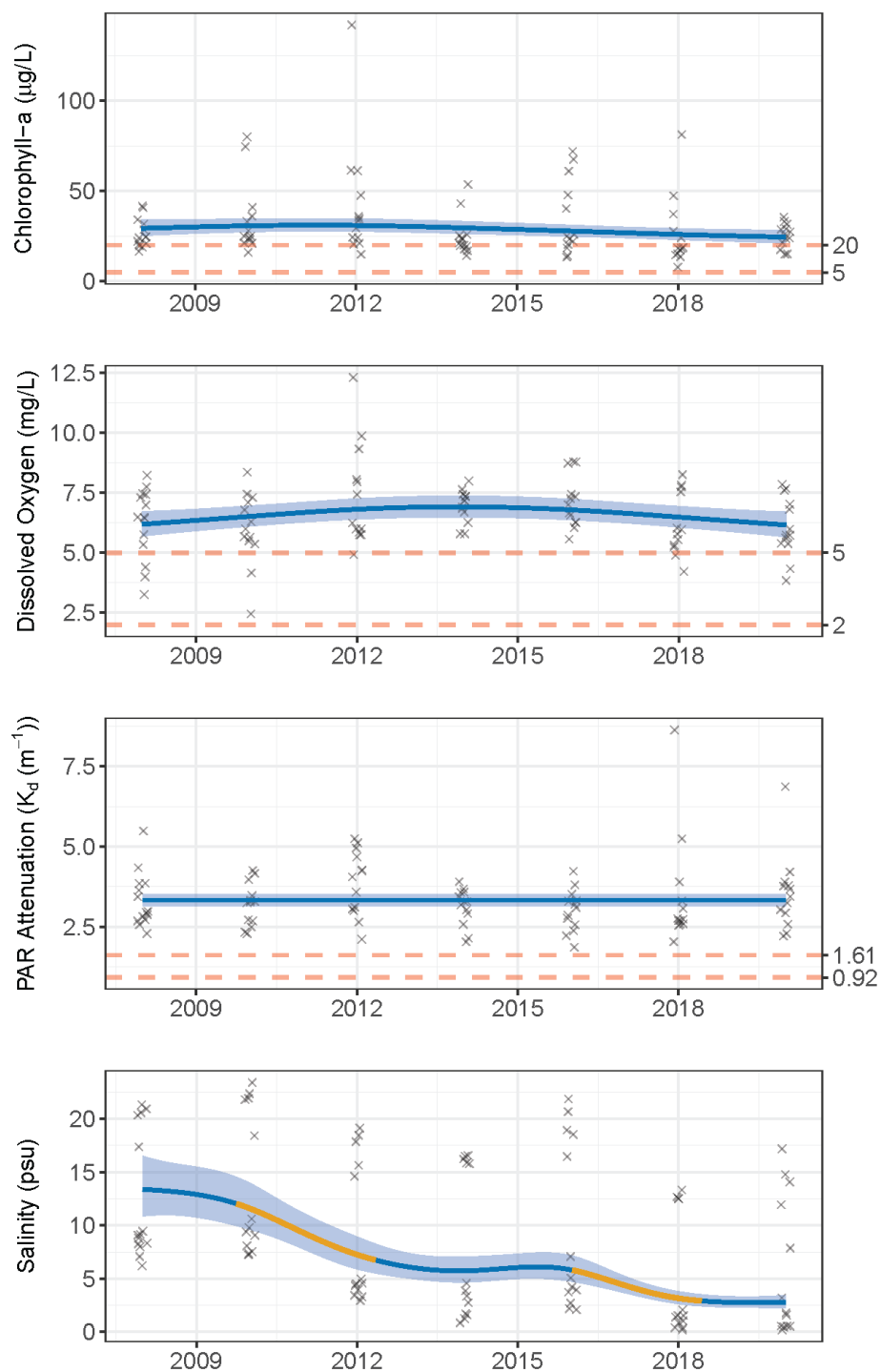


Figure 2-2. Trends in Colonial National Historical Park Thoroughfare water quality measures based on estimated generalized additive model smooths. The fitted smooth is plotted in dark blue with shaded 95% confidence interval. Periods of significant temporal change (increases/decreases), if present, are plotted in yellow/orange. Estuarine quality category thresholds (Table 2-1) plotted as red dashed lines. Note: in some instances, the y-axis is transformed to the $\log_{10}$ scale to better visualize the estimated trend.

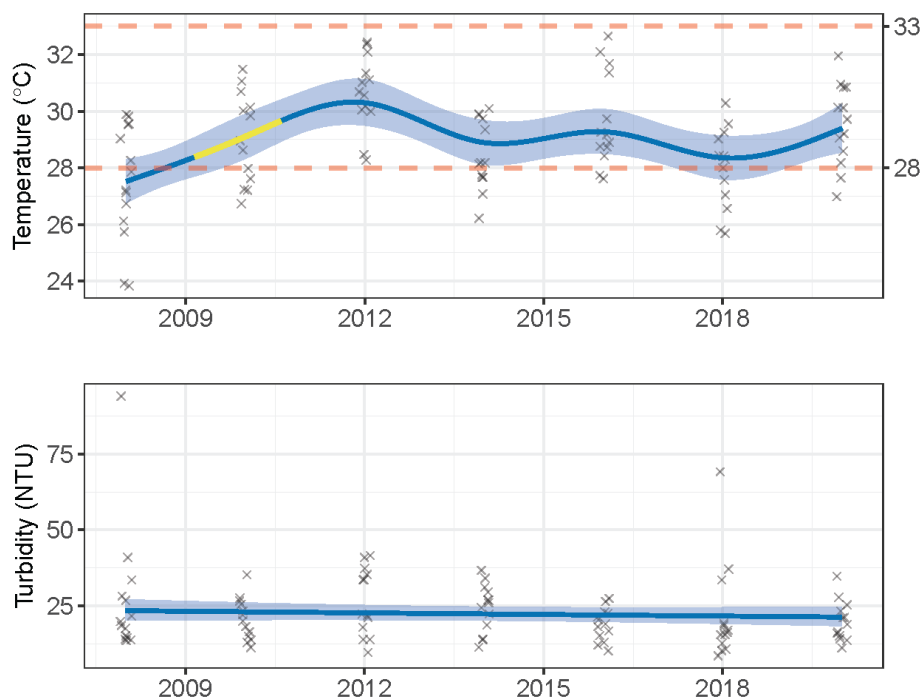


Figure 2-2 (continued). Trends in Colonial National Historical Park Thoroughfare water quality measures based on estimated generalized additive model smooths. The fitted smooth is plotted in dark blue with shaded 95% confidence interval. Periods of significant temporal change (increases/decreases), if present, are plotted in yellow/orange. Estuarine quality category thresholds (Table 2-1) plotted as red dashed lines. Note: in some instances, the y-axis is transformed to the $\log_{10}$ scale to better visualize the estimated trend.

2.3.3 Multivariate Mann-Kendall Trend Test Results

The multivariate Mann-Kendall test results indicate there is statistically significant decreasing monotonic trend in surface salinity (p -value = 0.031) within the Thoroughfare stratum during the index periods of 2008 to 2020 (Table 2-13).

Table 2-13. Summary of multivariate Mann-Kendall test for monotonic trend for each Colonial National Historical Park Thoroughfare 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	-0.26	0.797
Dissolved oxygen (mg/L)	Bottom	0.13	0.900
PAR attenuation (K_d (m^{-1}))	Water Column	0.11	0.910
Salinity (psu)	Surface	-2.16	0.031
Temperature ($^{\circ}\text{C}$)	Surface	0.82	0.411
Turbidity (NTU)	Surface	-0.53	0.599

2.4 All Strata Results

2.4.1 Area-Weighted Condition Results

Area-weighted condition summaries for each measure are as follows:

Chlorophyll-*a* (near surface)

Area-weighted average near surface chlorophyll-*a* ranged from 18.67 to 33.67 $\mu\text{g/L}$. Poor area-weighted average condition ($>20 \mu\text{g/L}$) was observed in all sampled years except 2018. The proportion of the estuary categorized as fair condition ranged from 13 to 79%. More than 70% of the estuary was in poor condition during 2010, 2012 and 2020.

Dissolved Oxygen (near bottom)

Area-weighted average near bottom dissolved oxygen concentration was good (i.e., $> 5 \text{ mg/L}$) for every index period and ranged from 5.98 to 7.32 mg/L . The majority of area ($>90\%$) within the estuary was characterized as in good condition. For each index period, no area was characterized as having poor dissolved oxygen condition (i.e., $< 2 \text{ mg/L}$).

PAR Attenuation K_d (SAV standard)

Area-weighted average PAR attenuation was poor for every index period (i.e. $K_d > 1.61 \text{ m}^{-1}$), and ranged from 2.82 to 4.23 m^{-1} (Table 2-14). The entirety of the estuary was characterized as being in poor condition (Table 2-15).

pH

Area-weighted average pH ranged from 7.48 to 7.72. Measurements of pH began in 2018.

Salinity

Area-weighted average salinity ranged from 1.20 (surface depth in 2018) to 9.53 psu (mid-depth in 2010).

Temperature (near surface)

Area-weighted average near surface water temperature ranged from 27.59 to 30.21 $^{\circ}\text{C}$, and exceeded 28 $^{\circ}\text{C}$ during five index periods. Excluding 2008 and 2014 (which were the coolest), temperatures in the 28-33 $^{\circ}\text{C}$ range were observed in at least 70% of the estuary (highest years were 2012 and 2016 with 99.2% and 98.2%, respectively).

Turbidity

Area-weighted average turbidity ranged from 15.96 (surface in 2018) to 50.76 NTU (bottom depth in 2014).

Table 2-14. Area-weighted average water quality condition at Colonial National Historical Park (allstrata) during the index periods of 2008-2020.

Measure	Depth	Sampling Year						
		2008	2010	2012	2014	2016	2018	2020
Chlorophyll <i>a</i> ($\mu\text{g/L}$)	Surface	25.51	28.04	33.67	20.39	20.87	18.67	24.59
Dissolved oxygen (% saturation)	Surface	90.36	91.56	97.24	96.59	92.99	96.17	101.02
Dissolved oxygen (% saturation)	Mid-Depth	88.62	86.61	94.79	96.41	90.69	94.20	89.86
Dissolved oxygen (% saturation)	Bottom	85.31	81.78	88.25	94.48	89.66	88.97	90.45
Dissolved oxygen (mg/L)	Surface	6.77	6.64	7.16	7.46	6.88	7.42	7.53
Dissolved oxygen (mg/L)	Mid-Depth	6.63	6.31	7.00	7.44	6.71	7.29	6.76
Dissolved oxygen (mg/L)	Bottom	6.42	5.98	6.57	7.32	6.65	6.92	6.81
PAR attenuation (K_d (m^{-1}))	Water Column	3.58	3.38	4.23	3.56	3.01	2.82	3.65
pH	Surface	NA	NA	NA	NA	NA	7.61	7.72
pH	Mid-Depth	NA	NA	NA	NA	NA	7.56	7.55
pH	Bottom	NA	NA	NA	NA	NA	7.48	7.55
Salinity (psu)	Surface	8.88	9.23	4.13	2.72	4.33	1.20	1.70
Salinity (psu)	Mid-Depth	8.62	9.53	4.34	2.89	4.42	1.20	2.31
Salinity (psu)	Bottom	8.82	9.42	4.36	2.79	4.49	1.25	1.89
Temperature ($^{\circ}\text{C}$)	Surface	27.59	29.41	30.04	27.90	29.85	28.37	30.21
Temperature ($^{\circ}\text{C}$)	Mid-Depth	27.74	29.12	29.82	27.90	29.83	28.20	29.48
Temperature ($^{\circ}\text{C}$)	Bottom	27.43	29.02	29.41	27.68	29.65	27.90	29.58
Turbidity (NTU)	Surface	27.53	25.86	30.58	28.41	20.07	15.96	21.79
Turbidity (NTU)	Mid-Depth	31.25	29.35	35.26	31.00	25.54	19.69	23.21
Turbidity (NTU)	Bottom	40.54	42.45	50.76	35.56	31.86	21.94	33.15

Table 2-15. Proportion of monitored estuarine resources falling into each quality category at Colonial National Historical Park (all strata) during the index periods of 2008-2020.

Measure	Condition	Sampling Year						
		2008	2010	2012	2014	2016	2018	2020
Bottom dissolved oxygen (mg/L)	Good	92.8	89.9	98.1	100.0	99.9	95.5	98.3
Bottom dissolved oxygen (mg/L)	Fair	7.2	10.1	1.9	0.0	0.1	4.5	1.7
Bottom dissolved oxygen (mg/L)	Poor	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Bottom dissolved oxygen (mg/L)	Missing	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Chlorophyll <i>a</i> (µg/L)	Good	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Chlorophyll <i>a</i> (µg/L)	Fair	43.0	13.3	10.3	67.1	57.7	78.5	29.0
Chlorophyll <i>a</i> (µg/L)	Poor	57.0	85.2	89.7	32.9	42.3	21.5	71.0
Chlorophyll <i>a</i> (µg/L)	Missing	0.0	1.5	0.0	0.0	0.0	0.0	0.0
PAR attenuation (K_d (m ⁻¹))	Good	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PAR attenuation (K_d (m ⁻¹))	Fair	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PAR attenuation (K_d (m ⁻¹))	Poor	100.0	100.0	100.0	100.0	100.0	100.0	100.0
PAR attenuation (K_d (m ⁻¹))	Missing	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Surface Temperature (°C)	<28	59.6	25.6	0.8	55.2	1.8	28.9	5.2
Surface Temperature (°C)	28-33	40.4	74.4	99.2	44.8	98.2	71.1	94.8
Surface Temperature (°C)	>33	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Surface Temperature (°C)	Missing	0.0	0.0	0.0	0.0	0.0	0.0	0.0

2.4.2 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 2-16). Notable observations from the multiple comparison tests (MCT) for each measure are as follows:

Chlorophyll-*a* (near surface)

The MCT results indicate that mean log near surface chlorophyll-*a* concentrations in 2010 and 2012 were statistically different from most other index periods (Figure 2-3; Table 2-17).

Dissolved Oxygen (near bottom)

The MCT results indicate that mean log bottom dissolved oxygen concentrations in 2008 and 2012 were statistically different from most other years (Figure 2-4; Table 2-17).

PAR Attenuation K_d (SAV standard)

The MCT results indicate that mean log PAR attenuation in 2012 and 2018 were statistically different compared to most other years (Figure 2-5; Table 2-19).

Salinity (near surface)

The MCT results indicate that mean log surface salinity for all but three pairs of years are statistically different (Figure 2-6; Table 2-20). The box plots indicate that within-year variability has been higher in the more recently sampled years.

Temperature (near surface)

Due to the cyclical nature of the surface water temperatures, the years with lower mean log surface water temperature tend to be statistically different from the years with higher temperatures (Figure 2-7; Table 2-21).

Turbidity (near surface)

The MCT results indicate that mean log surface turbidity in 2018 is statistically different compared to most other years (Figure 2-8; Table 2-22), and is estimated to be lower in each instance.

Additionally, mean log surface turbidity in 2012 is statistically different compared to years 2016 through 2020.

Table 2-16. Analysis of variance for the general linear mixed model fit to log-transformed water quality measurements collected within Colonial National Historical Park (all strata). 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. [D.F. = degrees of freedom]

Measure	Depths	D.F.	F-value	p-value	MCT Results
Chlorophyll <i>a</i> ($\mu\text{g/L}$)	Surface	(6, 251)	13.39	< 0.005	Figure 2-3
Dissolved oxygen (mg/L)	Bottom	(6, 250)	6.87	< 0.005	Figure 2-4
PAR attenuation (K_d (m^{-1}))	Water Column	(6, 252)	9.95	< 0.005	Figure 2-5
Salinity (psu)	Surface	(6, 252)	119.20	< 0.005	Figure 2-6
Temperature ($^{\circ}\text{C}$)	Surface	(6, 252)	24.78	< 0.005	Figure 2-7
Turbidity (NTU)	Surface	(6, 252)	11.71	< 0.005	Figure 2-8

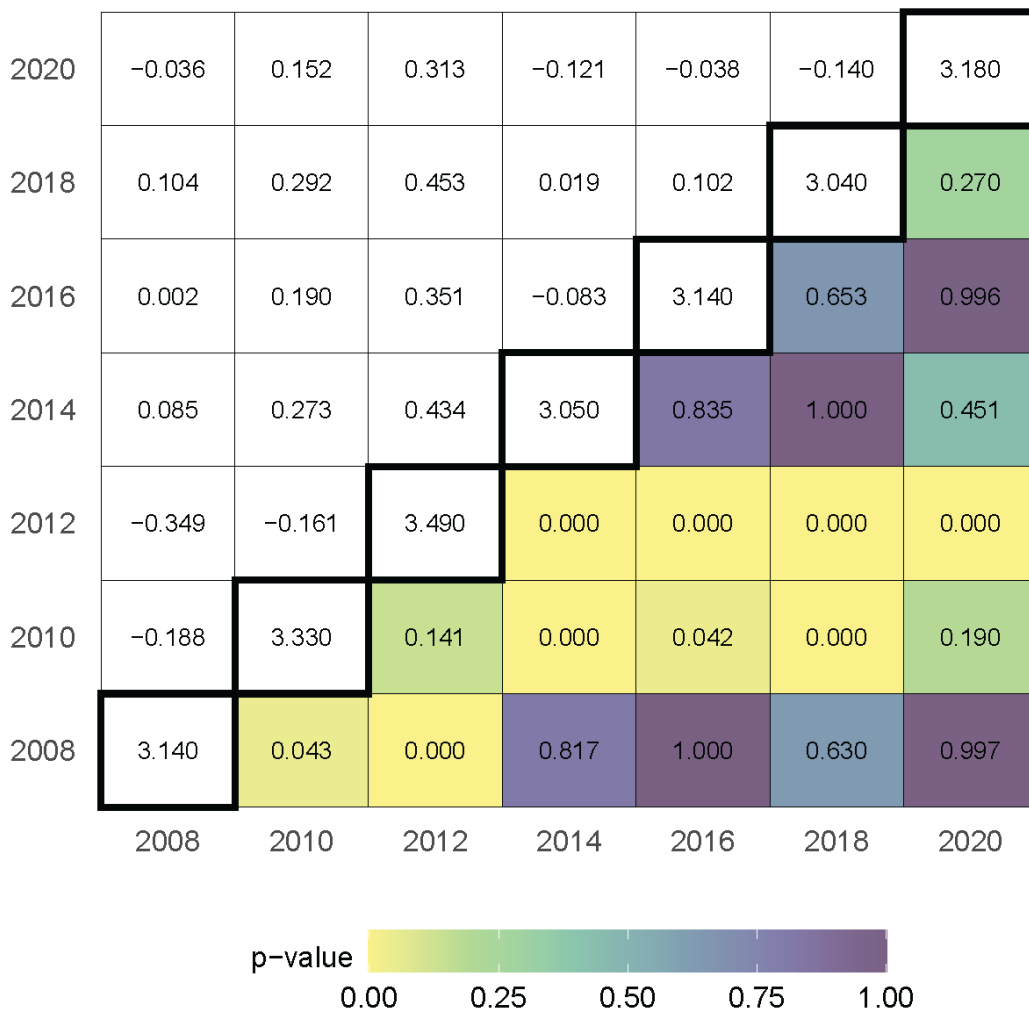
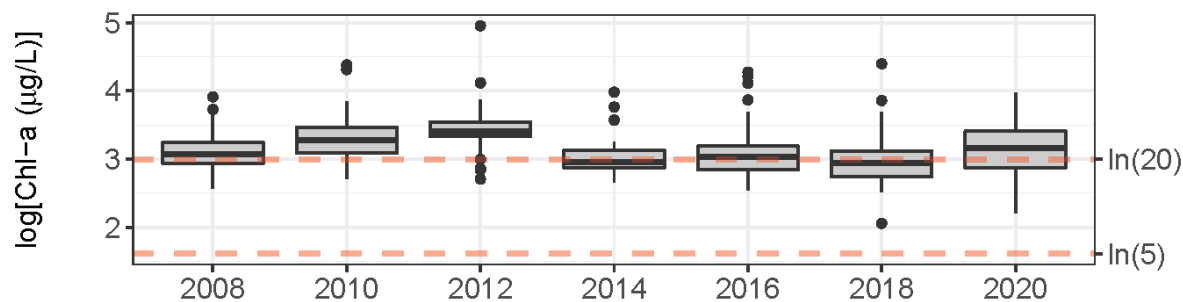


Figure 2-3. (Top): Boxplot of log surface chlorophyll-*a* concentration by year within Colonial National Historical Park (all strata). Estuarine quality category thresholds (Table 2-1) plotted as red dashed lines. (Bottom): Compact display of the multiple comparison test results. The bolded diagonal 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 2-17.

Table 2-17. Summary of the multiple comparison test results for all possible yearly contrasts of mean log surface chlorophyll-a concentration within Colonial National Historical Park (all strata). 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>₂₅₁	<i>p</i>-value	95% CI
2008 - 2010	-0.188	0.062	-3.022	0.043	(-0.373, -0.003)
2008 - 2012	-0.349	0.062	-5.636	0.000	(-0.533, -0.165)
2008 - 2014	0.085	0.062	1.370	0.817	(-0.099, 0.269)
2008 - 2016	0.002	0.062	0.029	1.000	(-0.182, 0.186)
2008 - 2018	0.104	0.062	1.681	0.629	(-0.080, 0.288)
2008 - 2020	-0.036	0.062	-0.585	0.997	(-0.220, 0.148)
2010 - 2012	-0.161	0.063	-2.567	0.141	(-0.346, 0.025)
2010 - 2014	0.273	0.063	4.366	0.000	(0.087, 0.459)
2010 - 2016	0.190	0.063	3.040	0.041	(0.004, 0.376)
2010 - 2018	0.292	0.063	4.674	0.000	(0.107, 0.478)
2010 - 2020	0.152	0.063	2.432	0.190	(-0.034, 0.338)
2012 - 2014	0.434	0.062	6.979	0.000	(0.249, 0.618)
2012 - 2016	0.351	0.062	5.643	0.000	(0.166, 0.535)
2012 - 2018	0.453	0.062	7.288	0.000	(0.268, 0.637)
2012 - 2020	0.313	0.062	5.032	0.000	(0.128, 0.497)
2014 - 2016	-0.083	0.062	-1.335	0.835	(-0.267, 0.102)
2014 - 2018	0.019	0.062	0.310	1.000	(-0.165, 0.204)
2014 - 2020	-0.121	0.062	-1.947	0.451	(-0.305, 0.064)
2016 - 2018	0.102	0.062	1.645	0.653	(-0.082, 0.287)
2016 - 2020	-0.038	0.062	-0.612	0.996	(-0.222, 0.146)
2018 - 2020	-0.140	0.062	-2.257	0.270	(-0.325, 0.044)

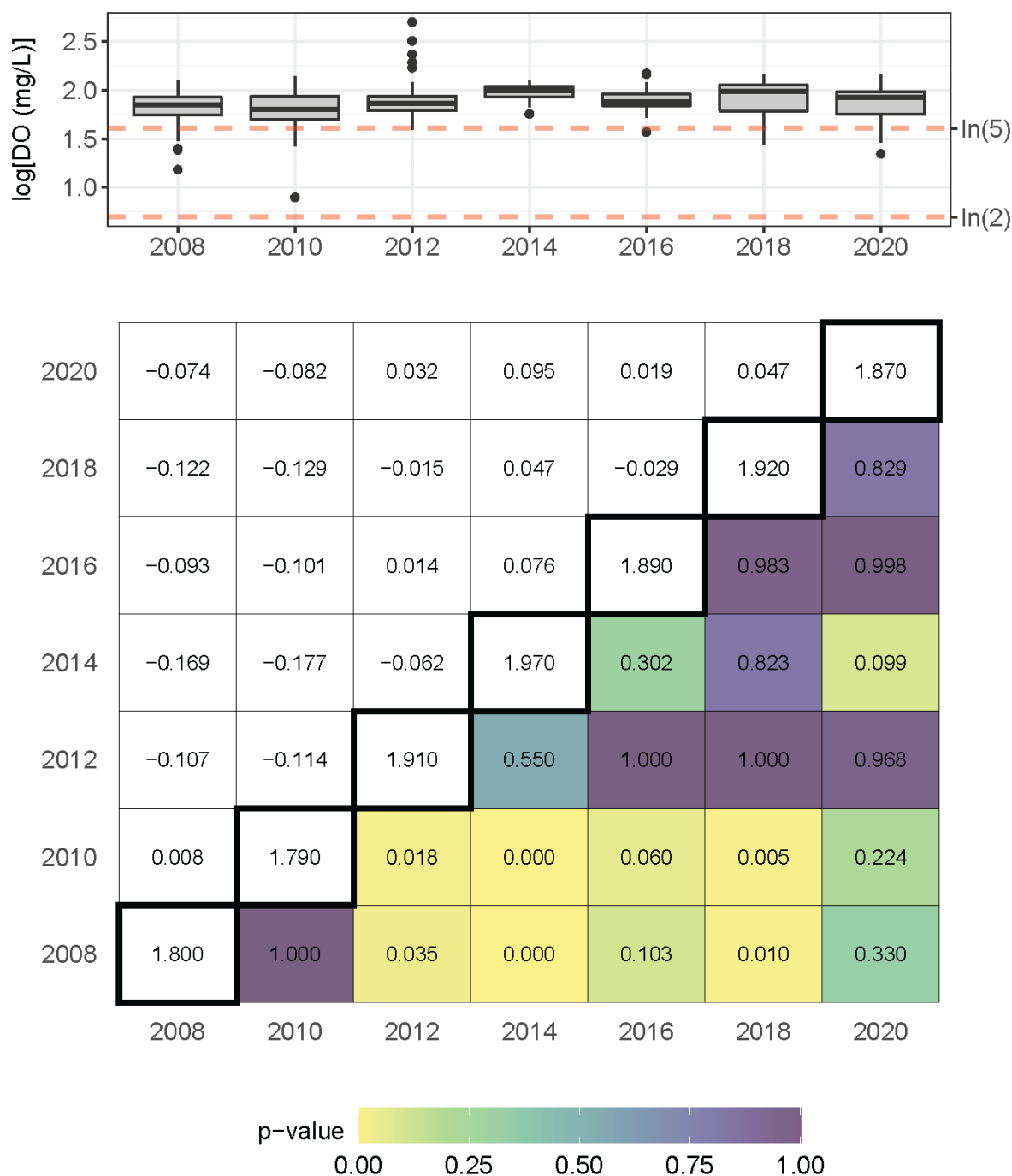


Figure 2-4. (Top): Boxplot of log bottom dissolved oxygen concentration by year within Colonial National Historical Park (all strata). Estuarine quality category thresholds (Table 2-1) plotted as red dashed lines. (Bottom): Compact display of the multiple comparison test results. The bolded diagonal 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 2-18.

