Florida Bay Water Quality Grab-Sampling Program Evaluation

Version 1.0—February 26, 2025



An aerial view of Florida Bay in Everglades National Park.
NPS Photo

Kelly R. McCaffrey (https://orcid.org/0000-0001-7705-5876)[1], Bradley A. Strickland[1], Zachary W. Fratto[2]

[1]Data Sciences Branch, South Florida Natural Resources Center, Everglades National Park, 950 N. Krome Avenue, Suite 301, Homestead, FL 33030-4443.

[2]Florida Bay Interagency Science Center, Everglades National Park, 98630 Overseas Highway, Key Largo, FL 33037

Correspondence: krmccaffrey@nps.gov or bastrickland@nps.gov

Abstract

Everglades National Park (ENP) monitors water quality in Florida Bay in part through a water quality grab sampling program, whose objective at establishment was to monitor picocyanobacterial blooms using measurements of chlorophyll-*a* concentrations and related water quality parameters. This evaluation integrates data from the ENP program and data from a similar South Florida Water Management District (SFWMD) program. We assess aspects of the current sampling design, including the sampling sites, water quality parameters, and sampling frequency. We emphasize the importance of establishing measurable program objectives and thoroughly documenting the program design. Additionally, we provide an overview of sampling design components which may need modification to align with specific programmatic goals.

Acknowledgements

We wish to acknowledge the efforts of the numerous South Florida Natural Resource Center scientists who contributed to the development of this work and the completion of related field and laboratory work. We would particularly like to thank G. Melodie Naja, Tylan Dean, Erik Stabenau, Laura Bauman, Christian Lopez, and Meagan Schrandt whose expertise and participation in an initial planning workshop was critical to the development of this work.

Executive Summary

- This report reviews the Everglades National Park (ENP) Florida Bay water quality grab sampling program, using program data and external data from a similar program by the South Florida Water Management District (SFWMD). We evaluate the existing ENP program design to assess potential sampling site redundancy, the similarity of water quality parameters, and the impact of altering sampling frequency on annual status.
- Three pairs of ENP sampling sites were identified as potentially redundant, as they were close to one another and the chlorophyll-*a* measurements obtained were similar throughout the time series (NEST01:NEST02, NMB02:NMB03, and RUS01:RUS02). However, alteration of the spatial sampling design should be considered in the context of program objectives. Regardless of the decided objective, the removal of one of these pairs is unlikely to significantly affect the overall effectiveness of the program and may yield nominal cost savings.
- Several water quality parameters were correlated, but it is unclear whether they are providing redundant information given the uncertainty of physical or ecological mechanisms for their correlation. Two pairs of water quality parameters were consistently more similar to one another than to other parameters, across SFWMD sampling sites (total nitrogen (TN) and total organic carbon (TOC); ammonia/ammonium (NH4) and nitrate+nitrite (NOX)). Many ENP parameters (e.g., soluble reactive phosphorus (SRP), nitrate+nitrite (NOX), nitrite (NO2)) have a high proportion of measurements below laboratory detection limits. Before reducing the number of water quality parameters collected, acceptable levels of similarity between parameters and analytical detection limits should be scrutinized by program managers. The suite of water quality parameters collected should align with program objectives.
- Altering the sampling frequency of SFWMD data (collection every one, two, three, four, or six months) did not significantly impact annual median estimates of chlorophyll-*a* concentrations for regions or sites, but did influence the maximum concentration values obtained. Most water quality parameters exhibit evidence of seasonality, trend, or both at some or all sites. The sampling periodicity will impact whether the peak intensity of algal bloom events is captured. Low frequency sampling may result in the start of algal blooms, or entire ephemeral events, being missed. Determination of adequate sampling frequency will be dependent on program objectives.
- **It is crucial that measurable program objectives be defined prior to altering aspects of the sampling design.** It is likely that the sampling design will need to be altered to fulfill one or more program objectives. Examples of water quality monitoring programs designed to meet different objectives are discussed. Once the objectives and design are agreed upon, program documentation should be created outlining critical aspects of monitoring program design.