Table 2-18. Summary of the multiple comparison test results for all possible yearly contrasts of mean log bottom dissolved oxygen concentration within Colonial National Historical Park (all strata). 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>₂₅₀	<i>p</i>-value	95% CI
2008 - 2010	0.008	0.034	0.219	1.000	(-0.095, 0.110)
2008 - 2012	-0.107	0.034	-3.098	0.035	(-0.209, -0.004)
2008 - 2014	-0.169	0.034	-4.905	0.000	(-0.271, -0.067)
2008 - 2016	-0.093	0.034	-2.699	0.103	(-0.195, 0.009)
2008 - 2018	-0.122	0.035	-3.507	0.009	(-0.225, -0.019)
2008 - 2020	-0.074	0.035	-2.145	0.330	(-0.177, 0.029)
2010 - 2012	-0.114	0.035	-3.301	0.019	(-0.217, -0.011)
2010 - 2014	-0.177	0.035	-5.100	0.000	(-0.279, -0.074)
2010 - 2016	-0.101	0.035	-2.905	0.060	(-0.203, 0.002)
2010 - 2018	-0.129	0.035	-3.708	0.005	(-0.233, -0.026)
2010 - 2020	-0.082	0.035	-2.352	0.224	(-0.185, 0.022)
2012 - 2014	-0.062	0.035	-1.799	0.550	(-0.165, 0.041)
2012 - 2016	0.014	0.035	0.396	1.000	(-0.089, 0.117)
2012 - 2018	-0.015	0.035	-0.428	1.000	(-0.118, 0.089)
2012 - 2020	0.032	0.035	0.928	0.968	(-0.071, 0.136)
2014 - 2016	0.076	0.035	2.195	0.302	(-0.027, 0.179)
2014 - 2018	0.047	0.035	1.360	0.822	(-0.056, 0.151)
2014 - 2020	0.095	0.035	2.716	0.099	(-0.009, 0.198)
2016 - 2018	-0.029	0.035	-0.821	0.983	(-0.132, 0.075)
2016 - 2020	0.019	0.035	0.535	0.998	(-0.085, 0.122)
2018 - 2020	0.047	0.035	1.347	0.829	(-0.057, 0.151)

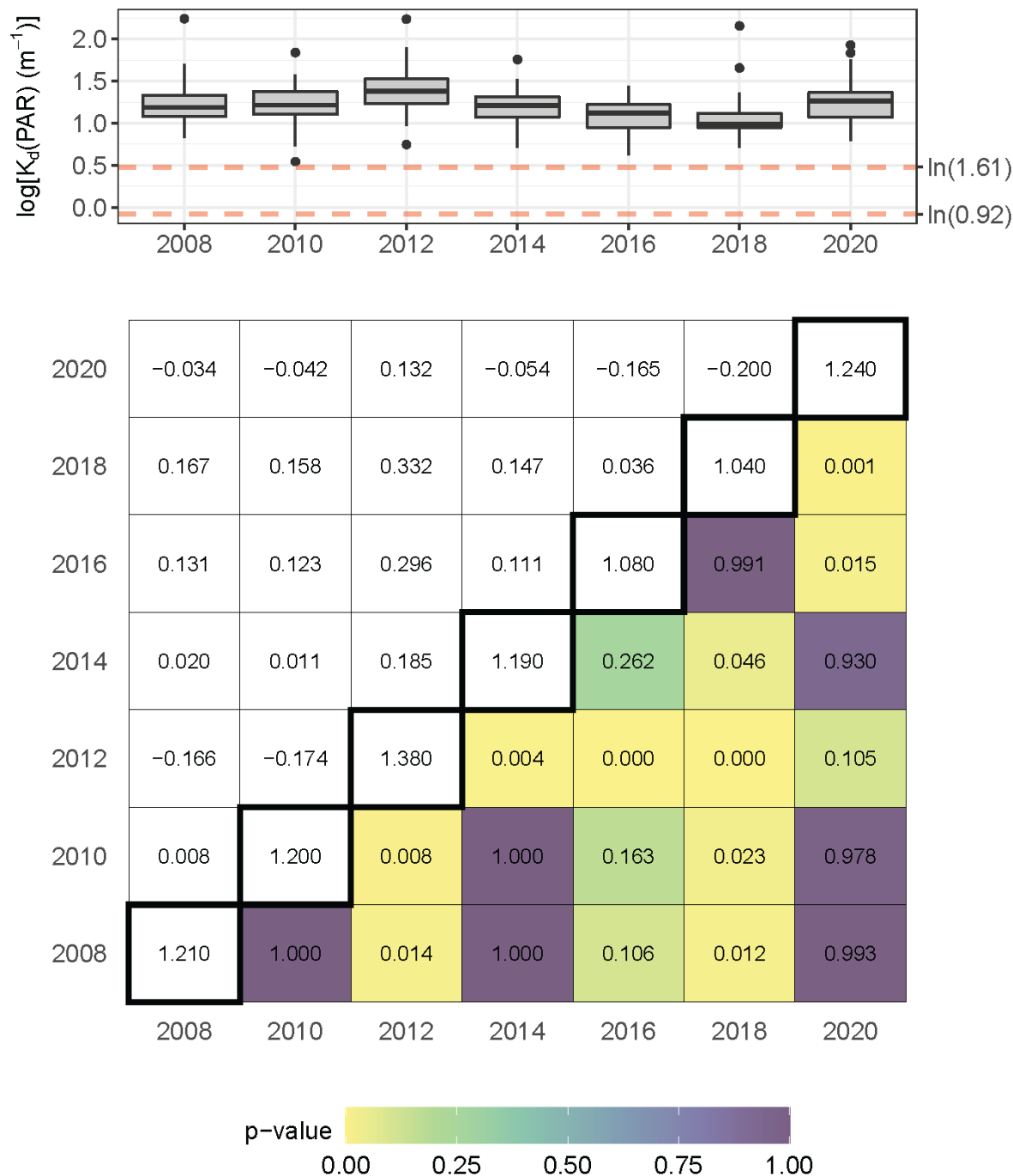


Figure 2-5. (Top): Boxplot of log PAR attenuation by year within Colonial National Historical Park (all strata). Estuarine quality category thresholds (Table 2-1) plotted as red dashed lines. (Bottom): Compact display of the multiple comparison test results. The bolded diagonal elements contain the mean 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 2-19.

Table 2-19. Summary of the multiple comparison test results for all possible yearly contrasts of mean log PAR attenuation within Colonial National Historical Park (all strata). 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> ₂₅₂	<i>p</i> -value	95% CI
2008 - 2010	0.008	0.049	0.172	1.000	(−0.137, 0.153)
2008 - 2012	−0.166	0.049	−3.397	0.014	(−0.311, −0.021)
2008 - 2014	0.020	0.049	0.404	1.000	(−0.125, 0.165)
2008 - 2016	0.131	0.049	2.686	0.106	(−0.014, 0.276)
2008 - 2018	0.167	0.049	3.420	0.013	(0.022, 0.312)
2008 - 2020	−0.034	0.049	−0.694	0.993	(−0.179, 0.111)
2010 - 2012	−0.174	0.049	−3.550	0.008	(−0.320, −0.028)
2010 - 2014	0.011	0.049	0.230	1.000	(−0.134, 0.157)
2010 - 2016	0.123	0.049	2.501	0.163	(−0.023, 0.268)
2010 - 2018	0.158	0.049	3.230	0.023	(0.013, 0.304)
2010 - 2020	−0.042	0.049	−0.862	0.978	(−0.188, 0.104)
2012 - 2014	0.185	0.049	3.781	0.004	(0.040, 0.331)
2012 - 2016	0.296	0.049	6.051	0.000	(0.151, 0.442)
2012 - 2018	0.332	0.049	6.781	0.000	(0.186, 0.478)
2012 - 2020	0.132	0.049	2.688	0.105	(−0.014, 0.277)
2014 - 2016	0.111	0.049	2.270	0.263	(−0.034, 0.257)
2014 - 2018	0.147	0.049	3.000	0.046	(0.001, 0.293)
2014 - 2020	−0.054	0.049	−1.092	0.930	(−0.199, 0.092)
2016 - 2018	0.036	0.049	0.730	0.991	(−0.110, 0.181)
2016 - 2020	−0.165	0.049	−3.363	0.015	(−0.310, −0.019)
2018 - 2020	−0.200	0.049	−4.092	0.001	(−0.346, −0.055)

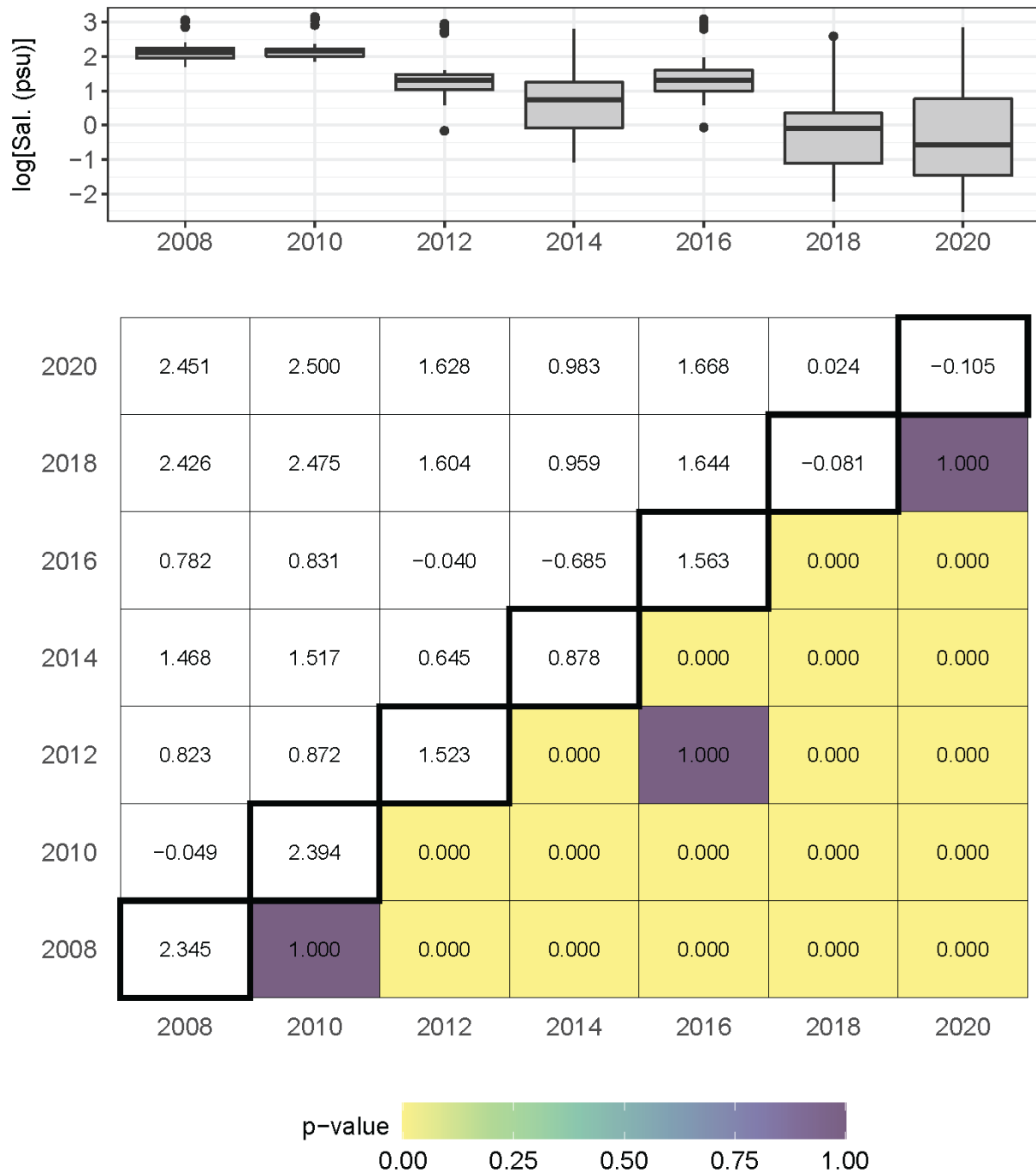


Figure 2-6. (Top): Boxplot of log surface salinity by year within Colonial National Historical Park (all strata). (Bottom): Compact display of the multiple comparison test results. The bolded diagonal elements contain the mean log surface salinity within each year. Upper diagonal cells contain the difference 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 2-20.

Table 2-20. Summary of the multiple comparison test results for all possible yearly contrasts of mean log surface salinity within Colonial National Historical Park (all strata). 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> ₂₅₂	<i>p</i> -value	95% CI
2008 - 2010	-0.049	0.134	-0.365	1.000	(-0.446, 0.349)
2008 - 2012	0.823	0.134	6.152	0.000	(0.425, 1.220)
2008 - 2014	1.468	0.134	10.974	0.000	(1.070, 1.865)
2008 - 2016	0.782	0.134	5.850	0.000	(0.385, 1.180)
2008 - 2018	2.426	0.134	18.143	0.000	(2.029, 2.824)
2008 - 2020	2.451	0.134	18.325	0.000	(2.053, 2.848)
2010 - 2012	0.872	0.134	6.497	0.000	(0.473, 1.270)
2010 - 2014	1.517	0.134	11.304	0.000	(1.118, 1.915)
2010 - 2016	0.831	0.134	6.196	0.000	(0.433, 1.230)
2010 - 2018	2.475	0.134	18.449	0.000	(2.077, 2.874)
2010 - 2020	2.500	0.134	18.630	0.000	(2.101, 2.898)
2012 - 2014	0.645	0.134	4.807	0.000	(0.246, 1.044)
2012 - 2016	-0.040	0.134	-0.301	1.000	(-0.439, 0.358)
2012 - 2018	1.604	0.134	11.953	0.000	(1.205, 2.002)
2012 - 2020	1.628	0.134	12.134	0.000	(1.229, 2.027)
2014 - 2016	-0.685	0.134	-5.108	0.000	(-1.084, -0.287)
2014 - 2018	0.959	0.134	7.146	0.000	(0.560, 1.357)
2014 - 2020	0.983	0.134	7.327	0.000	(0.584, 1.382)
2016 - 2018	1.644	0.134	12.254	0.000	(1.245, 2.043)
2016 - 2020	1.668	0.134	12.435	0.000	(1.270, 2.067)
2018 - 2020	0.024	0.134	0.181	1.000	(-0.374, 0.423)

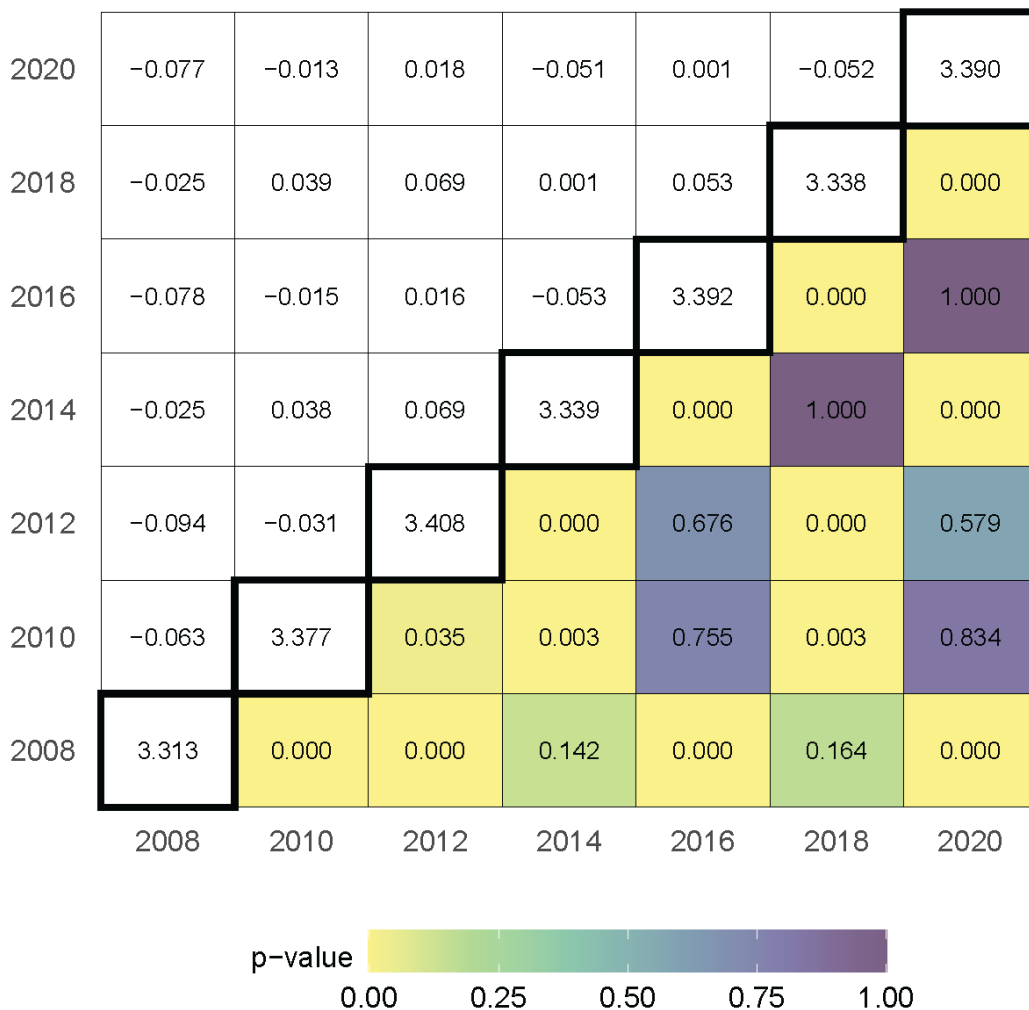
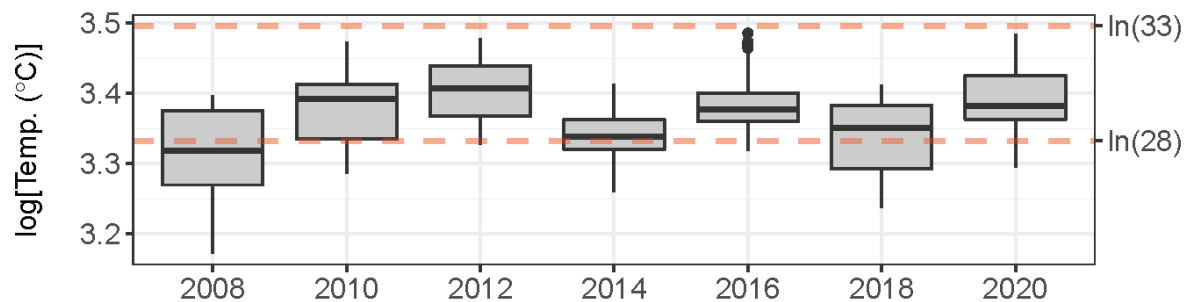


Figure 2-7. (Top): Boxplot of log surface temperature by year within Colonial National Historical Park (all strata). Estuarine quality category thresholds (Table 2-1) plotted as red dashed lines. (Bottom): Compact display of the multiple comparison test results. The bolded diagonal elements contain the 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 2-21.

Table 2-21. Summary of the multiple comparison test results for all possible yearly contrasts of mean log surface temperature within Colonial National Historical Park (all strata). Note: all estimates and confidence interval bounds are reported on the log scale. [SE = standard error; CI = confidence interval]

Contrast	Estimate	SE	t_{252}	p -value	95% CI
2008 - 2010	-0.063	0.010	-6.379	0.000	(-0.093, -0.034)
2008 - 2012	-0.094	0.010	-9.486	0.000	(-0.124, -0.065)
2008 - 2014	-0.025	0.010	-2.561	0.142	(-0.055, 0.004)
2008 - 2016	-0.078	0.010	-7.868	0.000	(-0.108, -0.049)
2008 - 2018	-0.025	0.010	-2.500	0.164	(-0.054, 0.005)
2008 - 2020	-0.077	0.010	-7.721	0.000	(-0.106, -0.047)
2010 - 2012	-0.031	0.010	-3.092	0.036	(-0.061, -0.001)
2010 - 2014	0.038	0.010	3.799	0.003	(0.008, 0.068)
2010 - 2016	-0.015	0.010	-1.482	0.755	(-0.045, 0.015)
2010 - 2018	0.039	0.010	3.860	0.003	(0.009, 0.068)
2010 - 2020	-0.013	0.010	-1.336	0.834	(-0.043, 0.016)
2012 - 2014	0.069	0.010	6.890	0.000	(0.039, 0.099)
2012 - 2016	0.016	0.010	1.609	0.677	(-0.014, 0.046)
2012 - 2018	0.069	0.010	6.951	0.000	(0.040, 0.099)
2012 - 2020	0.018	0.010	1.756	0.579	(-0.012, 0.047)
2014 - 2016	-0.053	0.010	-5.281	0.000	(-0.082, -0.023)
2014 - 2018	0.001	0.010	0.061	1.000	(-0.029, 0.030)
2014 - 2020	-0.051	0.010	-5.134	0.000	(-0.081, -0.022)
2016 - 2018	0.053	0.010	5.342	0.000	(0.024, 0.083)
2016 - 2020	0.001	0.010	0.146	1.000	(-0.028, 0.031)
2018 - 2020	-0.052	0.010	-5.196	0.000	(-0.082, -0.022)

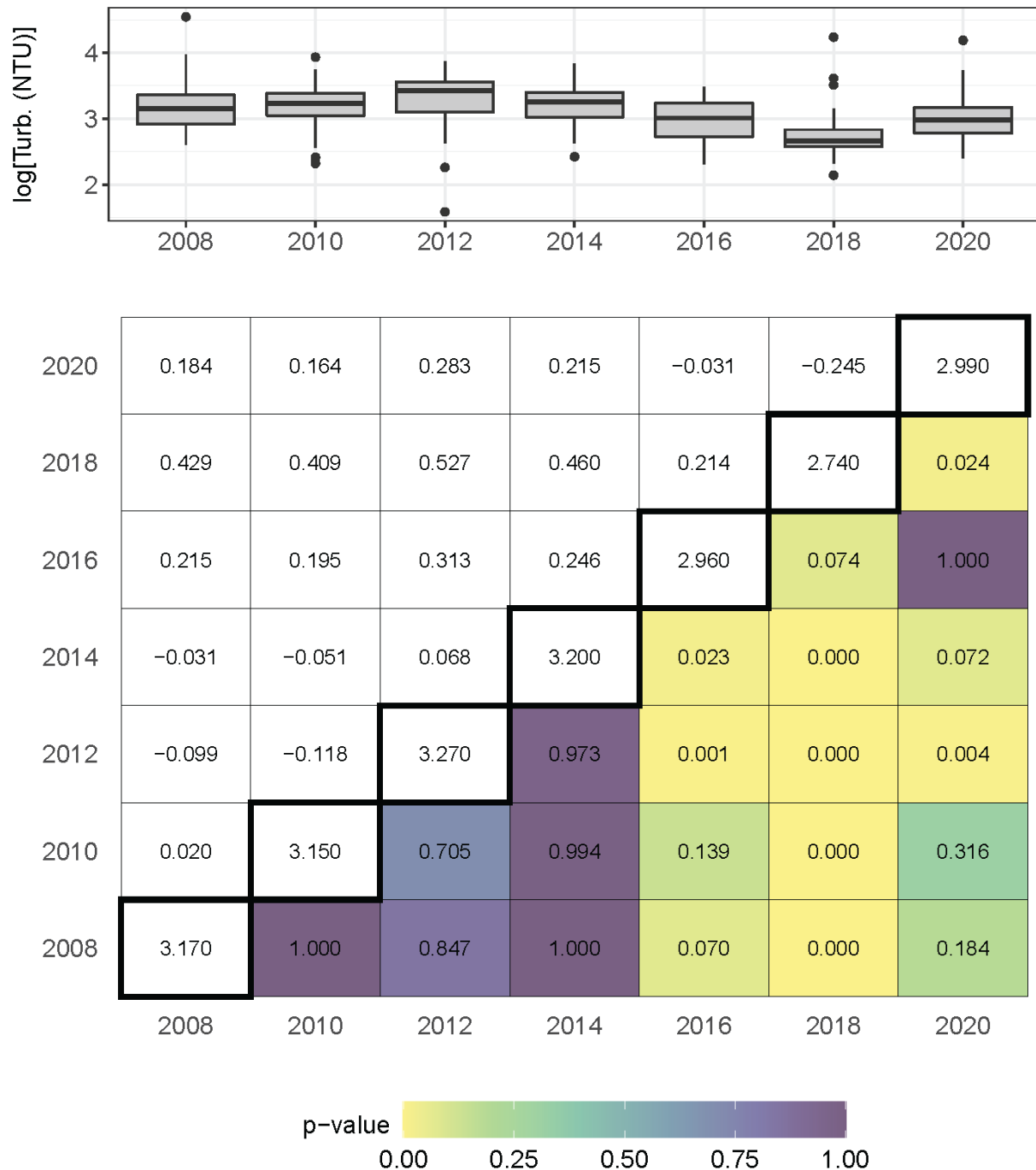


Figure 2-8. (Top): Boxplot of log surface turbidity by year within Colonial National Historical Park (all strata). (Bottom): Compact display of the multiple comparison test results. The bolded diagonal elements contain the mean log surface turbidity within each year. Upper diagonal cells contain the 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 2-22.

Table 2-22. Summary of the multiple comparison test results for all possible yearly contrasts of mean log surface turbidity within Colonial National Historical Park (all strata). Note: all estimates and confidence interval bounds are reported on the log scale. [SE = standard error; CI = confidence interval]

Contrast	Estimate	SE	t ₂₅₂	p-value	95% CI
2008 - 2010	0.020	0.075	0.263	1.000	(-0.204, 0.244)
2008 - 2012	-0.099	0.075	-1.310	0.847	(-0.323, 0.125)
2008 - 2014	-0.031	0.075	-0.408	1.000	(-0.255, 0.193)
2008 - 2016	0.215	0.075	2.850	0.069	(-0.009, 0.439)
2008 - 2018	0.429	0.075	5.691	0.000	(0.205, 0.653)
2008 - 2020	0.184	0.075	2.444	0.185	(-0.040, 0.408)
2010 - 2012	-0.119	0.076	-1.565	0.705	(-0.344, 0.107)
2010 - 2014	-0.051	0.076	-0.668	0.994	(-0.276, 0.175)
2010 - 2016	0.195	0.076	2.574	0.139	(-0.030, 0.420)
2010 - 2018	0.409	0.076	5.399	0.000	(0.184, 0.634)
2010 - 2020	0.164	0.076	2.170	0.316	(-0.061, 0.389)
2012 - 2014	0.068	0.076	0.897	0.973	(-0.157, 0.293)
2012 - 2016	0.313	0.076	4.138	0.001	(0.088, 0.539)
2012 - 2018	0.527	0.076	6.964	0.000	(0.302, 0.753)
2012 - 2020	0.283	0.076	3.735	0.004	(0.058, 0.508)
2014 - 2016	0.246	0.076	3.242	0.023	(0.020, 0.471)
2014 - 2018	0.460	0.076	6.067	0.000	(0.234, 0.685)
2014 - 2020	0.215	0.076	2.838	0.072	(-0.010, 0.440)
2016 - 2018	0.214	0.076	2.826	0.075	(-0.011, 0.439)
2016 - 2020	-0.031	0.076	-0.404	1.000	(-0.256, 0.195)
2018 - 2020	-0.245	0.076	-3.229	0.023	(-0.470, -0.019)

2.4.3 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 2-23). Notable observations from the fitted smooths for each measure are as follows:

Chlorophyll-*a* (near surface)

Chlorophyll-*a* exhibited a non-linear temporal trend from 2008 to 2014 when there were two periods of statistically significant changes (Figure 2-9). Higher concentrations of chlorophyll-*a* occur for sites located in the York River compared to the Thoroughfare (Figure 2-10). Within the Thoroughfare, chlorophyll-*a* appears to be slightly lower in Sandy Bay. Based on the fitted model, both time and location account for 51.4% of the variability in chlorophyll-*a* concentrations (Table 2-24).

Dissolved Oxygen (near bottom)

Bottom dissolved oxygen has a slightly increasing trend from 2008 to 2014 that is statistically significant. Concentrations have remained relatively constant since 2014. For sites located in the York River, bottom dissolved oxygen tends to be lower along the northwest shoreline compared to the southeast. Within the Thoroughfare stratum, concentrations tend to be lower in the Back River,

Sandy Bay, and near the outlet of Powhatan Creek (Figure 2-11). The fitted model does very poorly explaining variation in dissolved oxygen concentrations, with time and location accounting for only 16.1% of the variation (Table 2-24).

PAR Attenuation K_d (SAV standard)

The temporal trend in PAR attenuation for 2014 to 2020 is slightly concave upward with two periods of statistically significant change (Figure 2-9). For sites located in the York River, PAR attenuation tends to be higher along the northwest shoreline compared to the southeast. Within the Thoroughfare stratum, PAR attenuation tends to be higher in the Back River, Sandy Bay, and near the outlet of Powhatan Creek (Figure 2-12). Based on the fitted model, both time and location account for only 21.7% of the variability in PAR attenuation (Table 2-24).

Salinity (near surface)

Surface salinity exhibited a highly non-linear temporal trend ($\text{edf} = 5.83$) with statistically significant decreasing trends during the period from 2010 to 2012 and 2016 to 2018 (Figure 2-9). There was also a statistically significant increase from 2014 to 2016. Spatially, surface salinity concentrations are much higher for sites in the York River compared to the Thoroughfare (Figure 2-13). Within the Thoroughfare, salinity decreases with increasing proximity to the outlet of the Back River. The fitted GAM had the best fit amongst the measures, with time and location accounting for 76.5% of the variation in salinity (Table 2-24).

Temperature (near surface)

Surface water temperature exhibited a highly non-linear temporal trend ($\text{edf} = 5.82$) with cyclic periods of statistically significant increases and decreases (Figure 2-9 (continued)). The mean trend has been above 28 °C since 2010. Spatially, surface temperatures are the highest near the outlet of College Creek and the site near Yorktown (Figure 2-14). Within the Thoroughfare, temperatures tend to be cooler in the Back River, Sandy Bay, and near the outlet of Powhatan Creek. Based on the fitted model, both time and location account for only 34.7% of the variability in water temperature (Table 2-24).

Turbidity (near surface)

The temporal trend in surface turbidity indicates a statistically significant decrease occurred from 2014 to 2016 (Figure 2-9 (continued)). Spatially, turbidity is highest within the Thoroughfare stratum and the more northerly sites on the York River (Figure 2-15). Within the Thoroughfare, turbidity tends to be highest near the outlet of the Back River and within the Back River downstream of Sandy Bay. Time and location account for only 25.5% of the variation in turbidity (Table 2-24).

Table 2-23. Approximate significance of the generalized additive model smooth terms for each Colonial National Historical Park (all strata) 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(s)	Model Term	edf	F -value	p -value
Chlorophyll a ($\mu\text{g/L}$)	Surface	s(year)	5.48	10.86	< 0.001
Chlorophyll a ($\mu\text{g/L}$)	Surface	s(long, lat)	13.97	2.88	< 0.001
Dissolved oxygen (mg/L)	Bottom	s(year)	2.57	4.19	< 0.001
Dissolved oxygen (mg/L)	Bottom	s(long, lat)	4.72	0.48	< 0.001
PAR attenuation (K_d (m^{-1}))	Water Column	s(year)	3.72	9.77	< 0.001
PAR attenuation (K_d (m^{-1}))	Water Column	s(long, lat)	6.34	0.44	0.001
Salinity (psu)	Surface	s(year)	5.83	146.53	< 0.001
Salinity (psu)	Surface	s(long, lat)	6.90	12.07	< 0.001
Temperature ($^{\circ}\text{C}$)	Surface	s(year)	5.82	21.72	< 0.001
Temperature ($^{\circ}\text{C}$)	Surface	s(long, lat)	3.47	0.34	0.001
Turbidity (NTU)	Surface	s(year)	4.27	7.95	< 0.001
Turbidity (NTU)	Surface	s(long, lat)	5.98	0.51	< 0.001

Table 2-24. Proportion of null deviance explained by the explanatory variables for each Colonial National Historical Park (all strata) model.

Measure	Deviance Explained
Chlorophyll a ($\mu\text{g/L}$)	51.4%
Dissolved oxygen (mg/L)	16.1%
PAR attenuation (K_d (m^{-1}))	21.7%
Salinity (psu)	76.5%
Temperature ($^{\circ}\text{C}$)	34.7%
Turbidity (NTU)	25.5%

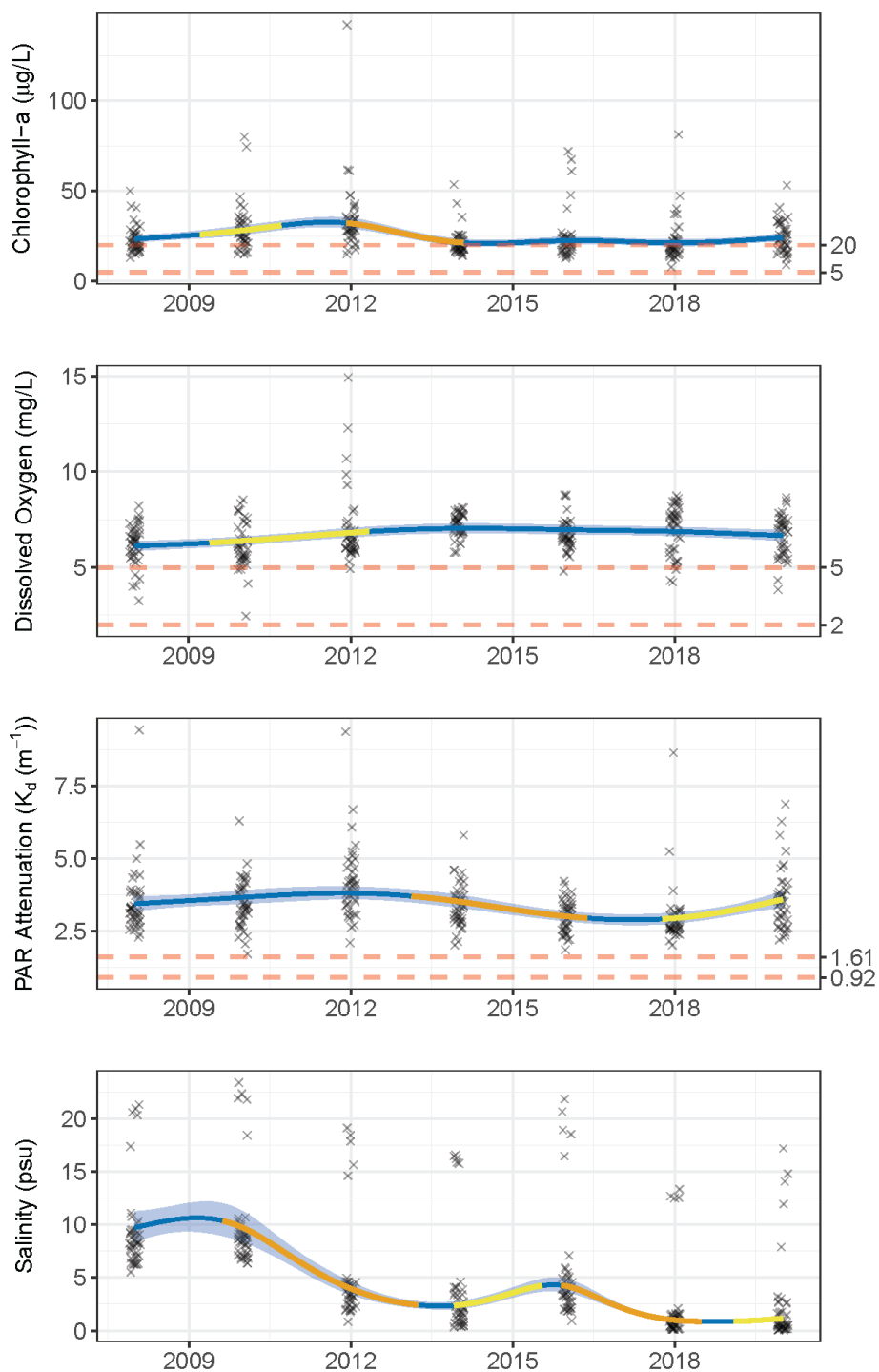


Figure 2-9. Trends in Colonial National Historical Park (all strata) water quality measures based on estimated generalized additive model smooths. The fitted smooth is plotted in dark blue with shaded 95% confidence interval. Periods of significant temporal change (increases/decreases), if present, are plotted in yellow/orange. Estuarine quality category thresholds (Table 2-1) plotted as red dashed lines.

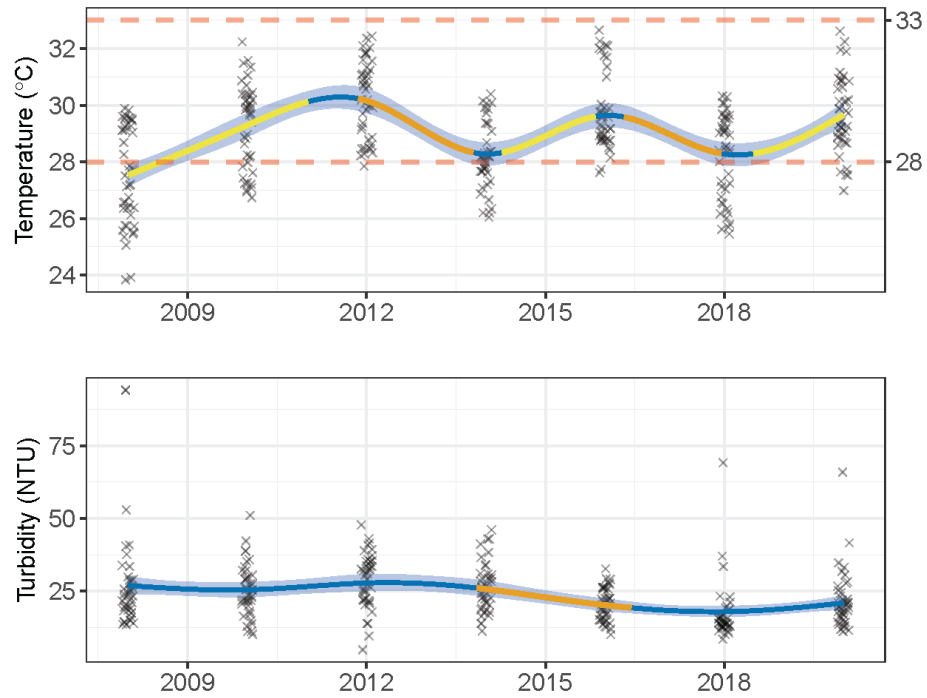


Figure 2-9 (continued). Trends in Colonial National Historical Park (all strata) water quality measures based on estimated generalized additive model smooths. The fitted smooth is plotted in dark blue with shaded 95% confidence interval. Periods of significant temporal change (increases/decreases), if present, are plotted in yellow/orange. Estuarine quality category thresholds (Table 2-1) plotted as red dashed lines.

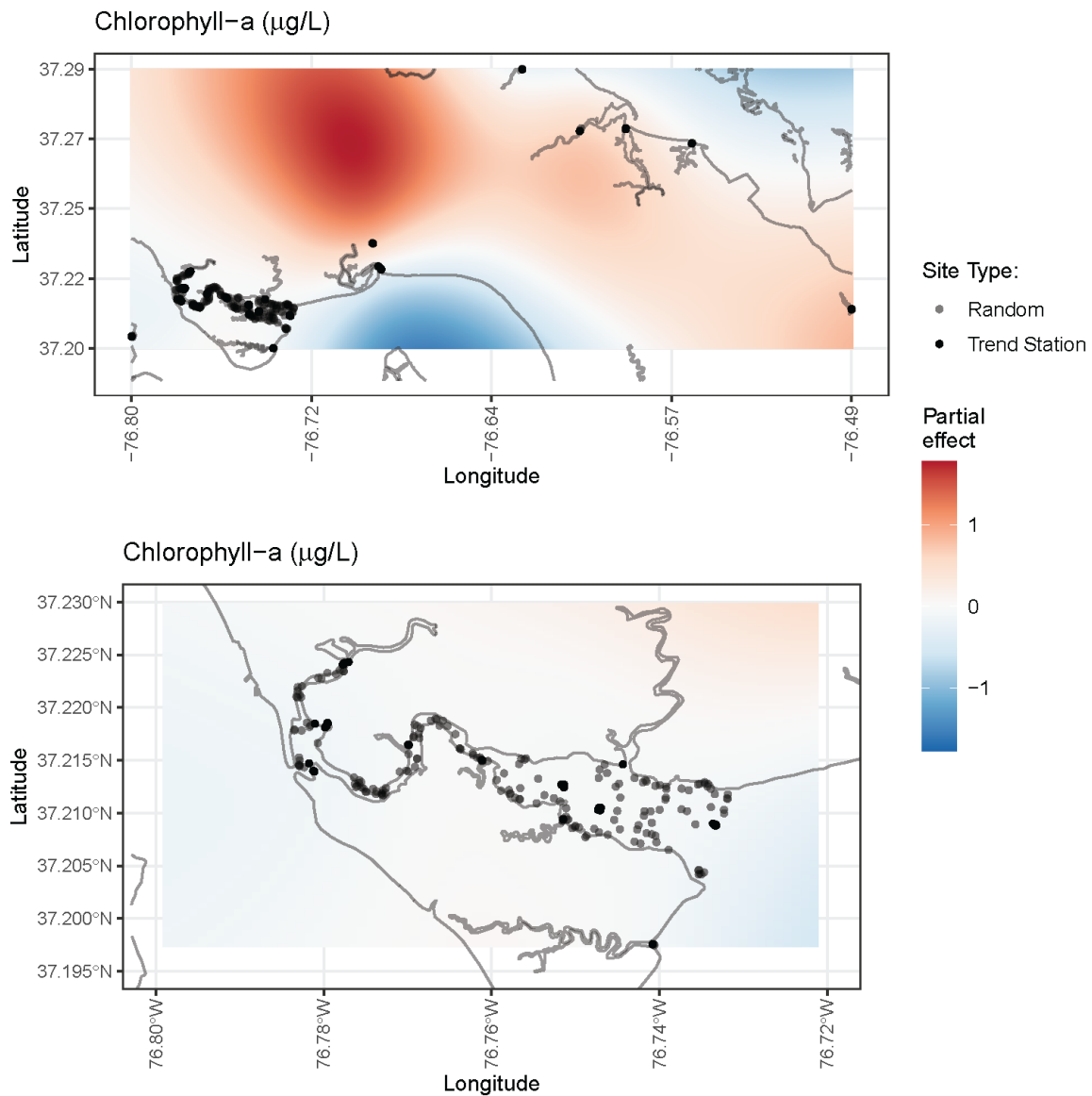


Figure 2-10. Mean surface for the generalized additive model latitude/longitude smooth of surface chlorophyll-a concentration within Colonial National Historical Park (top = all strata; bottom = Thoroughfare stratum inset) during 2008-2020.

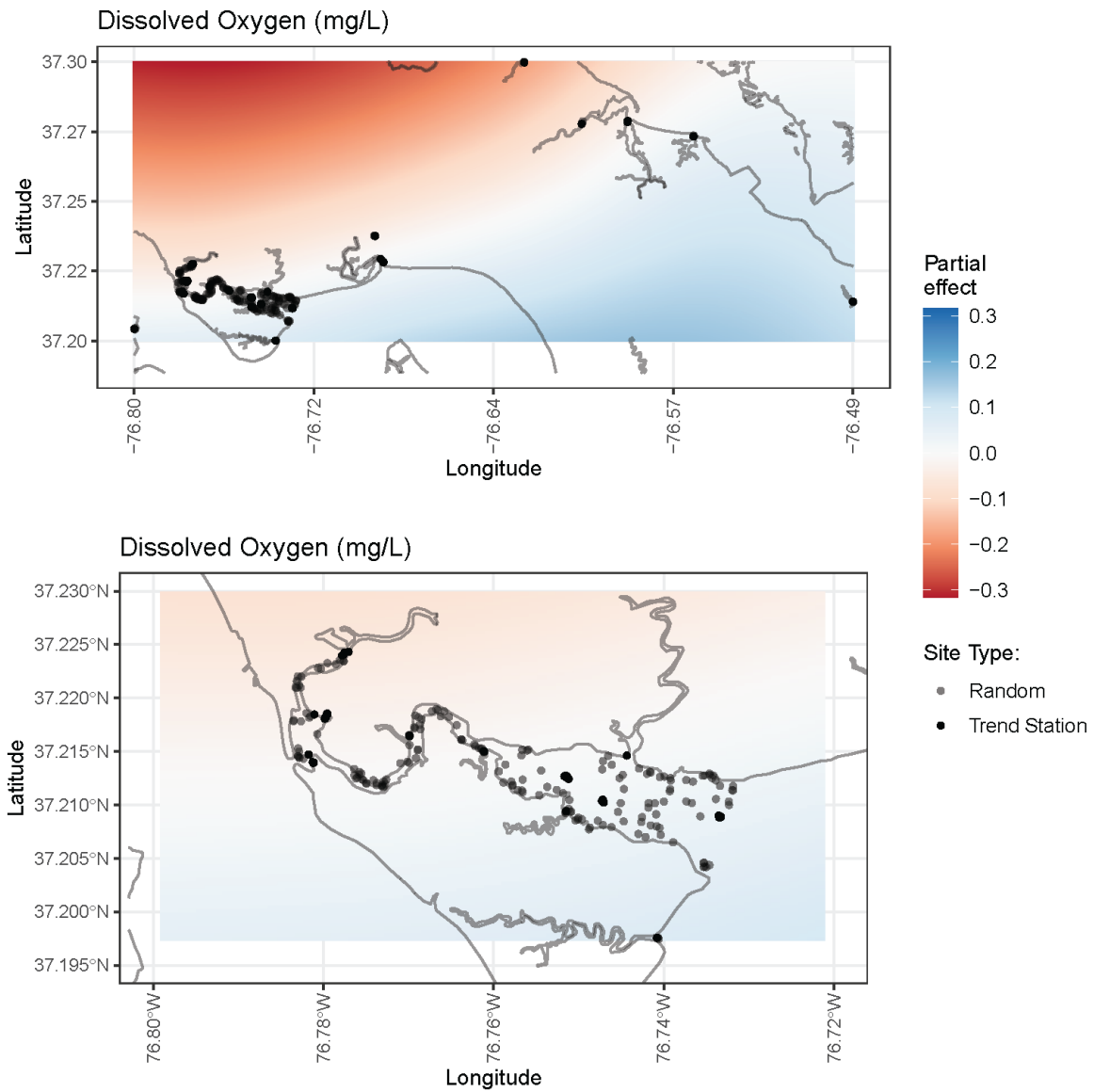


Figure 2-11. Mean surface for the generalized additive model latitude/longitude smooth of bottom dissolved oxygen concentration within Colonial National Historical Park (top = all strata; bottom = Thoroughfare stratum inset) during 2008-2020.

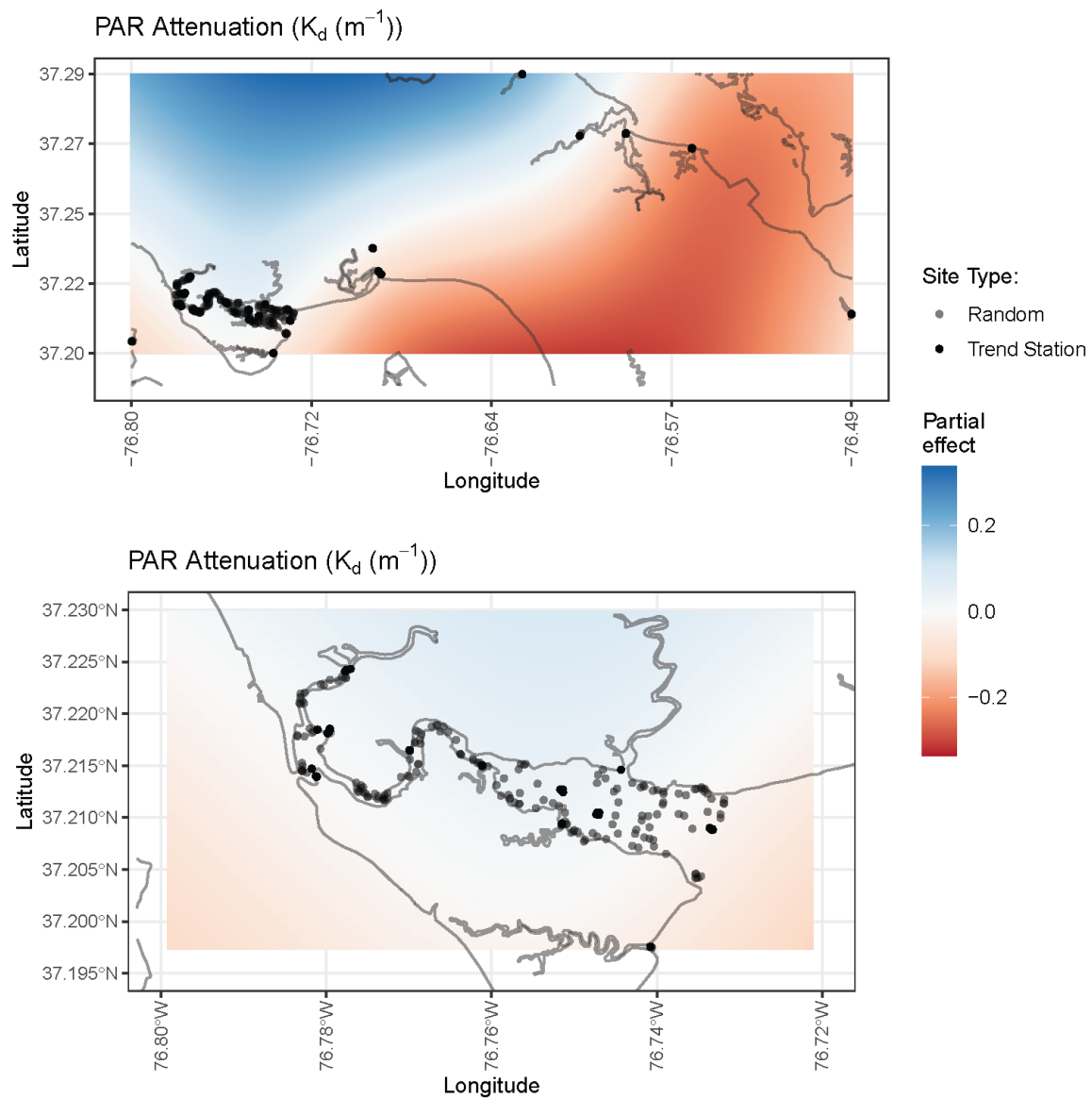


Figure 2-12. Mean surface for the generalized additive model latitude/longitude smooth of PAR attenuation within Colonial National Historical Park (top = all strata; bottom = Thoroughfare stratum inset) during 2008-2020.

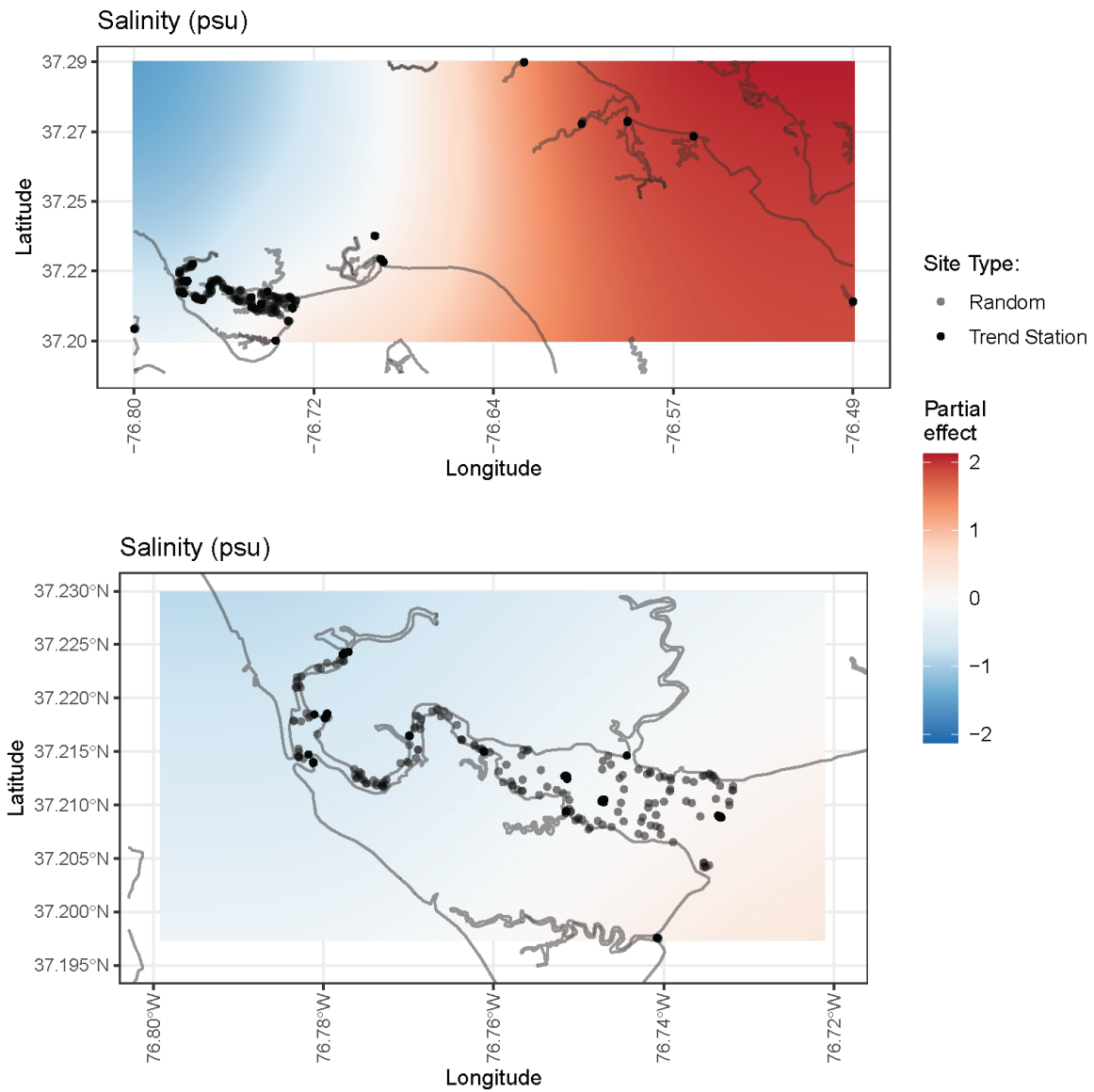


Figure 2-13. Mean surface for the generalized additive model latitude/longitude smooth of surface salinity within Colonial National Historical Park (top = all strata; bottom = Thoroughfare stratum inset) during 2008-2020.