Questions, Objectives, and Recommendations

Questions

- Are there sampling sites within the ENP Florida Bay water quality sampling program from which very similar (redundant) measurements of chlorophyll-*a* are being obtained?
- How similar are the water quality parameters being collected, and are there any that could be eliminated without significant loss of information?
- How frequently should samples be collected?

Objectives

- Determine if any ENP Florida Bay water quality sampling sites are redundant, based on chlorophyll-*a* concentrations, and provide resulting recommendations for the spatial design of the network.
- Assess how similar the collected water quality parameters are to determine if there are any which may be eliminated.
- Evaluate the impact of changes to the frequency of sampling on estimates of annual summary statistics and long-term trend estimation to determine how frequently samples may need to be collected to meet program goals.

Recommendations

- Program objectives must be defined or re-articulated as SMART objectives (Specific, Measurable, Achievable, Results-oriented, and Time-fixed). No alterations or changes to the sampling design should be made prior to the definition of measurable program objectives. This will best be achieved in an internal workshop and/or formal decision-making framework.
- Program objectives should guide the spatial and temporal sampling design of the program. Different sampling design strategies are required to achieve different types of objectives, and several design types may be necessary to meet program objectives. The information within this document may serve as guidance, but the spatial and temporal sampling design and water quality parameters collected should ultimately be determined by program objectives.
- It is recommended that the analytical detection limits of water quality parameters be considered relative to expected background concentrations. Managers will need to seek alternative laboratory methods with lower detection limits where necessary or consider the cost of measuring parameters which are routinely below detection limits.
- It is important that program documentation be created to record program design and structure. A suggested outline of the minimum content is presented.
- We suggest holding a Florida Bay water quality monitoring workshop with external stakeholders (e.g., SFWMD, FIU, EPA) to foster collaboration, improve data utilization, and maximize the compatibility of ENP's dataset with other sampling efforts.

- Future iterations of the program's database should be modified to include columns for laboratory detection limits and laboratory data quality flags associated with each measurement.
- Other practical considerations (e.g., cost, logistics) and complimentary sources of data (e.g., remote-sensing, sondes) are discussed only briefly herein but will be necessary to consider further once the program objectives are agreed on and design is refined.

Table of Contents

Abstract ii
Executive Summary iv
Questions, Objectives, and Recommendations v
Questions v
Objectives v
Recommendations v
Table of Contents vii
Introduction 1
Methods 3
Sampling Design 3
Sample Collection and Processing 4
External Data 5
Data Analysis 6
Sample Site Similarity 6
Water Quality Parameter Similarity 8
Sampling Frequency 9
Results 10
Sample Site Similarity 10
Water Quality Parameter Similarity 13
Sampling Frequency 15
Discussion 17
Sample Site Similarity 17
Water Quality Parameter Similarity 18
Sampling Frequency 19
Program Objectives Inform Program Design 19
Program Documentation 21
Other Considerations 22
Literature Cited 23
Appendix A. Supporting Tables and Figures 27

Introduction

Florida Bay is a shallow coastal estuary, which is composed of a series of basins separated by shallow banks, limiting water exchange between basins and between Florida Bay and the Gulf of Mexico (**Figure 1**; Cosby et al. 2004; Cosby et al. 2005; Lee et al. 2016). Following a period of drought and hypersalinity in 2015 and an associated major seagrass die-off, a water quality sampling project was established by Everglades National Park (ENP) in Florida Bay in October 2016 to determine the spatial extent and concentration of an algae bloom that was expected post die-off. Most sampling sites were established and routine sampling across the bay began in August 2017. In September 2017, Hurricane Irma exacerbated the already poor conditions in Florida Bay by resuspending nutrient-laden benthic substrates and causing mangrove mortality in coastal regions which drain into the bay. These environmental disturbances were expected to have long term impacts to the system. Thus, the objective of the network at its establishment appears to have been to monitor picocyanobacterial blooms (*Synechococcus* spp.) using chlorophyll-*a* concentration as a proxy measurement for phytoplankton biomass. Adjustments have not been made to the project following Hurricane Irma. The project, originally designed to provide periodic "snapshots" of conditions, was not intended to be a long-term monitoring program. Other water quality parameters, some of which may be associated with reduced water quality in the oligotrophic Florida Bay (e.g., nitrogen and phosphorous), are also routinely measured.