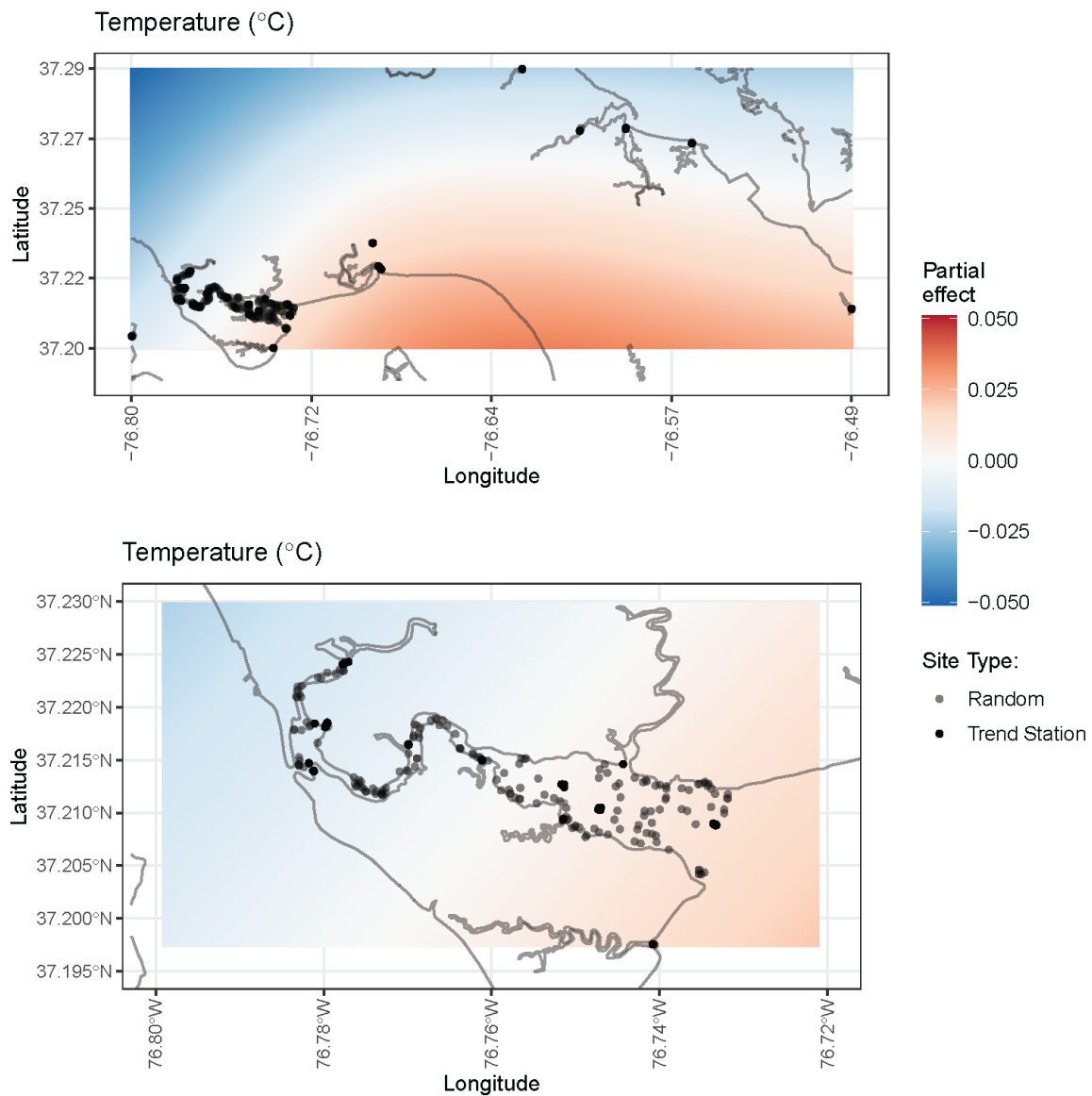


Figure 2-14. Mean surface for the generalized additive model latitude/longitude smooth of surface temperature within Colonial National Historical Park (top = all strata; bottom = Thoroughfare stratum inset) during 2008-2020.

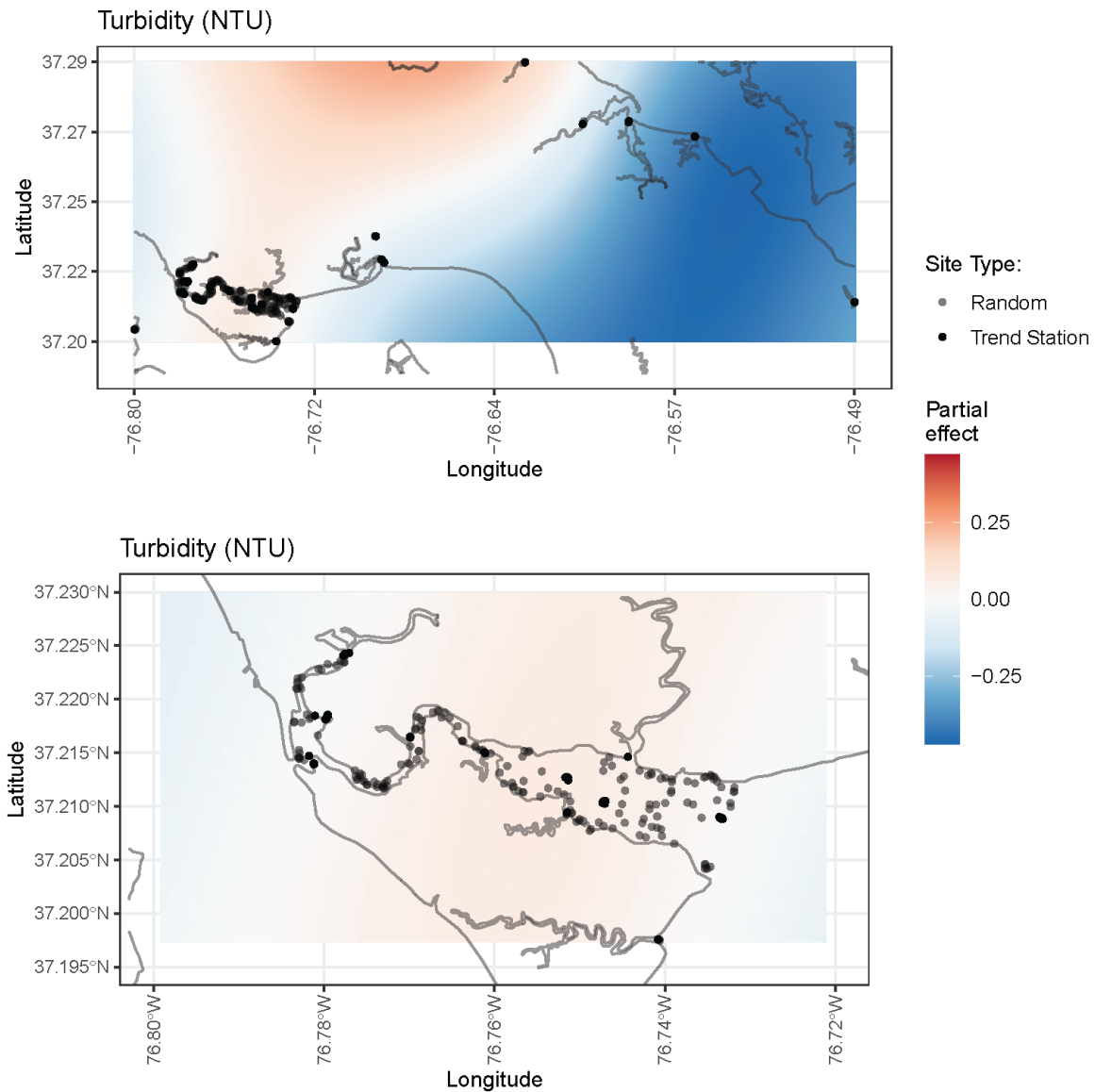


Figure 2-15. Mean surface for the generalized additive model latitude/longitude smooth of surface turbidity within Colonial National Historical Park (top = all strata; bottom = Thoroughfare stratum inset) during 2008-2020.

2.4.4 Multivariate Mann-Kendall Trend Test Results

The multivariate Mann-Kendall test results indicate that there is a statistically significant decreasing monotonic trend in surface salinity (p -value = 0.019) during the index periods of 2008 to 2020 (Table 2-25).

Table 2-25. Summary of multivariate Mann-Kendall test for monotonic trend for each Colonial National Historical Park (all strata) 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	−0.50	0.619
Dissolved oxygen (mg/L)	Bottom	1.69	0.091
PAR attenuation (K_d (m^{-1}))	Water Column	−0.71	0.477
Salinity (psu)	Surface	−2.35	0.019
Temperature ($^{\circ}\text{C}$)	Surface	0.83	0.407
Turbidity (NTU)	Surface	−1.54	0.122

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 Colonial National Historical Park trend stations. [SD = standard deviation]

Station	Statistic	Sample Year						
		2008	2010	2012	2014	2016	2018	2020
5	Mean	27.25	25.23	31.02	20.28	18.16	17.57	27.20
5	SD	17.22	2.76	7.70	5.82	2.18	1.69	8.27
5	n	4	3	4	4	4	4	4
10	Mean	20.24	35.20	32.56	24.89	15.88	15.94	29.31
10	SD	2.39	13.68	9.58	9.98	3.34	2.50	11.71
10	n	4	3	4	4	4	4	4
15	Mean	17.69	26.63	23.90	18.43	17.76	24.79	22.28
15	SD	1.99	4.36	3.99	0.94	4.21	6.37	10.06
15	n	4	3	4	4	4	4	4
20	Mean	17.61	26.80	30.70	18.53	20.15	19.78	28.29
20	SD	4.29	9.82	5.29	3.61	3.05	2.14	10.17
20	n	4	3	4	4	4	4	4
25	Mean	18.17	21.93	28.20	20.72	22.29	23.89	19.36
25	SD	2.70	4.99	9.06	1.34	7.18	7.76	6.57
25	n	4	3	4	4	4	4	4
30	Mean	22.65	25.00	28.33	22.48	21.97	22.11	25.33
30	SD	5.40	5.70	7.11	3.19	1.96	10.62	8.78
30	n	4	4	4	4	4	4	4
33	Mean	23.83	37.17	35.72	25.78	38.12	18.80	28.84
33	SD	11.87	29.05	9.53	4.04	16.93	6.60	11.29
33	n	4	4	4	4	4	4	4
37	Mean	19.12	30.23	32.73	19.12	19.00	15.99	24.11
37	SD	9.11	4.12	12.54	5.40	2.05	1.24	7.07
37	n	4	4	4	4	4	4	4

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

Station	Statistic	Sample Year						
		2008	2010	2012	2014	2016	2018	2020
5	Mean	6.52	5.77	6.49	7.44	6.66	7.47	7.04
5	SD	0.38	0.19	0.22	0.63	0.33	0.48	0.16
5	n	4	3	4	4	4	4	4
10	Mean	6.00	5.56	6.79	7.56	6.69	7.45	7.14
10	SD	0.41	0.68	0.26	0.63	0.15	0.67	0.27
10	n	4	3	4	4	4	4	4
15	Mean	6.47	6.77	7.72	7.68	6.59	7.64	6.90
15	SD	0.70	0.78	2.08	0.63	0.49	0.77	1.10
15	n	4	3	4	4	4	4	4
20	Mean	6.43	5.69	6.31	7.23	6.33	7.29	6.53
20	SD	0.70	0.44	0.18	0.60	0.42	0.88	0.30
20	n	4	3	4	4	4	4	4
25	Mean	6.29	6.39	10.93	7.68	7.81	7.38	6.76
25	SD	0.85	1.26	3.41	0.45	2.48	0.93	1.76
25	n	4	3	4	4	4	4	4
30	Mean	5.28	5.84	5.71	6.15	6.37	6.08	6.62
30	SD	1.26	1.18	0.87	0.51	1.68	1.86	1.08
30	n	4	4	4	4	4	4	4
33	Mean	6.78	5.45	7.05	6.37	8.19	6.11	5.29
33	SD	0.87	1.57	1.10	0.81	2.53	1.35	1.30
33	n	4	4	4	4	4	4	4
37	Mean	7.08	6.77	6.59	7.32	6.97	6.91	7.01
37	SD	0.59	1.21	0.85	0.54	0.30	0.86	0.66
37	n	4	4	4	4	4	3	4

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

Station	Statistic	Sample Year						
		2008	2010	2012	2014	2016	2018	2020
5	Mean	3.56	3.32	3.60	2.97	2.80	2.65	3.45
5	SD	0.97	0.50	0.86	1.06	0.69	0.30	0.62
5	n	4	3	4	4	4	4	4
10	Mean	4.49	3.79	3.59	3.22	3.09	2.43	3.37
10	SD	0.49	0.79	0.44	0.71	0.57	0.20	0.62
10	n	4	3	4	4	4	4	4
15	Mean	3.18	3.18	3.18	3.75	2.62	2.94	2.98
15	SD	0.22	0.45	0.52	0.58	0.21	0.36	0.72
15	n	4	3	4	4	4	4	4
20	Mean	2.90	3.25	3.55	3.11	3.17	2.58	3.23
20	SD	0.25	0.38	0.58	0.78	0.62	0.25	0.57
20	n	4	3	4	4	4	4	4
25	Mean	3.57	3.77	5.27	3.65	2.50	2.95	4.06
25	SD	0.42	0.25	1.85	1.47	0.59	0.40	1.32
25	n	4	3	4	4	4	4	4
30	Mean	2.95	3.23	3.57	3.42	2.50	2.95	3.17
30	SD	0.16	0.61	0.59	0.56	0.18	1.04	0.33
30	n	4	4	4	4	4	4	4
33	Mean	2.50	2.87	3.59	2.99	2.83	2.89	3.00
33	SD	0.49	1.10	1.60	1.16	1.04	0.57	0.63
33	n	4	4	4	4	4	4	4
37	Mean	2.49	2.44	3.44	2.86	3.06	2.61	3.38
37	SD	0.62	0.26	0.71	0.63	0.89	0.14	0.82
37	n	4	4	4	4	4	4	4

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

Station	Statistic	Sample Year						
		2008	2010	2012	2014	2016	2018	2020
5	Mean	9.38	9.71	4.32	3.21	4.66	1.45	1.11
5	SD	1.58	1.25	0.62	1.28	1.69	0.39	1.18
5	n	4	3	4	4	4	4	4
10	Mean	8.26	9.01	3.81	2.74	4.23	1.17	0.95
10	SD	1.38	1.09	0.86	1.28	1.10	0.32	1.15
10	n	4	3	4	4	4	4	4
15	Mean	7.42	7.37	2.49	1.19	2.78	0.35	0.33
15	SD	1.20	1.54	0.61	0.66	1.26	0.20	0.30
15	n	4	3	4	4	4	4	4
20	Mean	8.46	9.10	3.52	2.32	4.17	1.00	0.89
20	SD	2.24	1.13	0.65	1.12	1.69	0.61	1.18
20	n	4	3	4	4	4	4	4
25	Mean	7.38	7.31	2.20	1.15	2.80	0.36	0.31
25	SD	1.16	1.44	0.65	0.58	1.29	0.19	0.23
25	n	4	3	4	4	4	4	4
30	Mean	6.86	7.57	1.45	0.82	6.83	0.18	0.27
30	SD	1.50	1.20	0.82	0.61	9.30	0.08	0.19
30	n	4	4	4	4	4	4	4
33	Mean	21.19	22.62	18.62	16.85	20.07	12.80	14.06
33	SD	0.37	0.50	0.74	0.72	0.98	0.32	2.31
33	n	4	4	4	4	4	4	4
37	Mean	10.47	10.21	4.72	3.45	5.31	1.72	1.44
37	SD	1.22	1.24	0.67	1.46	1.49	0.31	1.43
37	n	4	4	4	4	4	4	4

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

Station	Statistic	Sample Year						
		2008	2010	2012	2014	2016	2018	2020
5	Mean	27.54	29.32	29.91	28.26	30.27	28.35	29.62
5	SD	1.60	1.13	1.78	1.11	1.34	1.17	1.54
5	n	4	3	4	4	4	4	4
10	Mean	27.02	28.92	29.99	28.58	30.17	28.11	30.10
10	SD	1.77	1.31	1.58	1.12	1.50	1.45	1.38
10	n	4	3	4	4	4	4	4
15	Mean	27.19	30.11	30.51	28.12	29.83	28.19	29.61
15	SD	1.72	2.56	0.96	1.34	1.47	1.35	1.24
15	n	4	3	4	4	4	4	4
20	Mean	27.44	29.11	30.16	28.22	29.93	28.02	29.94
20	SD	1.42	1.86	1.84	1.50	1.28	1.37	1.28
20	n	4	3	4	4	4	4	4
25	Mean	27.04	29.16	30.33	28.19	29.99	27.80	28.92
25	SD	1.80	1.74	1.95	0.81	1.95	1.25	1.82
25	n	4	3	4	4	4	4	4
30	Mean	27.27	29.21	29	28.07	30.35	27.44	30.38
30	SD	1.33	1.31	0.64	1.44	1.88	2.44	2.23
30	n	4	4	4	4	4	4	4
33	Mean	26.52	28.14	29.41	27.74	29.71	27.72	29.14
33	SD	1.94	0.84	0.81	1.86	2.44	0.87	1.10
33	n	4	4	4	4	4	4	4
37	Mean	27.57	29.46	29.81	28.67	29.85	28.20	29.43
37	SD	1.71	1.45	1.39	1.03	1.04	1.31	1.37
37	n	4	4	4	4	4	4	4

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

Station	Statistic	Sample Year						
		2008	2010	2012	2014	2016	2018	2020
5	Mean	28.50	24.70	25.70	24.77	19.77	13.11	22.20
5	SD	17.39	5.75	7.54	12.40	8.39	0.63	6.78
5	n	4	3	4	4	4	4	4
10	Mean	31.57	28.73	26.62	24.45	21.25	13.57	18.77
10	SD	6.44	10.05	3.95	11.61	7.59	1.05	5.19
10	n	3	3	4	4	4	4	4
15	Mean	23.25	32.30	40.35	29.92	18.12	14.81	17.60
15	SD	4.04	15.85	19.21	12.50	3.41	4.70	2.19
15	n	4	3	4	4	4	4	4
20	Mean	20.00	24.40	25.32	23.20	21.50	12.94	18.25
20	SD	2.80	5.56	5.09	5.87	5.72	1.59	4.74
20	n	4	3	4	4	4	4	4
25	Mean	24.40	29.07	18.95	28.53	13.12	13.87	37.67
25	SD	4.64	5.56	11.74	12.21	7.87	2.06	23.20
25	n	4	3	4	4	4	4	3
30	Mean	16.85	25.00	24.15	22.48	13.68	17.21	20.53
30	SD	2.92	9.94	7.76	5.22	1.86	8.76	8.15
30	n	4	4	4	4	4	4	4
33	Mean	14.00	16.30	19.23	20.82	21.98	14.16	15.86
33	SD	6.67	7.42	10.09	2.11	11.63	3.98	3.96
33	n	4	4	4	4	4	4	4
37	Mean	20.50	18.45	24.25	22.80	20.82	13.91	22.43
37	SD	8.97	4.22	5.92	8.07	7.52	1.47	6.12
37	n	4	4	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 non-linear and differ by station (Figure 3-1). Most mean trends are estimated to be above 20 $\mu\text{g/L}$ (i.e., poor quality condition). No periods of significant temporal change were observed.

3.3.2 Dissolved Oxygen (near bottom)

There is slightly increasing temporal trend in bottom dissolved oxygen concentrations for most trend stations, although no periods of statistically significant change were observed (Figure 3-2). The within-year variability is much higher for hexagon 25 compared to all other trend stations. The yearly average concentrations are all above the 5 mg/L threshold (i.e., good quality category).

3.3.3 PAR Attenuation K_d (SAV standard)

Temporal trends indicate that PAR attenuation has remained relatively constant in most hexagons except hexagon 10, which has a slight decreasing trend over time (Figure 3-3). No periods of significant temporal change were identified. The yearly average concentrations are all above the 1.61 m^{-1} threshold (i.e., poor quality category).

3.3.4 Salinity (near surface)

Temporal trends are consistent across all trend stations with salinity being relatively constant since 2013 (Figure 3-4). No periods of significant temporal change were identified despite a slight decrease observed during the period of 2009 to 2013.

3.3.5 Temperature (near surface)

Temporal trends in salinity are non-linear have been generally decreasing in all trend stations except for hexagon 33 which has remained constant (Figure 3-4). Periods of statistically significant decreases were observed in hexagons 15, 25, and 30.

3.3.6 Turbidity (near surface)

Temporal trends in turbidity vary by trend station and no periods of statistically significant change were observed (Figure 3-5). Within-year variability tends to be higher in hexagons 5, 10, 15, and 25 compared to the others.

Table 3-7. Approximate significance of the generalized additive model smooth terms for each Colonial National Historical Park 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 <i>a</i> ($\mu\text{g/L}$)	Surface	s(year, hex)	25.26 (55)	1.09	< 0.001
Dissolved oxygen (mg/L)	Bottom	s(year, hex)	18.87 (55)	0.98	< 0.001
PAR attenuation (K_d (m^{-1}))	Water Column	s(year, hex)	16.82 (55)	0.70	0.001
Salinity (psu)	Surface	s(year, hex)	42.90 (55)	17.02	< 0.001
Temperature ($^{\circ}\text{C}$)	Surface	s(year, hex)	8.59 (55)	0.25	0.049
Turbidity (NTU)	Surface	s(year, hex)	18.53 (55)	1.03	< 0.001

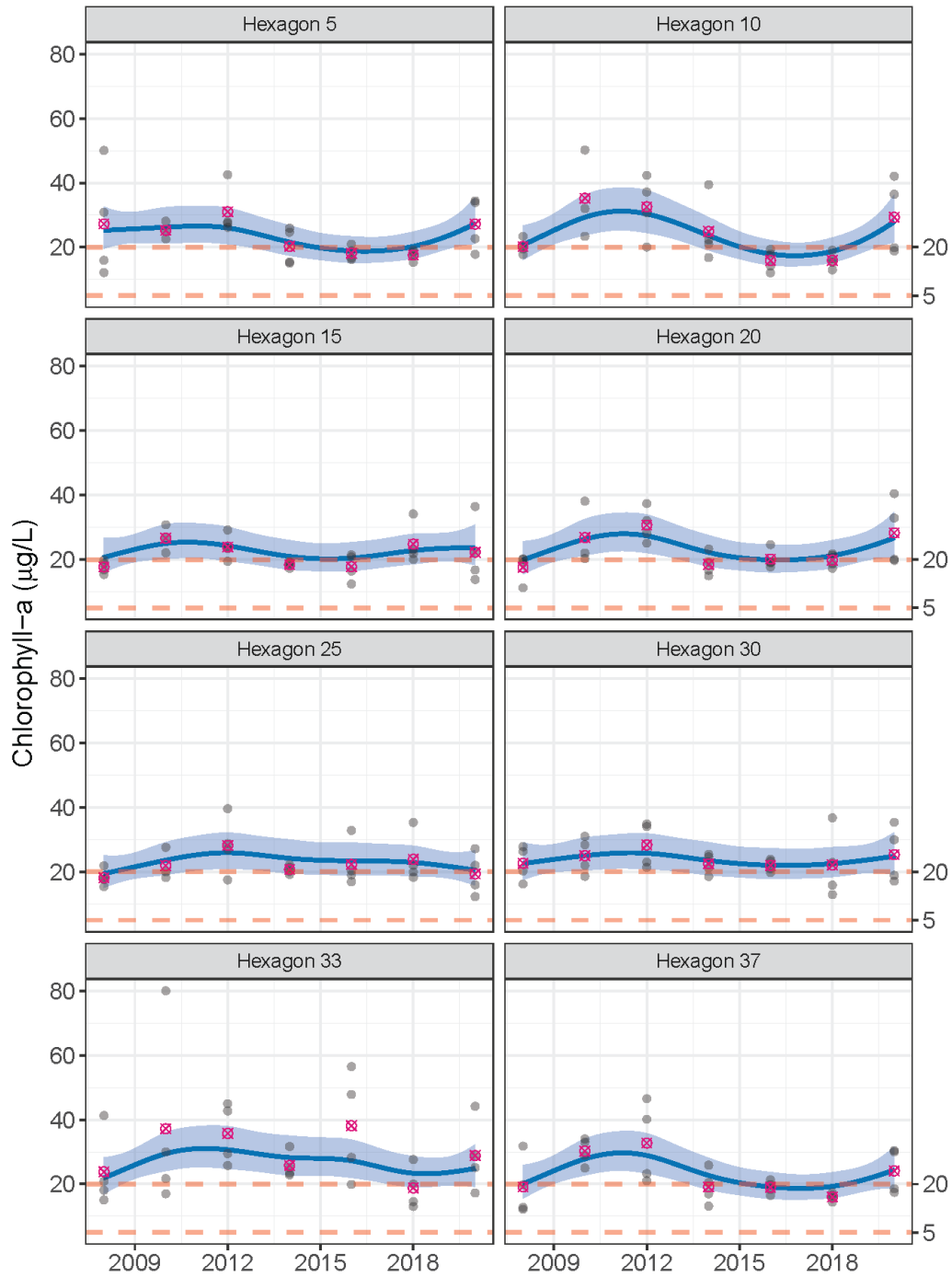


Figure 3-1. Mean surface chlorophyll-a concentration ($\mu\text{g/L}$) trend at each Colonial National Historical Park trend station based on estimated generalized additive model smooths. The fitted smooth is plotted in dark blue with shaded 95% confidence interval. Periods of significant temporal change (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) plotted as red dashed lines.

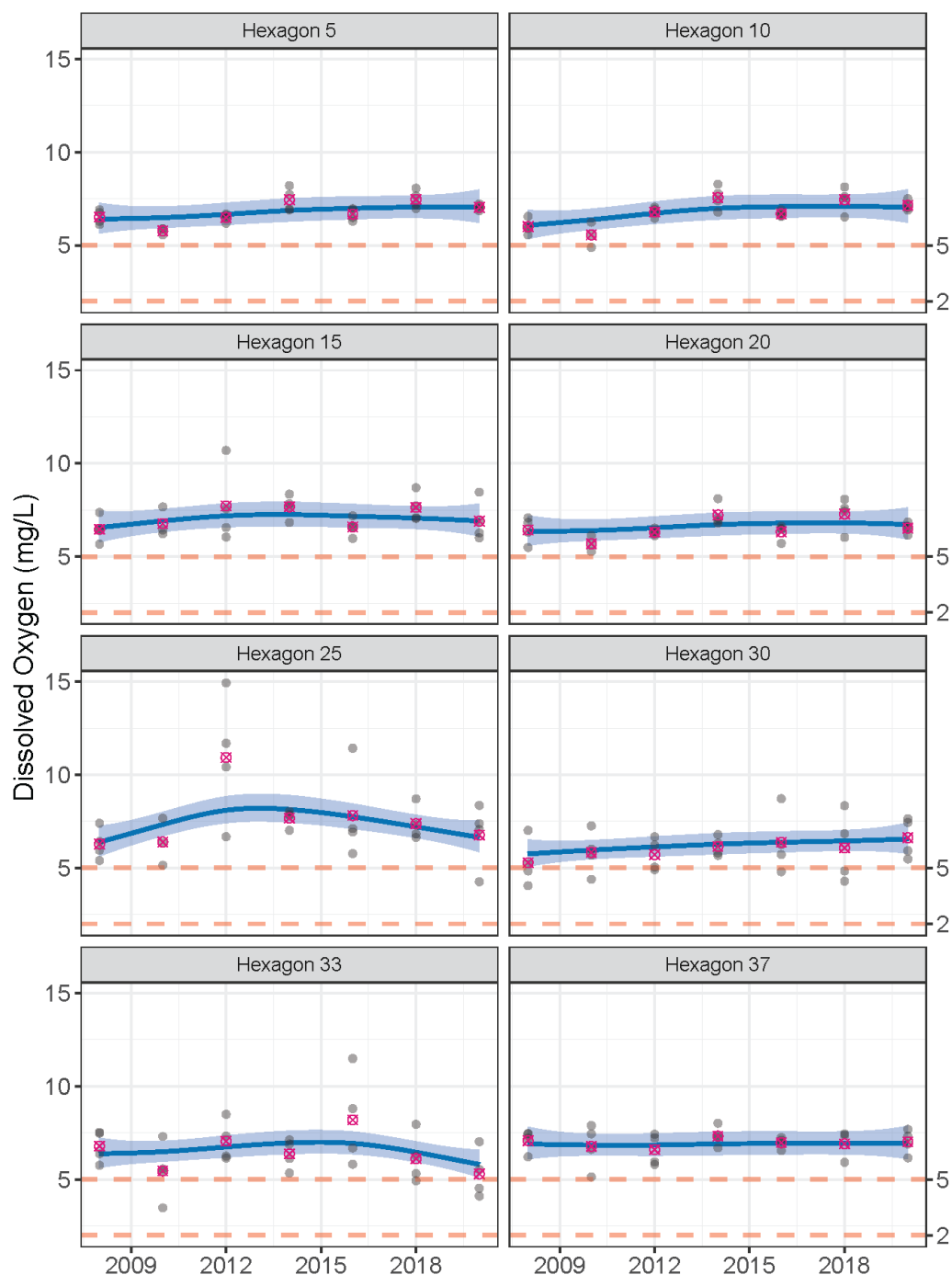


Figure 3-2. Mean bottom dissolved oxygen concentration (mg/L) trend at each Colonial National Historical Park trend station based on estimated generalized additive model smooths. The fitted smooth is plotted in dark blue with shaded 95% confidence interval. Periods of significant temporal change (increases/decreases), 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) plotted as red dashed lines.

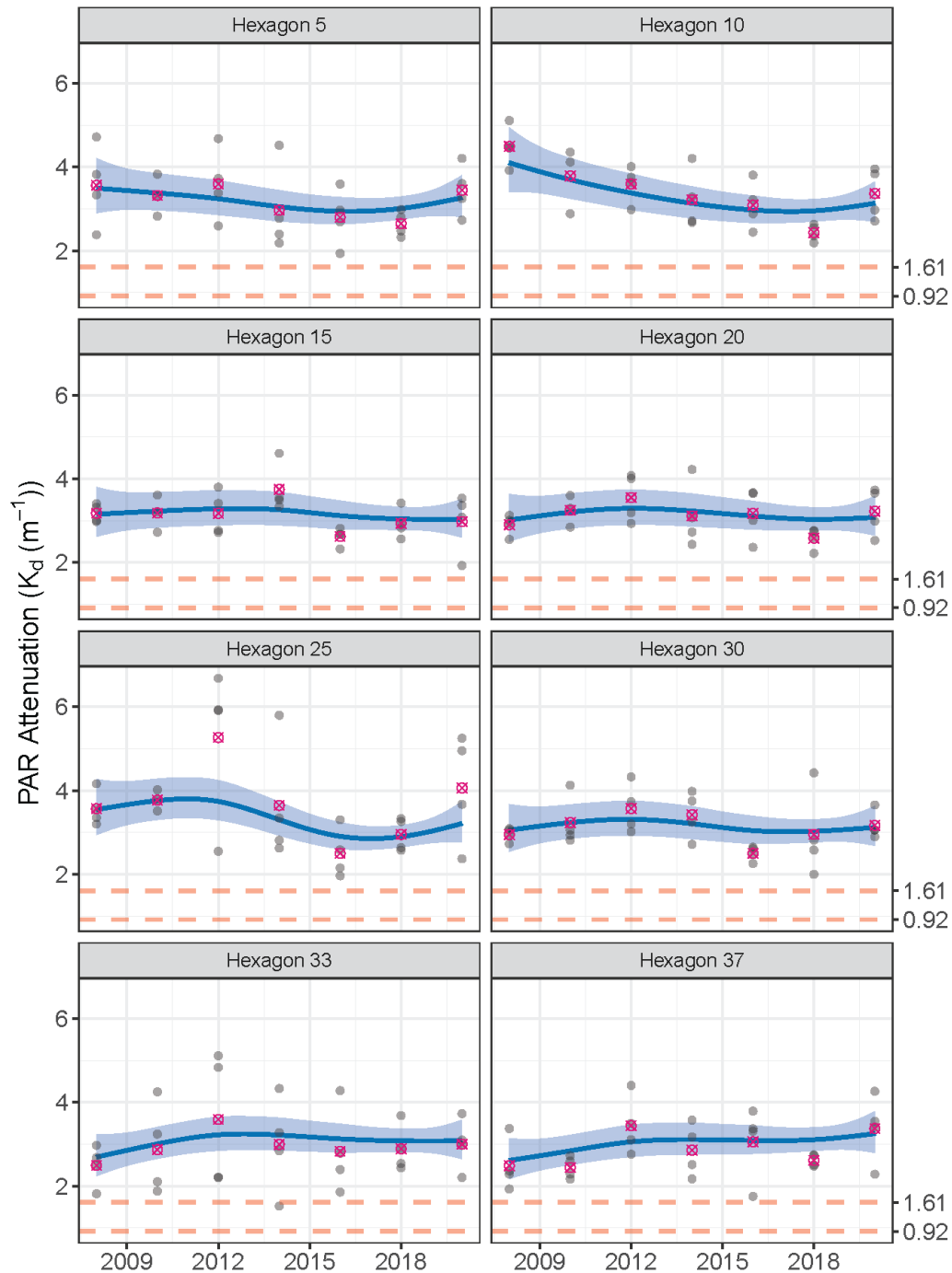


Figure 3-3. Mean PAR attenuation (K_d (m^{-1})) Mean PAR attenuation trend at each Colonial National Historical Park trend station based on estimated generalized additive model smooths. The fitted smooth is plotted in dark blue with shaded 95% confidence interval. Periods of significant temporal change (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) plotted as red dashed lines.

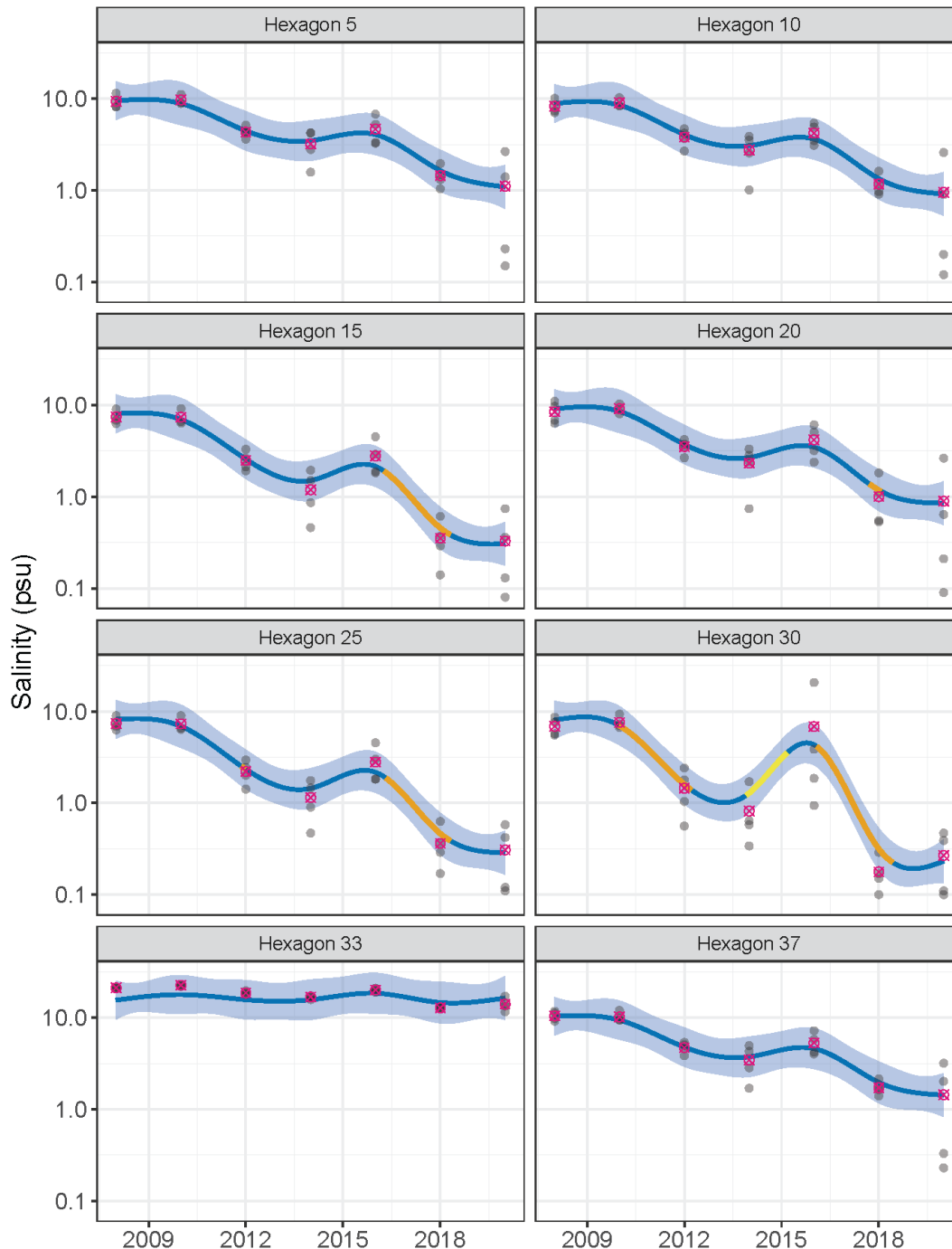


Figure 3-4. Mean surface salinity (psu) trend at each Colonial National Historical Park trend station based on estimated generalized additive model smooths. The fitted smooth is plotted in dark blue with shaded 95% confidence interval. Periods of significant temporal change (increases/decreases), if 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 $\log_{10}$ scale to better visualize the estimated trend

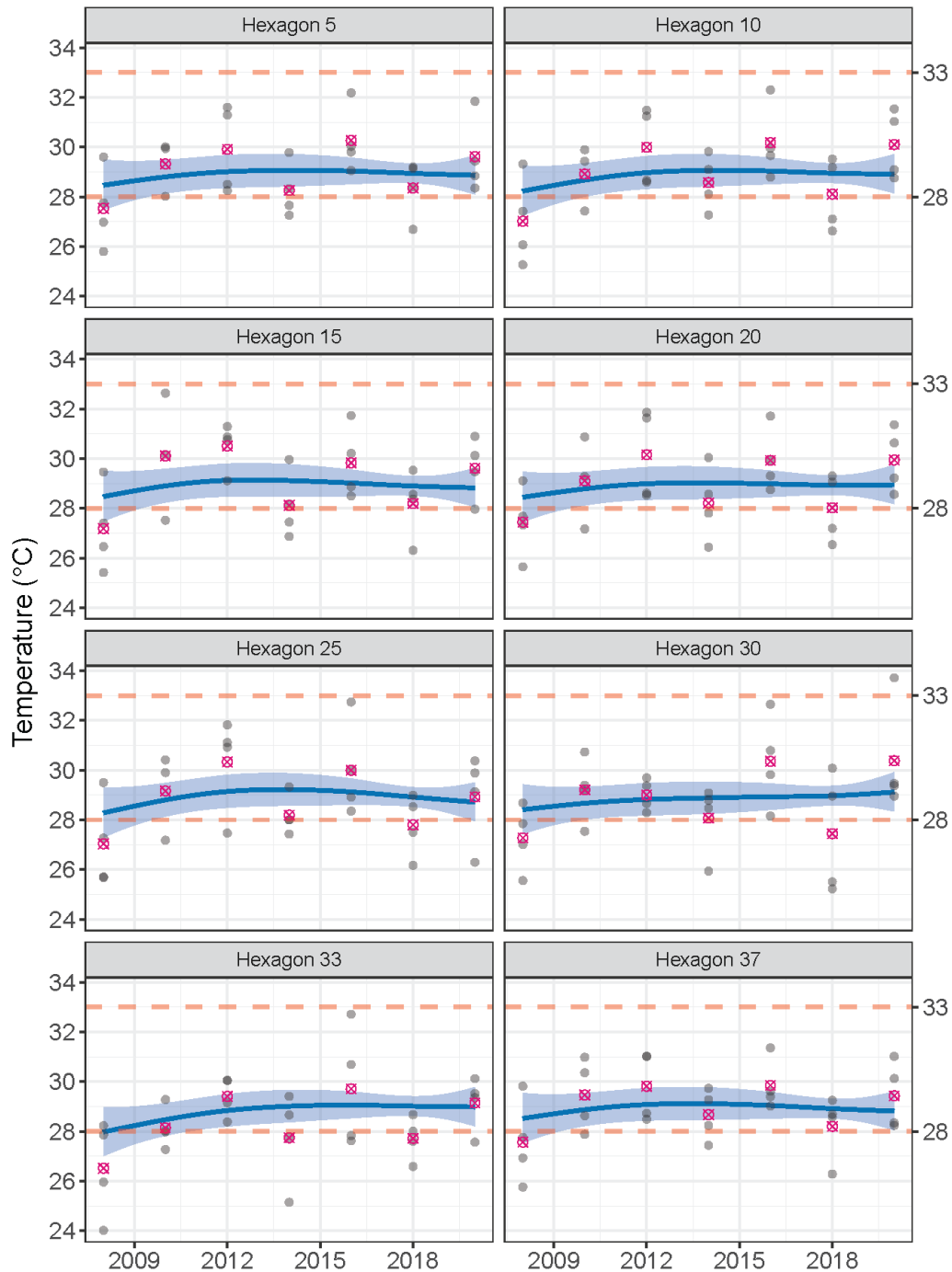


Figure 3-5. Mean surface temperature (°C) trend at each Colonial National Historical Park trend station based on estimated generalized additive model smooths. The fitted smooth is plotted in dark blue with shaded 95% confidence interval. Periods of significant temporal change (increases/decreases), if 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) plotted as red dashed lines.

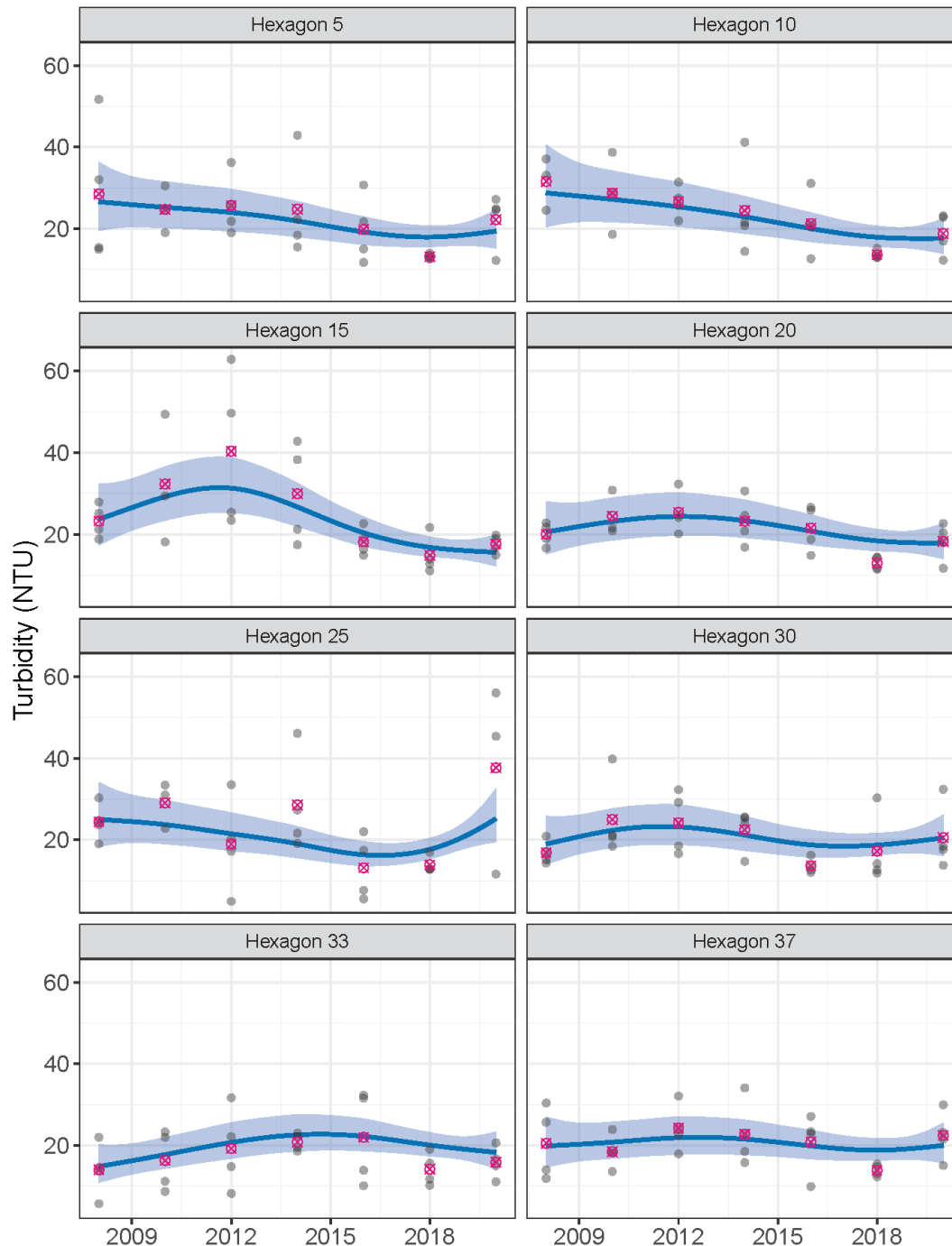


Figure 3-6. Mean surface turbidity (NTU) trend at each Colonial National Historical Park trend station based on estimated generalized additive model smooths. The fitted smooth is plotted in dark blue with shaded 95% confidence interval. Periods of significant temporal change (increases/decreases), if present, are plotted in yellow/orange. Within each year, individual measurements and the average are plotted in gray and magenta, respectively.

3.4 Multivariate Mann-Kendall Trend Test Results

The multivariate Mann-Kendall test results that there is statistical evidence of an increasing monotonic trend in bottom dissolved oxygen (p -value = 0.066) and a statistically significant decreasing monotonic trend in salinity (p -value = 0.032) at the trend stations during the period of 2008 to 2020 (Table 3-8).

Table 3-8. Summary of multivariate Mann-Kendall test for monotonic trend in each measure for Colonial National Historical Park trend stations: 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	0.11	0.911
Dissolved oxygen (mg/L)	Bottom	1.84	0.066
PAR attenuation (K_d (m^{-1}))	Water Column	0.00	1.000
Salinity (psu)	Surface	-2.14	0.032
Temperature ($^{\circ}\text{C}$)	Surface	1.09	0.277
Turbidity (NTU)	Surface	-0.99	0.322

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.

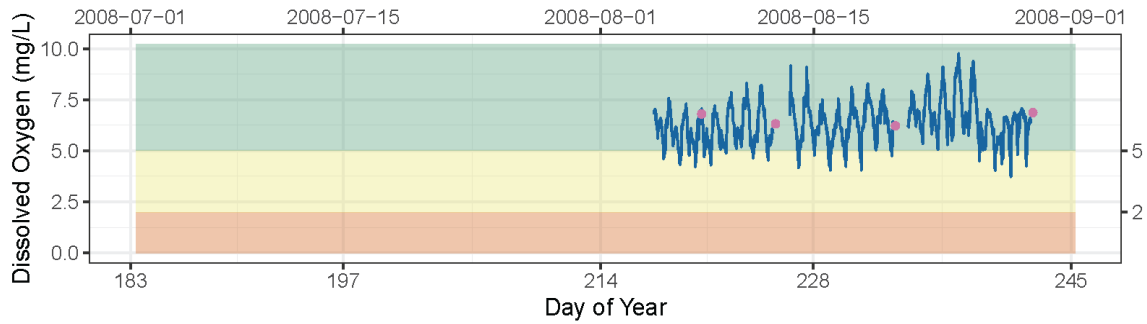


Figure 4-1. Continuous dissolved oxygen readings from sonde deployed at the Colonial National Historical Park logging station in 2008. Magenta points represent discrete grab samples collected every week for quality assurance. Estuarine quality category thresholds (Table 2-1) are delineated by the background 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 `mcmc` 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 Thoroughfare Stratum 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 most years and are above 20 $\mu\text{g/L}$ for all years except 2018 (Figure 4-2). The distribution of chlorophyll-a measurements was highly right skewed during 2016 (Figure 4-3). Chlorophyll-a condition during the day is generally poor with some fair conditions observed during the day (Figure 4-8).

4.3.2 Dissolved Oxygen

Intra-annual mean trends in dissolved oxygen are non-linear and are above 5 mg/L for all years (Figure 4-2). The distribution of measurements often partially extends below 5 mg/L threshold, particularly in 2012 (Figure 4-3). 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

Intra-annual mean trends in PAR attenuation tend to be decreasing during the month of July and relatively constant during August (Figure 4-2). The distribution of measurements tend to be long tailed (Figure 4-3). Across all years, PAR attenuation conditions for a majority of each day were in poor condition (Figure 4-9).

4.3.4 pH

Measurements of pH are somewhat periodic with measurements in 2020 having both a higher amplitude and lower frequency (Figure 4-6). The distribution of measurements in 2020 are more right skewed compared to 2018 (Figure 4-7).

4.3.5 Salinity

Intra-annual mean trends in salinity tend to be relatively constant during July and increasing during August (Figure 4-4). Salinity measurements within each year tend to be either long tailed or bimodal (Figure 4-5).

4.3.6 Temperature

Yearly mean surface water temperature at the logging station was above 28 °C for six out the eight years sampled (Table 4-1). Intra-annual mean trends in water temperature are highly non-linear with minimal difference between July and August (Figure 4-4). Water temperature measurements within each year typically span approximately 6 °C (Figure 4-5).

4.3.7 Turbidity

Intra-annual mean trends in turbidity were relatively smoother during July compared to August (Figure 4-4). Turbidity measurements in 2010 were highly variable and more left-skewed than other years (Figure 4-5).

Table 4-1. Yearly summary statistics for measures collected at the Colonial National Historical Park Thoroughfare logging station. [SD = standard deviation]

Measure	Statistic	Sampling Year							
		2003	2008	2010	2012	2014	2016	2018	2020
Chlorophyll <i>a</i> (µg/L)	Mean	38.61	28.09	22.57	27.62	23.88	22.11	17.82	27.73
Chlorophyll <i>a</i> (µg/L)	SD	9.77	7.89	6.01	4.28	4.82	5.06	4.03	6.46
Chlorophyll <i>a</i> (µg/L)	n	1713	2183	1376	1160	2592	2580	2202	2015
Depth (m)	Mean	0.81	0.82	1.02	1.11	1.06	0.82	1.28	0.78
Depth (m)	SD	0.26	0.23	0.27	0.27	0.28	0.27	0.26	0.29
Depth (m)	n	2619	2187	3088	2102	2592	2584	2203	2015
dissolved oxygen (mg/L)	Mean	7.25	6.34	6.15	5.94	7.07	6.54	7.33	6.62
dissolved oxygen (mg/L)	SD	0.74	1.01	1.06	1.11	0.90	1.17	0.90	1.05
dissolved oxygen (mg/L)	n	2619	2187	2053	2103	2592	2584	2203	2015
PAR attenuation (K_d (m ⁻¹))	Mean	3.42	3.07	2.98	2.64	2.52	2.75	2.47	3.45
PAR attenuation (K_d (m ⁻¹))	SD	1.01	0.52	0.58	0.71	0.52	0.52	0.49	0.71
PAR attenuation (K_d (m ⁻¹))	n	558	443	767	538	644	558	564	376
pH	Mean	NA	NA	NA	NA	NA	NA	7.62	7.47
pH	SD	NA	NA	NA	NA	NA	NA	0.26	0.28
pH	n	0	0	0	0	0	0	2203	2015
Salinity (psu)	Mean	0.66	8.81	8.50	3.90	2.40	3.98	1.35	1.03
Salinity (psu)	SD	0.60	1.08	0.81	0.62	0.91	1.19	0.35	0.82
Salinity (psu)	n	2619	2187	3088	2102	2589	2584	2203	2015
Temperature (°C)	Mean	28.54	27.28	28.97	30.11	27.90	30.00	28.20	29.05
Temperature (°C)	SD	0.83	1.52	1.25	1.33	1.19	1.47	1.29	1.55
Temperature (°C)	n	2619	2187	3088	2103	2592	2584	2203	2015
Turbidity (NTU)	Mean	38.06	35.98	41.96	39.03	36.92	33.85	19.89	22.48
Turbidity (NTU)	SD	14.25	15.13	25.90	9.36	9.84	8.36	6.18	6.51
Turbidity (NTU)	n	2619	2181	2895	2072	2462	2581	2194	2013

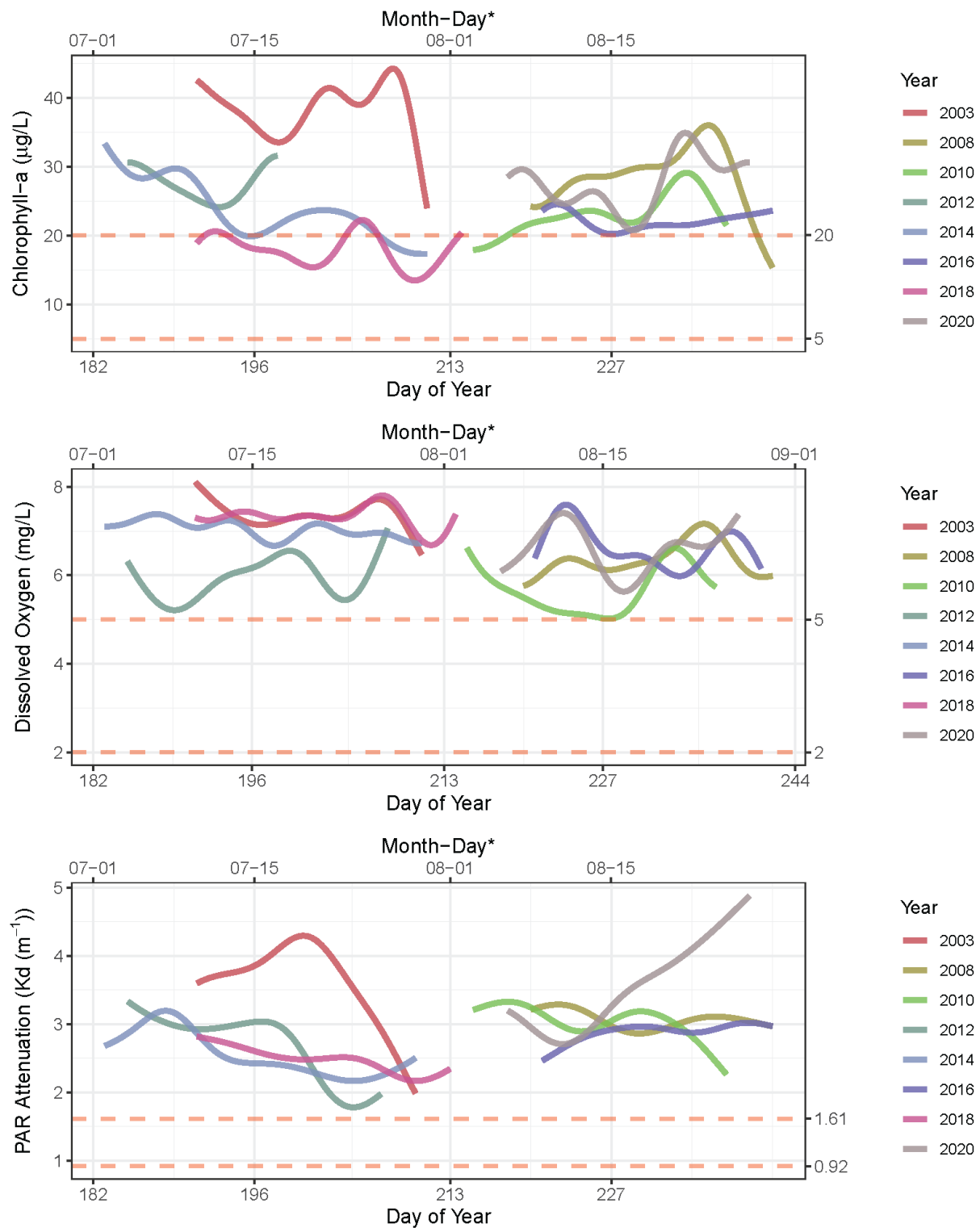


Figure 4-2. Intra-annual trends in surface chlorophyll-a concentration (top), bottom dissolved oxygen concentration (middle), and PAR attenuation (bottom) at the Colonial National Historical Park Thoroughfare logging station based on generalized additive model smooths. Estuarine quality category thresholds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

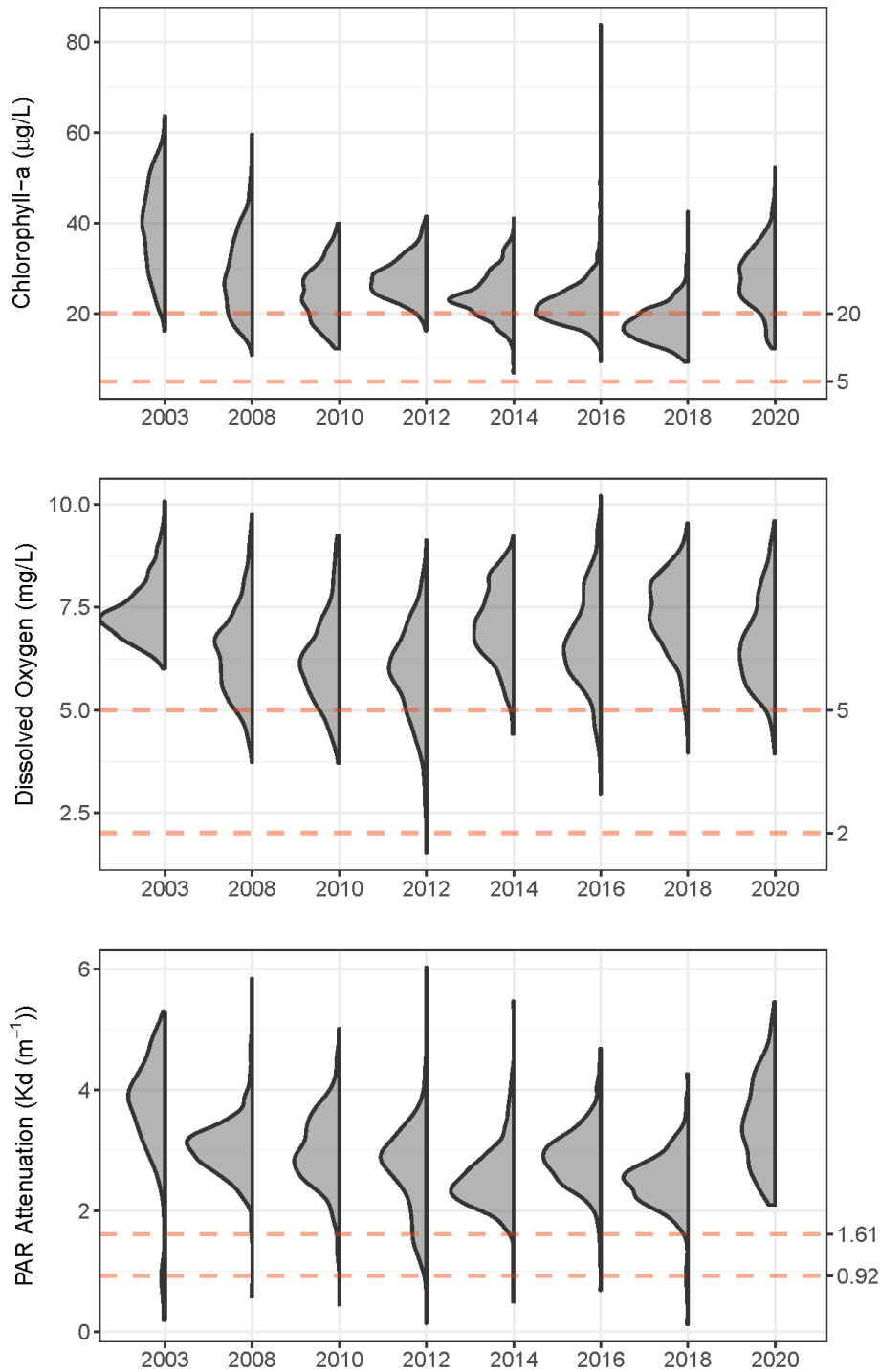


Figure 4-3. Distribution of index period time series values for surface chlorophyll-*a* concentration (top), bottom dissolved oxygen concentration (middle), and PAR attenuation (bottom) at the Colonial National Historical Park Thoroughfare logging station. Estuarine quality category thresholds (Table 2-1) are plotted as red dashed lines. Note: the surface chlorophyll-*a* concentration y-axis is transformed to the log₁₀ scale to better visualize the estimated distributions.

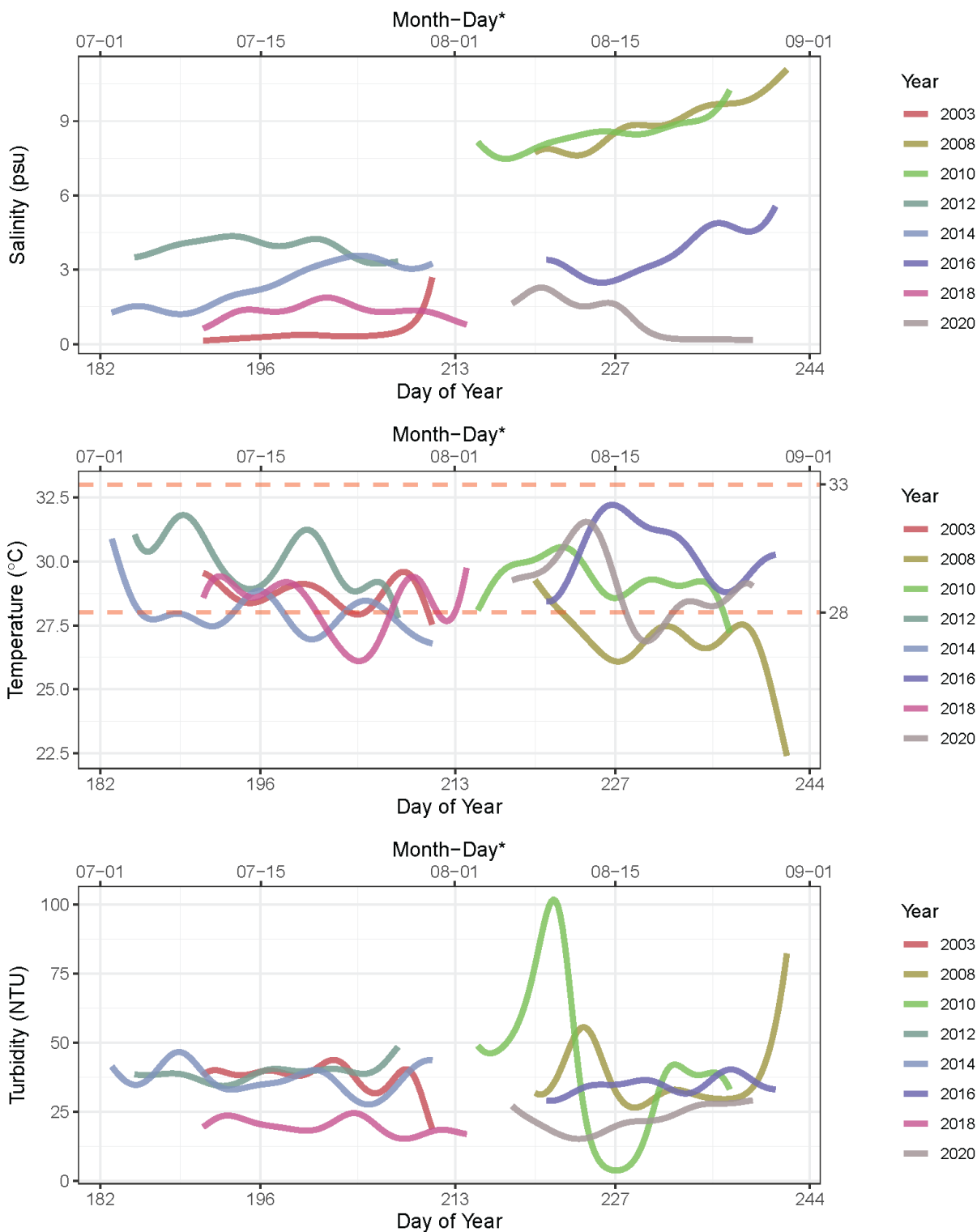


Figure 4-4. Intra-annual trends in surface salinity (top), surface temperature (middle), and surface turbidity (bottom) at Colonial National Historical Park Thoroughfare logging station based on generalized additive model smooths. Estuarine quality category thresholds (Table2-1) are plotted as red dashed lines.

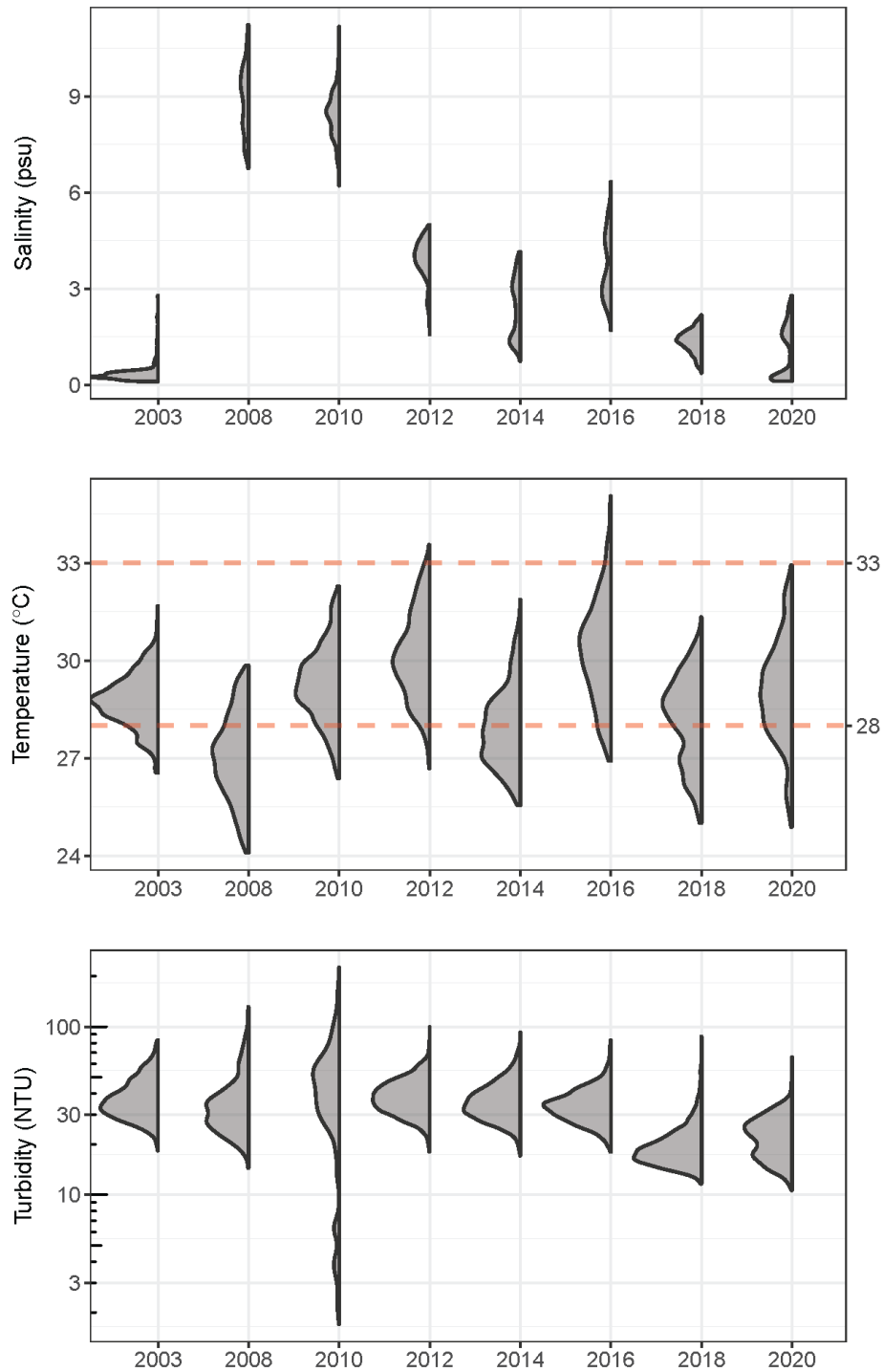


Figure 4-5. Distribution of index period time series values for surface salinity (top), surface temperature (middle), and surface turbidity (bottom) at the Colonial National Historical Park Thoroughfare logging station. Note: the surface turbidity y-axis is transformed to the log₁₀ scale to better visualize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

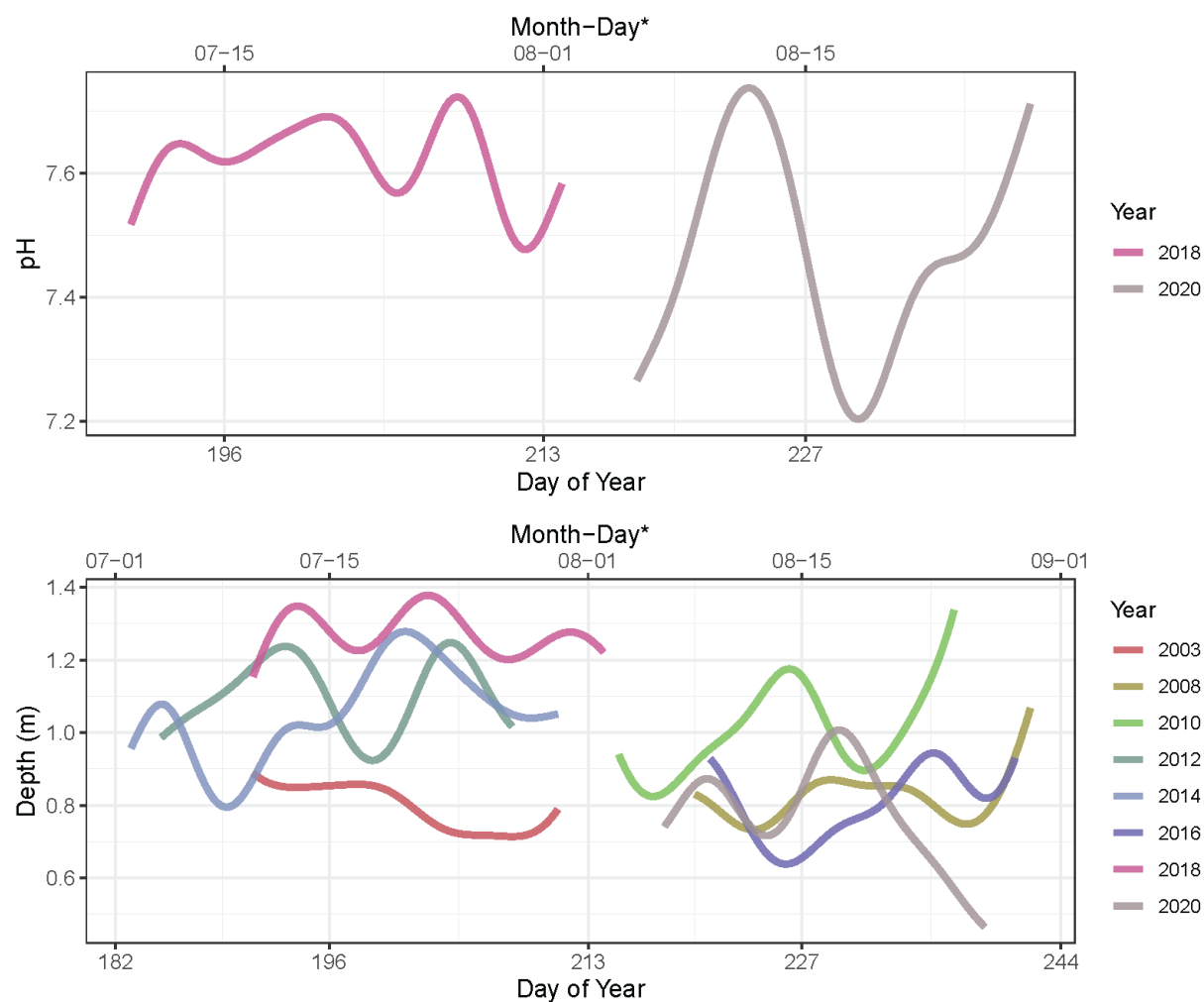


Figure 4-6. Intra-annual trends in pH (top) and depth (bottom) at the Colonial National Historical Park Thoroughfare logging station based on generalized additive model smooths.

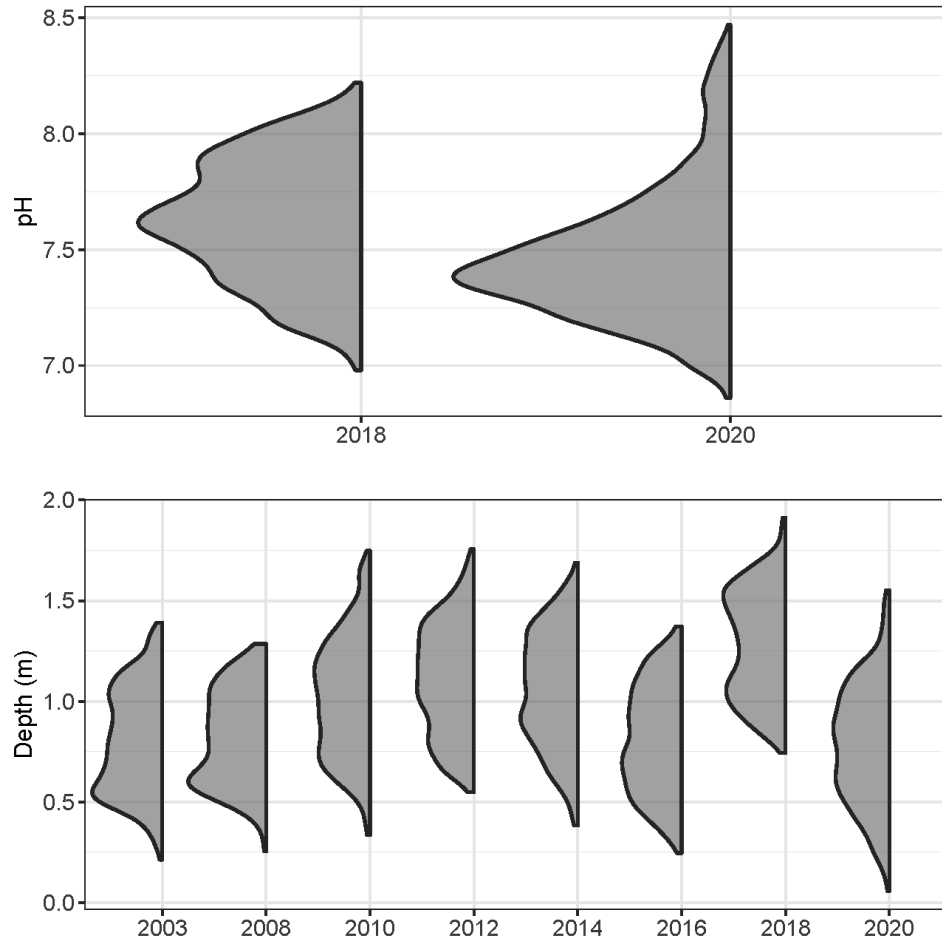


Figure 4-7. Distribution of index period time series values for pH (top) and depth (bottom) at the Colonial National Historical Park Thoroughfare logging station.

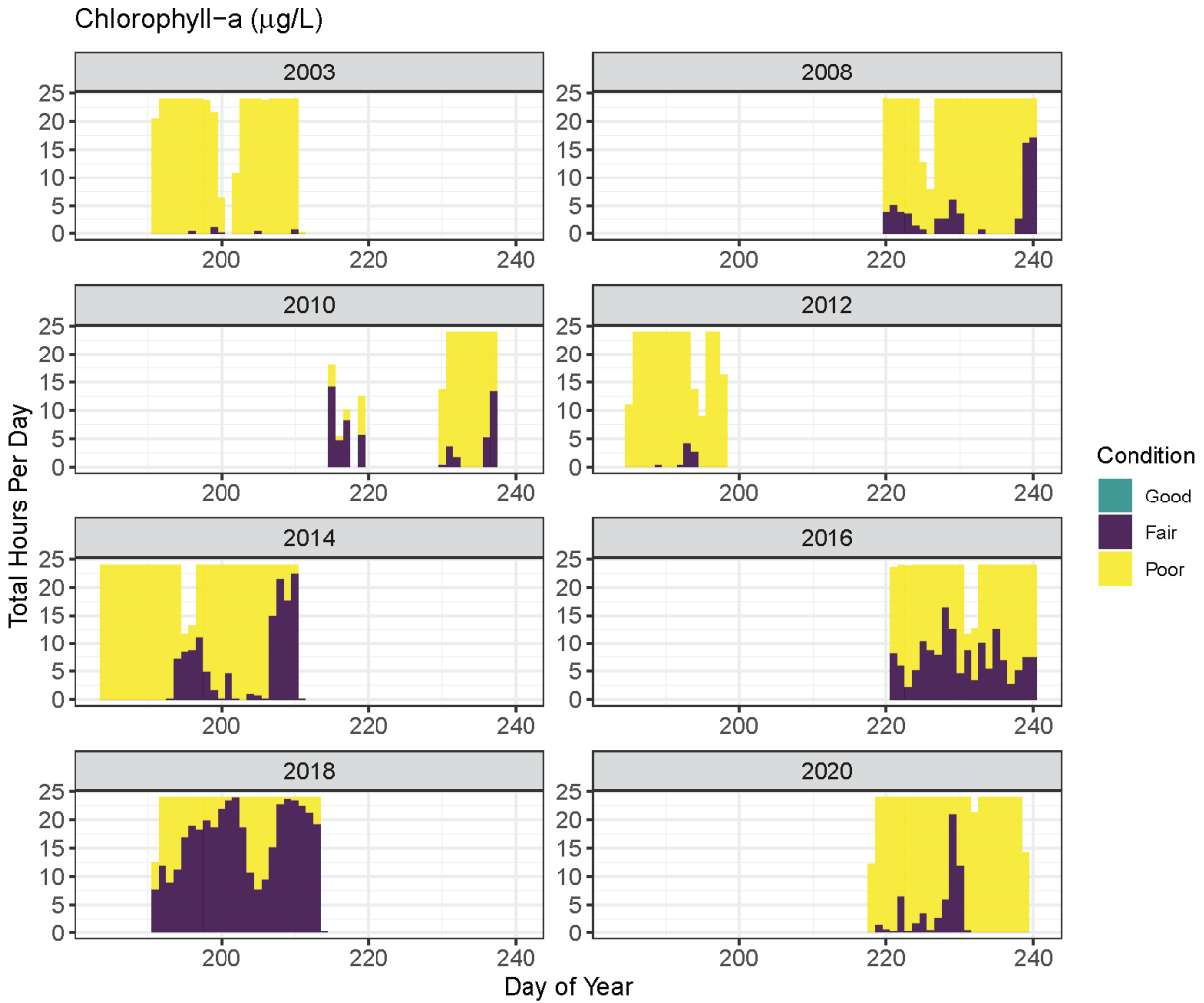


Figure 4-8. Total hours per day that surface chlorophyll-a ($\mu\text{g/L}$) concentration was observed within each condition category at the Colonial National Historical Park Thoroughfare logging station.

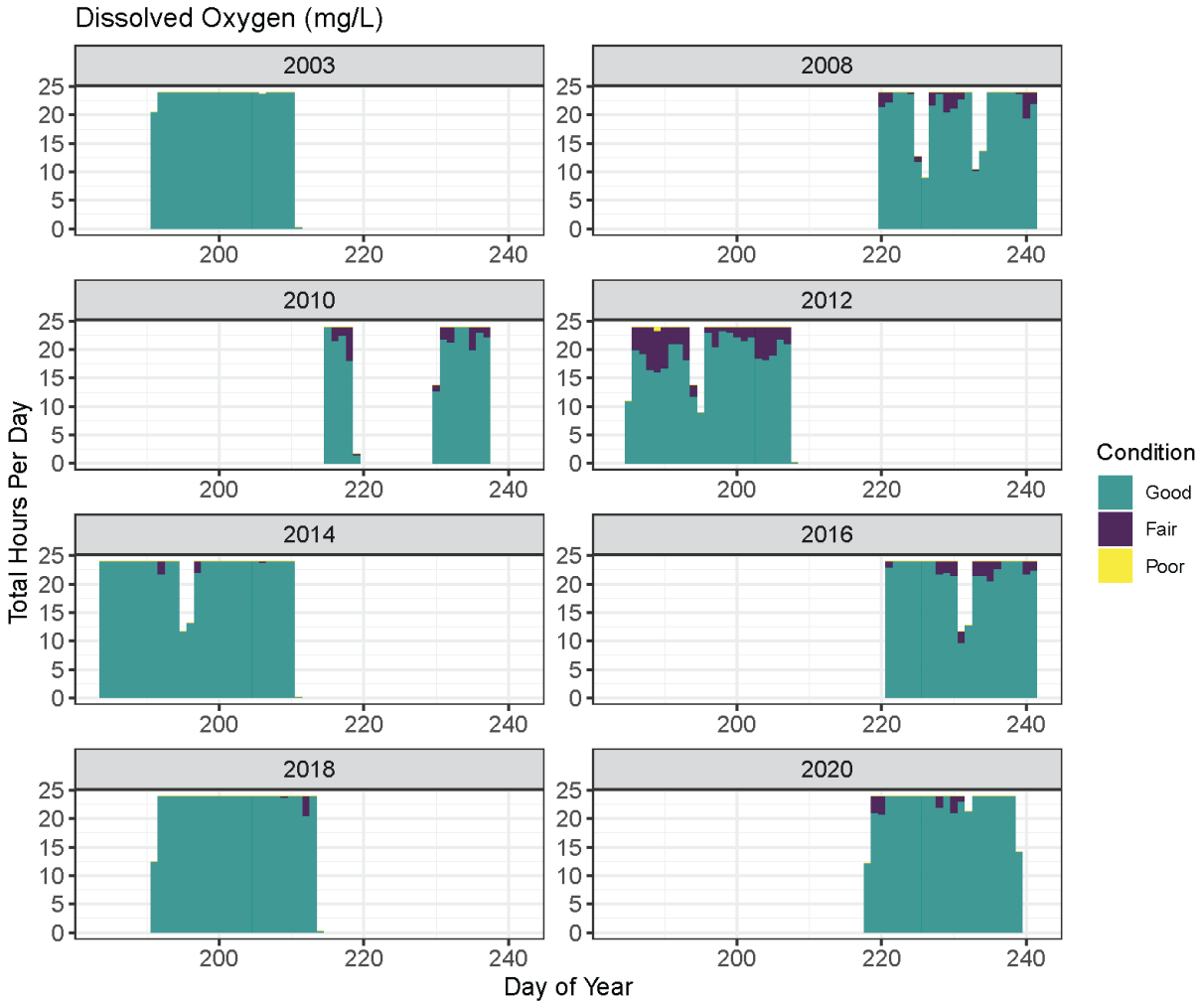


Figure 4-9. Total hours per day that bottom dissolved oxygen concentration (mg/L) was observed within each condition category at the Colonial National Historical Park Thoroughfare logging station.

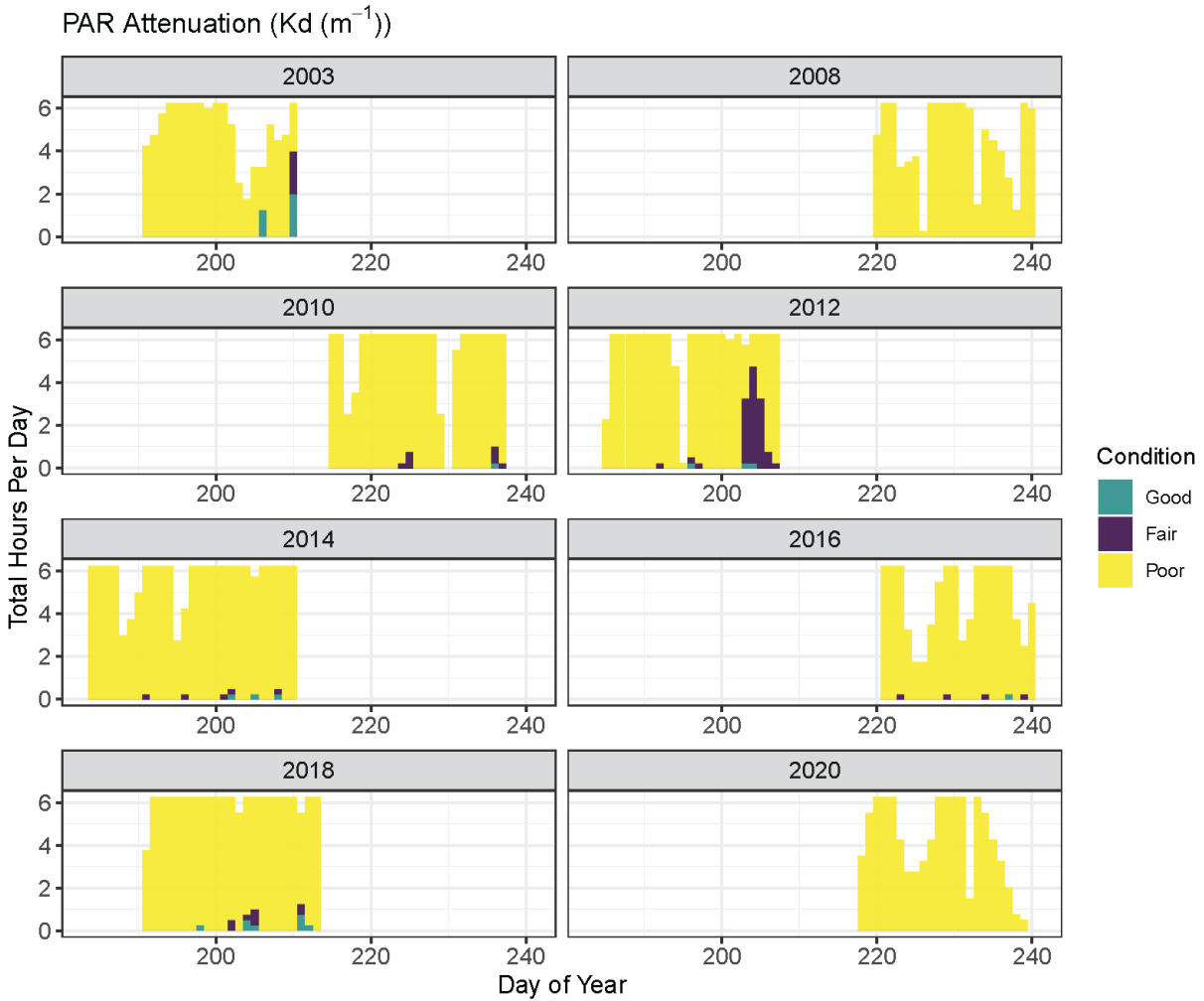


Figure 4-10. Total hours per day that PAR attenuation (K_d (m^{-1})) was observed within each condition category at the Colonial National Historical Park Thoroughfare logging station.

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 framework⁶ 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

⁶ 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 6-1 which corresponds to an approximate 5.1% increase in the mean each year ($\exp(0.05) \approx 1.051$).

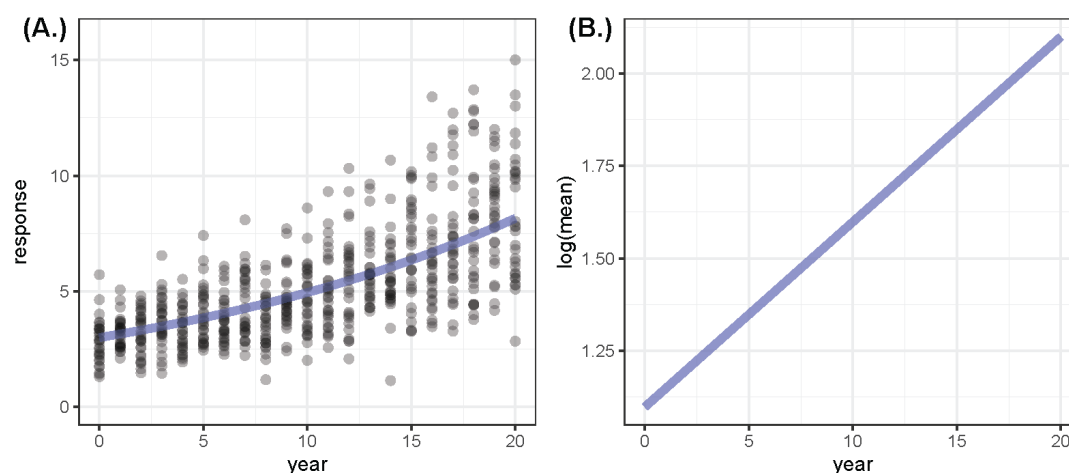


Figure 5-1. (A.) Process model realizations (black dots) representing samples from 30 sites within a hypothetical estuary. The i th hexagon's response value is simulated from $\text{Gamma}(\mu_i, 0.1)$, where $\ln(\mu_i) = 3 + 0.05 \cdot \text{year}_i$. The blue curve represents the true underlying mean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

Measure	Tidal Creek	Thoroughfare	All Strata
Chlorophyll <i>a</i> ($\mu\text{g/L}$)	0.069	0.157	0.085
Dissolved oxygen (mg/L)	0.024	0.048	0.031
PAR attenuation (K_d (m^{-1}))	0.062	0.072	0.067

5.2 Simulation Results

5.2.1 Tidal Creek Stratum

Estimated power for detecting a linear trend within the Tidal Creek Stratum 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).

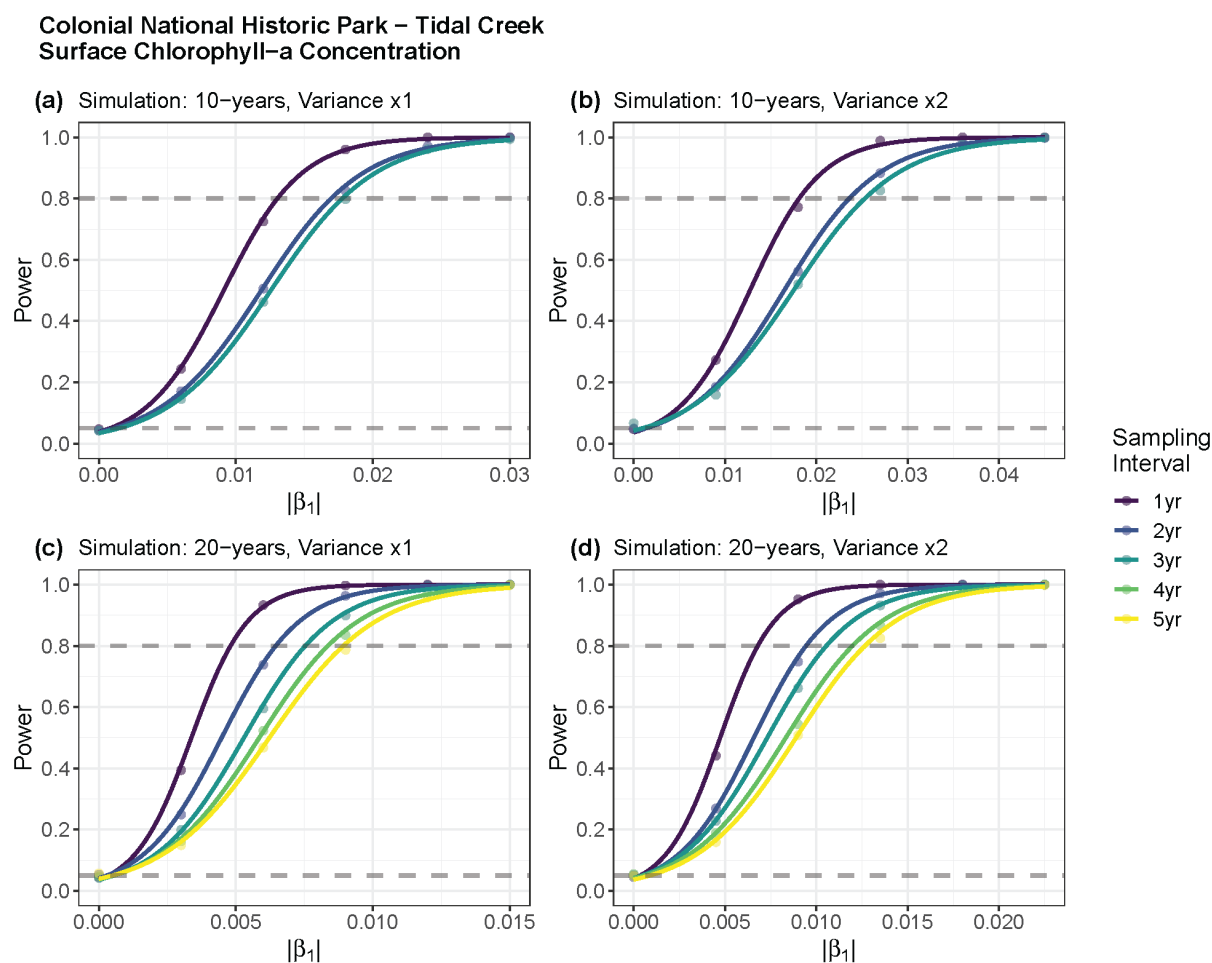


Figure 5-2. Estimated power to detect a linear trend in mean surface chlorophyll-*a* concentration within the Tidal Creek stratum of Colonial National Historical Park. Power estimates vary based on the assumed sampling interval (see legend) and duration: (a) 10-year period, (c) 20-year period. Power 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 the Tidal Creek stratum of Colonial National Historical Park. 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 2.2.2 is provided in the table footnotes.

Period	Variance	Threshold	β_1	Power by Sampling Interval				
				1yr	2yr	3yr	4yr	5yr
10 yrs	×1	5	−0.138	1.00	1.00	1.00	NA	NA
10 yrs	×1	20	0.000	0.04	0.04	0.04	NA	NA
10 yrs	×2	5	−0.138	1.00	1.00	1.00	NA	NA
10 yrs	×2	20	0.000	0.04	0.04	0.05	NA	NA
20 yrs	×1	5	−0.069	1.00	1.00	1.00	1.00	1.00
20 yrs	×1	20	0.000	0.04	0.05	0.04	0.04	0.04
20 yrs	×2	5	−0.069	1.00	1.00	1.00	1.00	1.00
20 yrs	×2	20	0.000	0.05	0.04	0.05	0.04	0.04

2018 Estimate: 19.92 $\mu\text{g/L}$

**Colonial National Historic Park – Tidal Creek
Bottom Dissolved Oxygen Concentration**

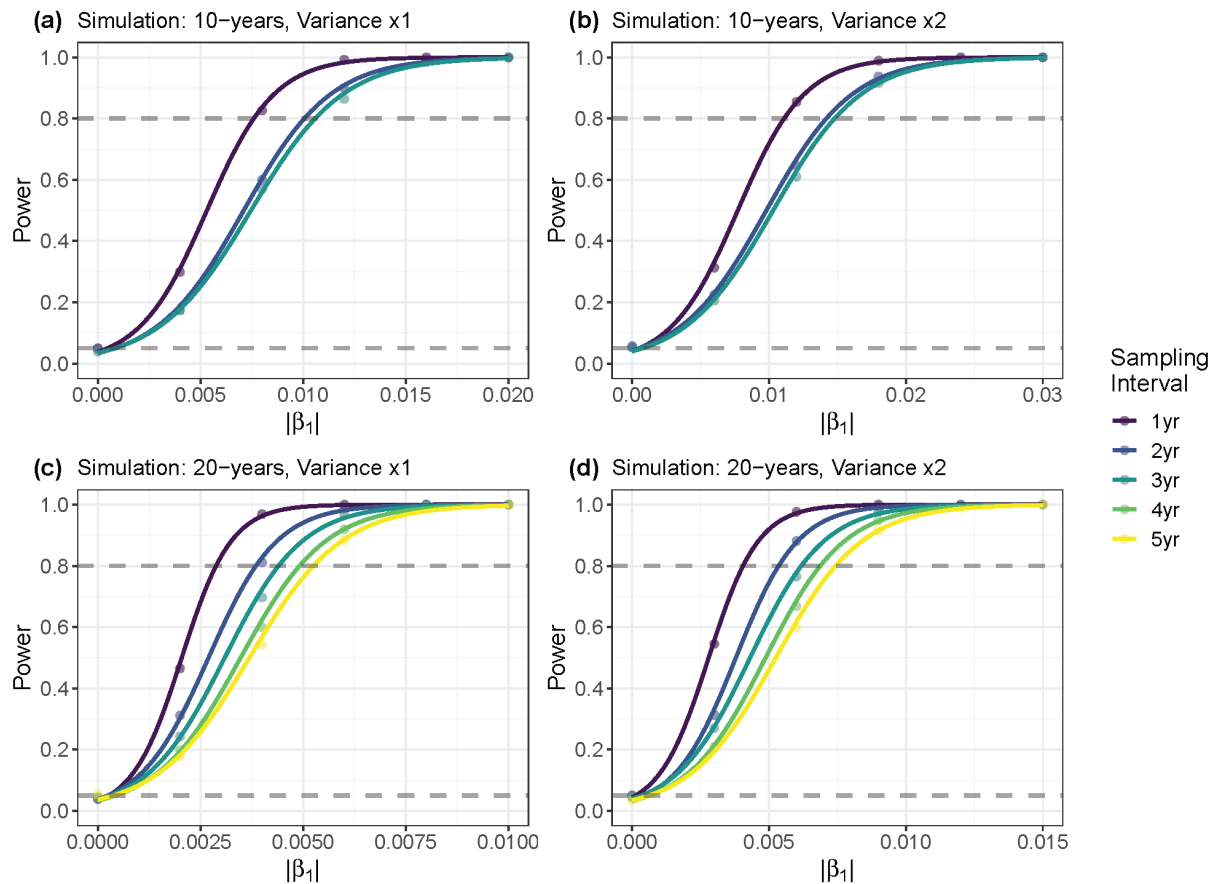


Figure 5-3. Estimated power to detect a linear trend in mean bottom dissolved oxygen concentration within the Tidal Creek stratum of Colonial National Historical Park. Power estimates vary based on the assumed sampling interval (see legend) and duration: (a) 10-year period, (c) 20-year period. Power 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 the Tidal Creek stratum of Colonial National Historical Park. 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 2.2.2 is provided in the table footnotes

Period	Variance	Threshold	β_1	Power by Sampling Interval				
				1yr	2yr	3yr	4yr	5yr
10 yrs	×1	2	−0.128	1.00	1.00	1.00	NA	NA
10 yrs	×1	5	−0.037	1.00	1.00	1.00	NA	NA
10 yrs	×2	2	−0.128	1.00	1.00	1.00	NA	NA
10 yrs	×2	5	−0.037	1.00	1.00	1.00	NA	NA
20 yrs	×1	2	−0.064	1.00	1.00	1.00	1.00	1.00
20 yrs	×1	5	−0.018	1.00	1.00	1.00	1.00	1.00
20 yrs	×2	2	−0.064	1.00	1.00	1.00	1.00	1.00
20 yrs	×2	5	−0.018	1.00	1.00	1.00	1.00	1.00

2018 Estimate: 7.22 mg/L

Colonial National Historic Park – Tidal Creek
PAR Attenuation

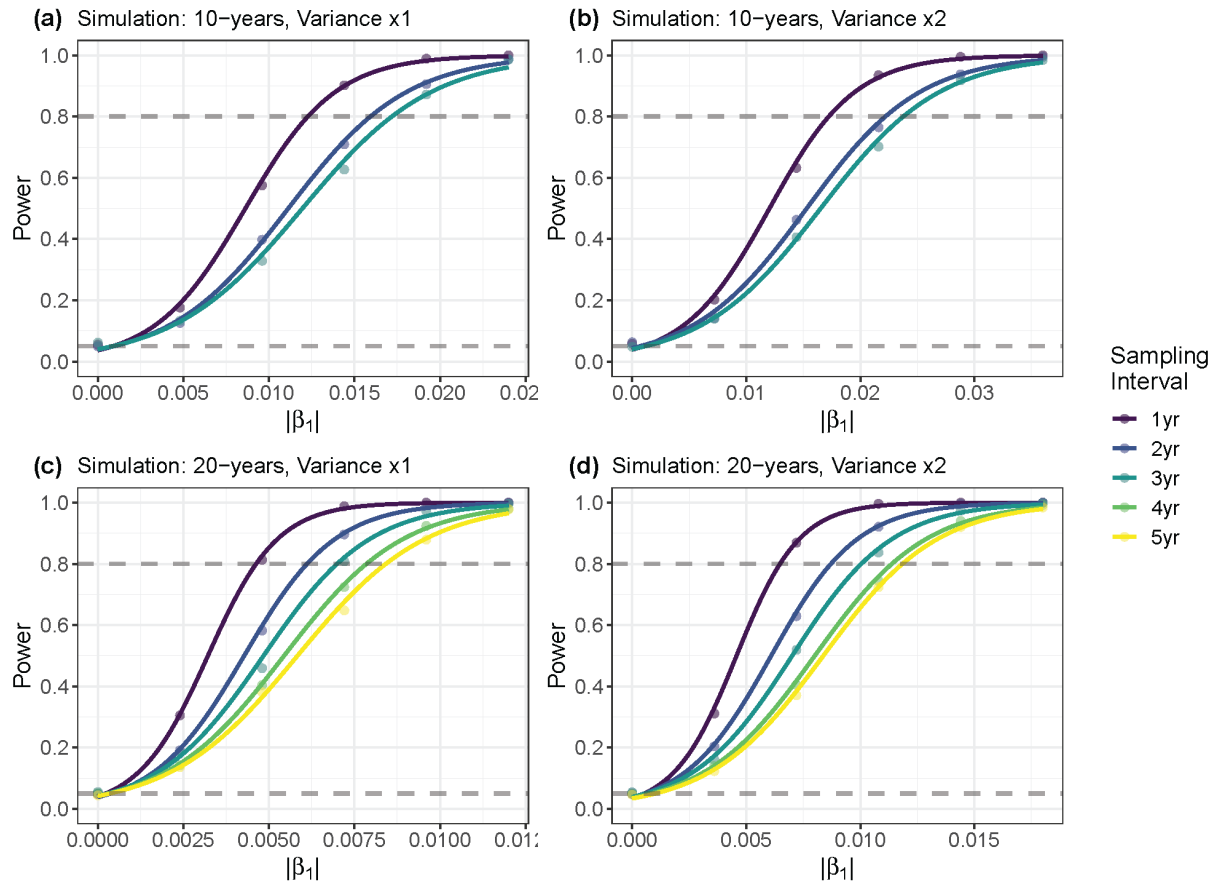


Figure 5-4. Estimated power to detect a linear trend in mean PAR attenuation within the Tidal Creek stratum of Colonial National Historical Park. Power estimates vary based on the assumed sampling interval (see legend) and duration: (a) 10-year period, (c) 20-year period. Power estimates in (b) and (d) 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 the Tidal Creek stratum of Colonial National Historical Park. 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 2.3.2 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	−0.111	1.00	1.00	1.00	NA	NA
10 yrs	×1	1.61	−0.055	1.00	1.00	1.00	NA	NA
10 yrs	×2	0.92	−0.111	1.00	1.00	1.00	NA	NA
10 yrs	×2	1.61	−0.055	1.00	1.00	1.00	NA	NA
20 yrs	×1	0.92	−0.055	1.00	1.00	1.00	1.00	1.00
20 yrs	×1	1.61	−0.027	1.00	1.00	1.00	1.00	1.00
20 yrs	×2	0.92	−0.055	1.00	1.00	1.00	1.00	1.00
20 yrs	×2	1.61	−0.027	1.00	1.00	1.00	1.00	1.00

2018 Estimate: 2.78 m^{−1}

5.2.2 Thoroughfare Stratum

Estimated power for detecting a linear trend within the Thoroughfare Stratum is summarized in both graphical and tabular format for surface chlorophyll-*a* concentration (Figure 5-5; Table 5-5), bottom dissolved oxygen concentration (Figure 5-6; Table 5-6), and PAR attenuation (Figure 5-7; Table 5-7).

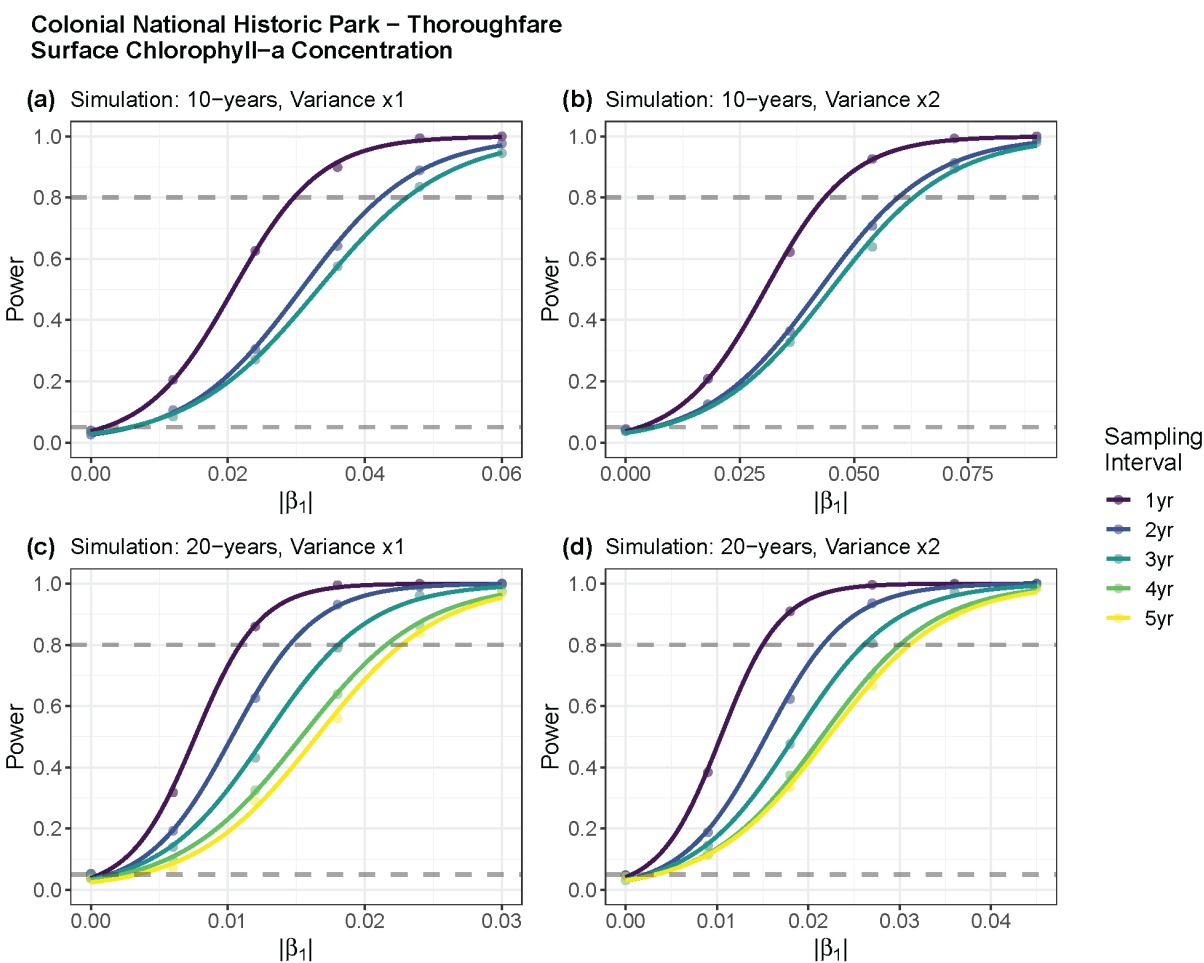


Figure 5-5. Estimated power to detect a linear trend in mean surface chlorophyll-*a* concentration within the Thoroughfare stratum of Colonial National Historical Park. Power estimates vary based on the assumed sampling interval (see legend) and duration: (a) 10-year period, (c) 20-year period. Power 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 5-5. Estimated power to detect a linear trend in mean surface chlorophyll-a concentration large enough to cross both quality thresholds within the Thoroughfare stratum of Colonial National Historical Park. 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 2.3.2 is provided in the table footnotes.

Period	Variance	Threshold	β_1	Power by Sampling Interval				
				1yr	2yr	3yr	4yr	5yr
10 yrs	×1	5	−0.159	1.00	1.00	1.00	NA	NA
10 yrs	×1	20	−0.020	0.48	0.22	0.20	NA	NA
10 yrs	×2	5	−0.159	1.00	1.00	1.00	NA	NA
10 yrs	×2	20	−0.020	0.25	0.14	0.13	NA	NA
20 yrs	×1	5	−0.079	1.00	1.00	1.00	1.00	1.00
20 yrs	×1	20	−0.010	0.74	0.48	0.33	0.23	0.19
20 yrs	×2	5	−0.079	1.00	1.00	1.00	1.00	1.00
20 yrs	×2	20	−0.010	0.47	0.24	0.18	0.14	0.13

2018 Estimate: 24.46 $\mu\text{g/L}$

**Colonial National Historic Park – Thoroughfare
Bottom Dissolved Oxygen Concentration**

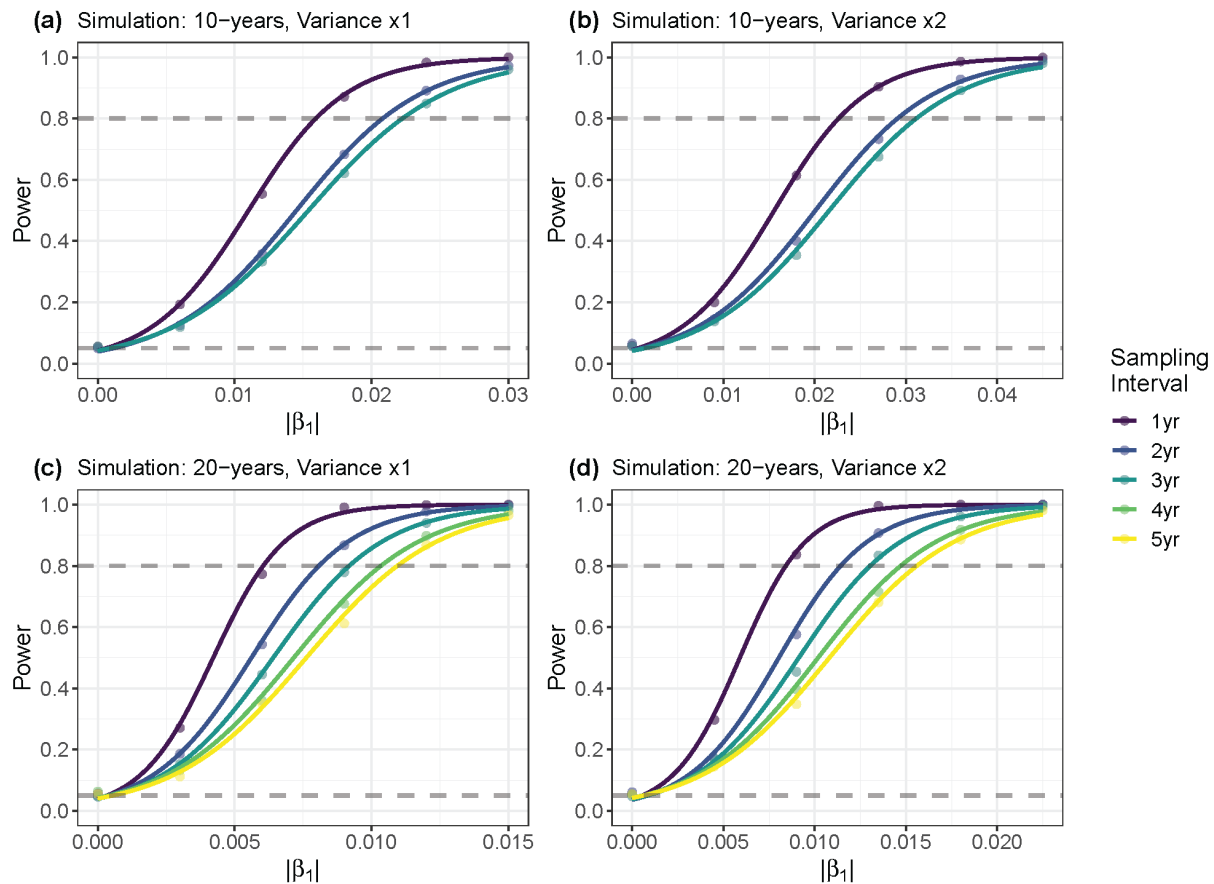


Figure 5-6. Estimated power to detect a linear trend in mean bottom dissolved oxygen concentration within the Thoroughfare stratum of Colonial National Historical Park. Power estimates vary based on the assumed sampling interval (see legend) and duration: (a) 10-year period, (c) 20-year period. Power 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 5-6. Estimated power to detect a linear trend in mean bottom dissolved oxygen concentration large enough to cross both quality thresholds within the Thoroughfare stratum of Colonial National Historical Park. 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 2.3.2 is provided in the table footnotes

Period	Variance	Threshold	β_1	Power by Sampling Interval				
				1yr	2yr	3yr	4yr	5yr
10 yrs	×1	2	−0.119	1.00	1.00	1.00	NA	NA
10 yrs	×1	5	−0.027	0.99	0.94	0.91	NA	NA
10 yrs	×2	2	−0.119	1.00	1.00	1.00	NA	NA
10 yrs	×2	5	−0.027	0.90	0.74	0.68	NA	NA
20 yrs	×1	2	−0.059	1.00	1.00	1.00	1.00	1.00
20 yrs	×1	5	−0.013	1.00	0.99	0.97	0.94	0.92
20 yrs	×2	2	−0.059	1.00	1.00	1.00	1.00	1.00
20 yrs	×2	5	−0.013	0.98	0.90	0.82	0.73	0.68

2018 Estimate: 6.54 mg/L

Colonial National Historic Park – Thoroughfare PAR Attenuation

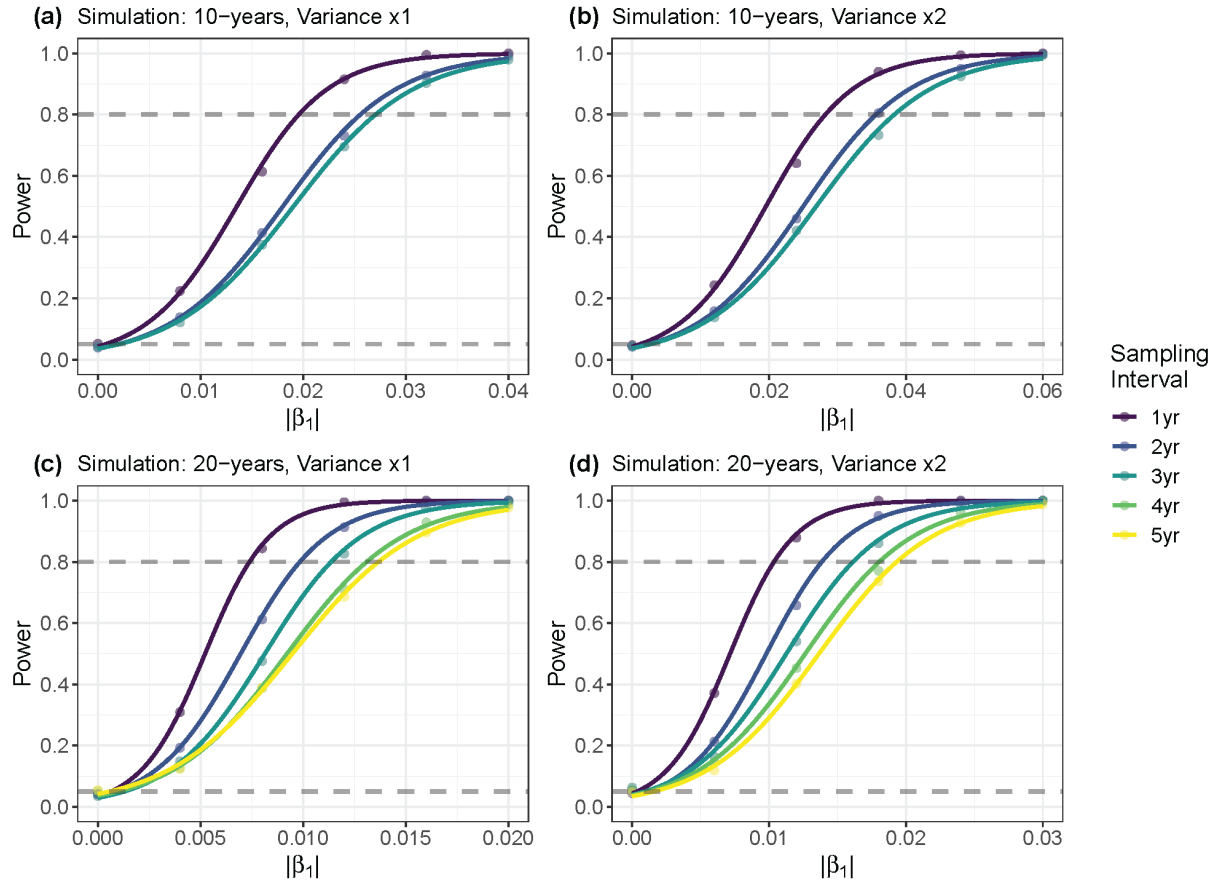


Figure 5-7. Estimated power to detect a linear trend in mean PAR attenuation within the Thoroughfare stratum of Colonial National Historical Park. Power estimates vary based on the assumed sampling interval (see legend) and duration: (a) 10-year period, (c) 20-year period. Power estimates in (b) and (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 5-7. Estimated power to detect a linear trend in mean PAR attenuation large enough to cross both quality thresholds within the Thoroughfare stratum of Colonial National Historical Park. 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 2.3.2 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	−0.127	1.00	1.00	1.00	NA	NA
10 yrs	×1	1.61	−0.071	1.00	1.00	1.00	NA	NA
10 yrs	×2	0.92	−0.127	1.00	1.00	1.00	NA	NA
10 yrs	×2	1.61	−0.071	1.00	1.00	1.00	NA	NA
20 yrs	×1	0.92	−0.064	1.00	1.00	1.00	1.00	1.00
20 yrs	×1	1.61	−0.036	1.00	1.00	1.00	1.00	1.00
20 yrs	×2	0.92	−0.064	1.00	1.00	1.00	1.00	1.00
20 yrs	×2	1.61	−0.036	1.00	1.00	1.00	1.00	1.00

2018 Estimate: 3.29 m^{−1}

5.2.3 All Strata

Estimated power for detecting a linear trend within all strata is summarized in both graphical and tabular format for surface chlorophyll-*a* concentration (Figure 5-8; Table 5-8), bottom dissolved oxygen concentration (Figure 5-9; Table 5-9), and PAR attenuation (Figure 5-10; Table 5-10).

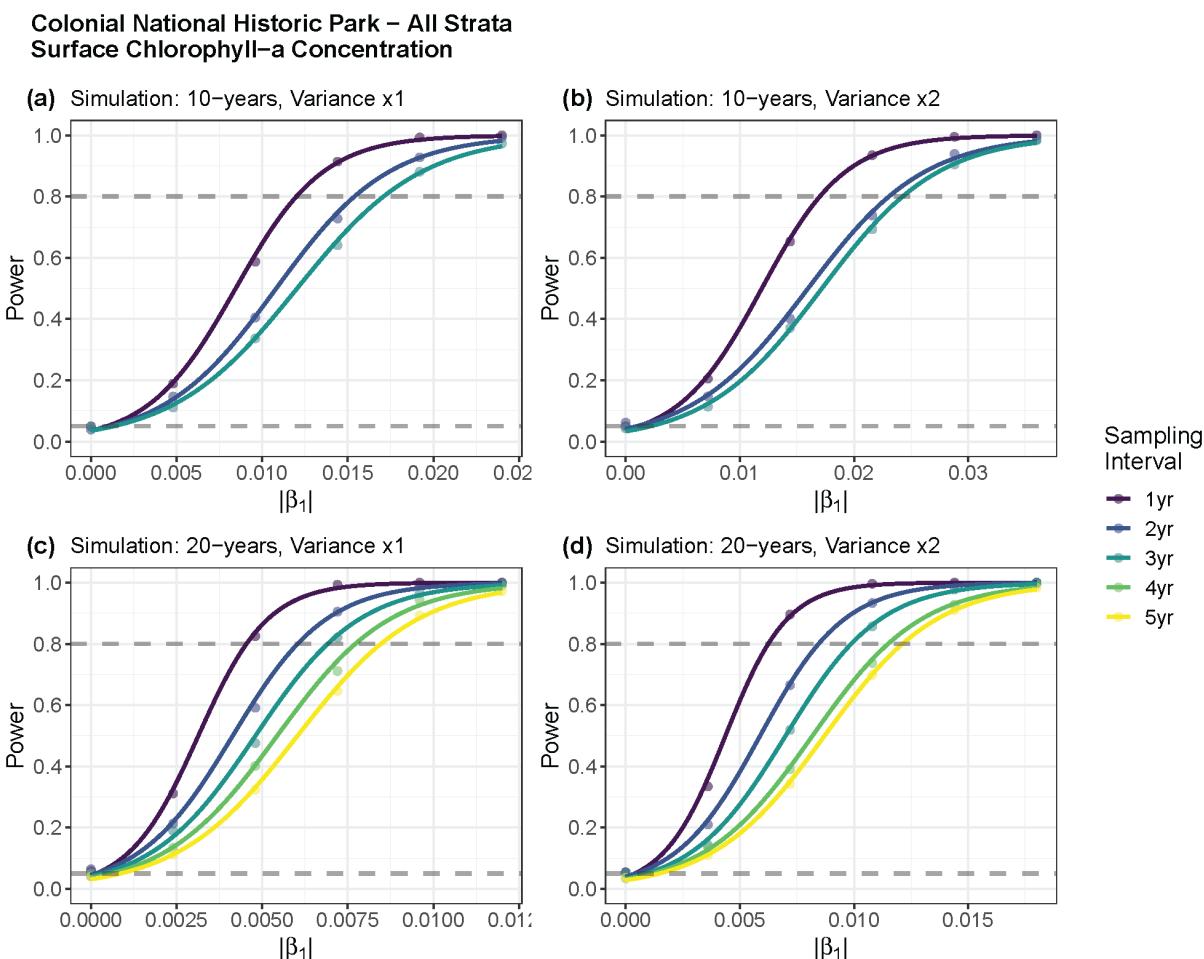


Figure 5-8. Estimated power to detect a linear trend in mean surface chlorophyll-*a* concentration within all strata of Colonial National Historical Park. Power estimates vary based on the assumed sampling interval (see legend) and duration: (a) 10-year period, (c) 20-year period. Power estimates in (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 5-8. Estimated power to detect a linear trend in mean surface chlorophyll-a concentration large enough to cross both quality thresholds within all stratum of Colonial National Historical Park. 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 2.4.3 is provided in the table footnotes.

Period	Variance	Threshold	β_1	Power by Sampling Interval				
				1yr	2yr	3yr	4yr	5yr
10 yrs	×1	5	−0.143	1.00	1.00	1.00	NA	NA
10 yrs	×1	20	−0.004	0.17	0.13	0.11	NA	NA
10 yrs	×2	5	−0.143	1.00	1.00	1.00	NA	NA
10 yrs	×2	20	−0.004	0.12	0.10	0.08	NA	NA
20 yrs	×1	5	−0.072	1.00	1.00	1.00	1.00	1.00
20 yrs	×1	20	−0.002	0.29	0.20	0.16	0.13	0.10
20 yrs	×2	5	−0.072	1.00	1.00	1.00	1.00	1.00
20 yrs	×2	20	−0.002	0.17	0.12	0.09	0.08	0.07

2018 Estimate: 20.92 $\mu\text{g/L}$

**Colonial National Historic Park – All Strata
Bottom Dissolved Oxygen Concentration**

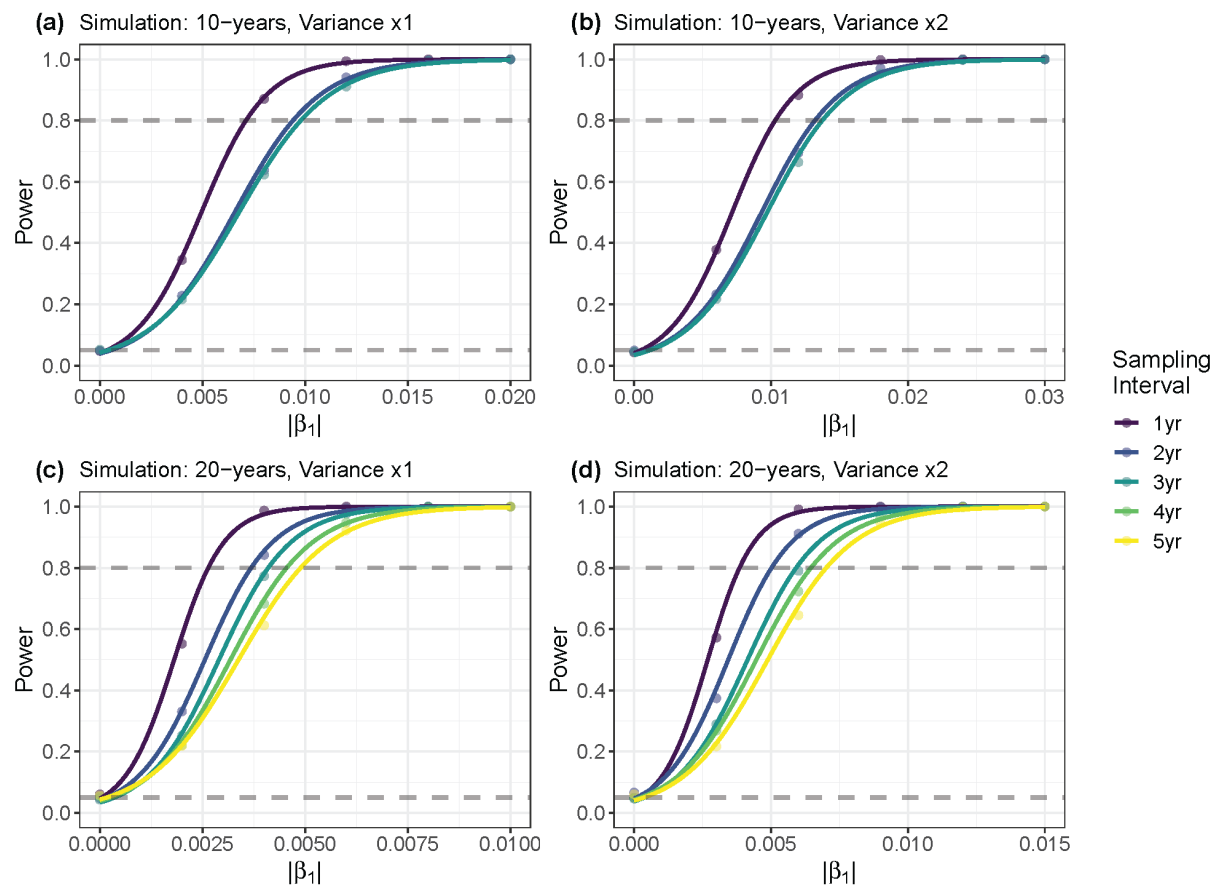


Figure 5-9. Estimated power to detect a linear trend in mean bottom dissolved oxygen concentration within all strata of Colonial National Historical Park. Power estimates vary based on the assumed sampling interval (see legend) and duration: (a) 10-year period, (c) 20-year period. Power estimates in (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 5-9. Estimated power to detect a linear trend in mean bottom dissolved oxygen concentration large enough to cross both quality thresholds within all strata of Colonial National Historical Park. 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 2.4.3 is provided in the table footnotes

Period	Variance	Threshold	β_1	Power by Sampling Interval				
				1yr	2yr	3yr	4yr	5yr
10 yrs	×1	2	−0.124	1.00	1.00	1.00	NA	NA
10 yrs	×1	5	−0.032	1.00	1.00	1.00	NA	NA
10 yrs	×2	2	−0.124	1.00	1.00	1.00	NA	NA
10 yrs	×2	5	−0.032	1.00	1.00	1.00	NA	NA
20 yrs	×1	2	−0.062	1.00	1.00	1.00	1.00	1.00
20 yrs	×1	5	−0.016	1.00	1.00	1.00	1.00	1.00
20 yrs	×2	2	−0.062	1.00	1.00	1.00	1.00	1.00
20 yrs	×2	5	−0.016	1.00	1.00	1.00	1.00	1.00

2018 Estimate: 6.89 mg/L

**Colonial National Historic Park – All Strata
PAR Attenuation**

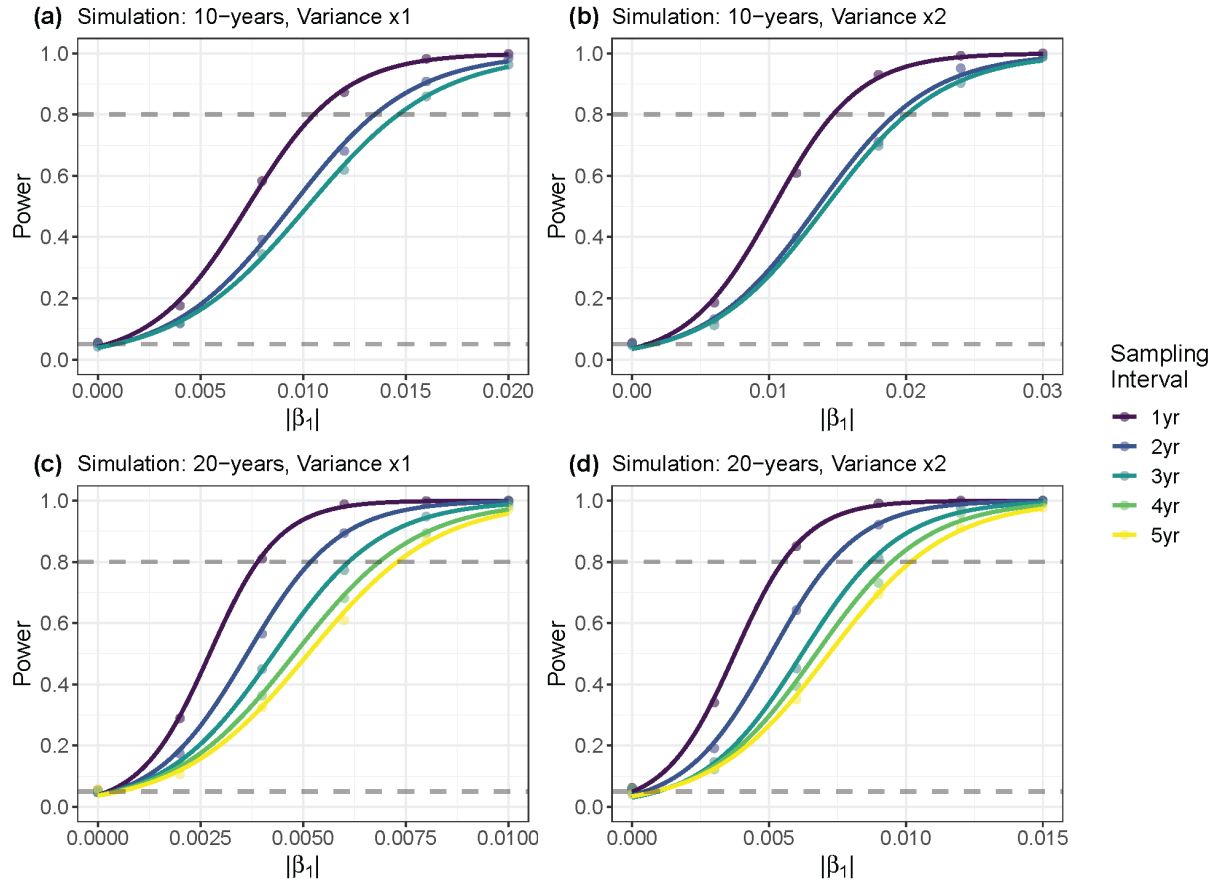


Figure 5-10. Estimated power to detect a linear trend in mean PAR attenuation within all strata of Colonial National Historical Park. Power estimates vary based on the assumed sampling interval (see legend) and duration: (a) 10-year period, (c) 20-year period. Power estimates in (b) and (d) correspond 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 5-10. Estimated power to detect a linear trend in mean PAR attenuation large enough to cross both quality thresholds within all strata of Colonial National Historical Park. 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 2.4.3 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	−0.116	1.00	1.00	1.00	NA	NA
10 yrs	×1	1.61	−0.060	1.00	1.00	1.00	NA	NA
10 yrs	×2	0.92	−0.116	1.00	1.00	1.00	NA	NA
10 yrs	×2	1.61	−0.060	1.00	1.00	1.00	NA	NA
20 yrs	×1	0.92	−0.058	1.00	1.00	1.00	1.00	1.00
20 yrs	×1	1.61	−0.030	1.00	1.00	1.00	1.00	1.00
20 yrs	×2	0.92	−0.058	1.00	1.00	1.00	1.00	1.00
20 yrs	×2	1.61	−0.030	1.00	1.00	1.00	1.00	1.00

2018 Estimate: 2.94 m^{−1}

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 5-8). From the table's footnotes, the estimated mean surface chlorophyll-*a* concentration in 2018 was 20.92 $\mu\text{g/L}$. 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 + x1	10 yrs + x2	20 yrs + x1	20 yrs + x2
Surface chlorophyll- <i>a</i> (poor)	1 yr	1 yr	1 yr	1 yr
Bottom dissolved oxygen (good)	3 yr	3 yr	5 yr	5 yr
PAR attenuation (poor)	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 'poor' to 'fair' category for PAR attenuation and surface chlorophyll-*a*, and from 'good' to 'fair' in bottom dissolved oxygen.

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

- Bauer, Laurie J., and Thomas J. Miller. 2010a. Spatial and Interannual Variability in Winter Mortality of the Blue Crab (*Callinectes Sapidus*) in the Chesapeake Bay. *Estuaries and Coasts* 33(3):678–87.
- Bauer, Laurie J., and Thomas J. Miller. 2010b. Temperature-, Salinity-, and Size-Dependent Winter Mortality of Juvenile Blue Crabs (*Callinectes Sapidus*). *Estuaries and Coasts* 33(3):668–77.
- 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.
- 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/>.
- 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 Generalized 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 functions⁷:

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

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

⁷ 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 $\hat{f}(x)$.

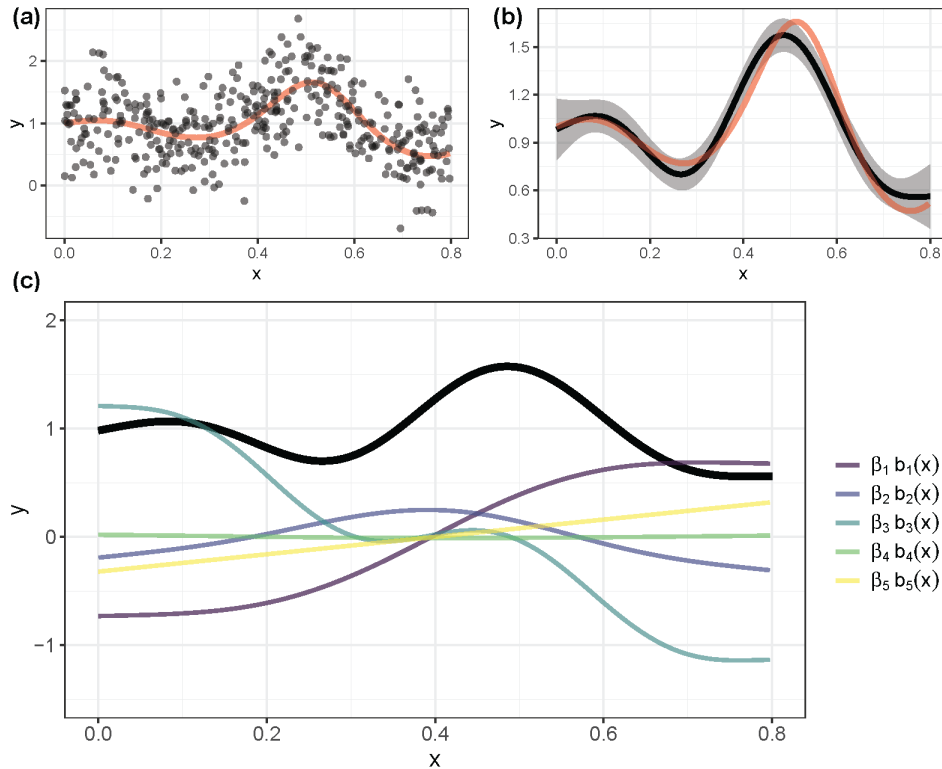


Figure A-1. (a) Data simulated from the nonlinear model given by Equation 2 and depicted in red. (b) Fitted generalized additive model smooth in black with corresponding 95% pointwise confidence interval in gray. True underlying curve provided in red for reference. (c) The six thin plate regression splines 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, Colonial National Historical Park has 7 years of monitoring data and therefore has 21 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 21 tests approximately 66% 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.