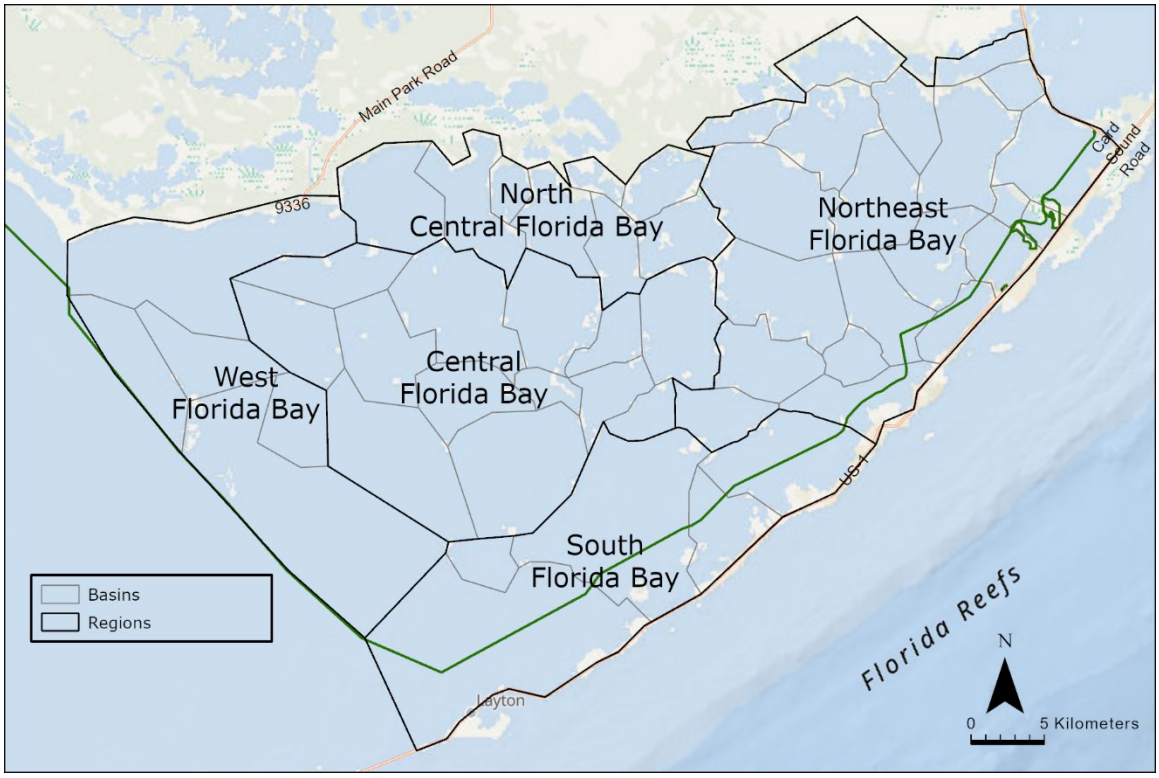


Figure 1. A map of Florida Bay, showing basins defined by Cosby et al. (2004) (light gray lines) grouped into regions (black lines) used for World Heritage Site reporting (Naja et al. 2022).

The design of this network and its ongoing objectives have not been well-documented. Some objectives of the program articulated and inferred by staff include:

- Delineation of the extent of algal blooms,
- Documenting the peak intensity of algal blooms,
- Determination of the drivers of algal blooms,
- Establishment of baseline and abnormal water quality conditions,
- Tracking changes in water quality during periods of hydrological transition (e.g., seasonal, restoration…), and
- Monitoring of water quality conditions within ENP to meet reporting needs (e.g., for the Comprehensive Everglades Restoration Plan).

Given the time that has passed since program establishment, increasing interest in the water quality data products, improvements in scientific knowledge, and workload and budget considerations, the water quality monitoring program would benefit from review and optimization. Optimization of any monitoring program involves an appraisal of program elements to verify that objectives are being met (Behmel et al. 2016). Relevant objectives of a program appraisal for the Florida Bay water quality monitoring program, modified from Behmel et al. (2016), should include:

1. Identification or assessment of monitoring objectives,
2. Appraisal of the spatial representativeness of monitoring sites,
3. Evaluation of collected water quality parameters,
4. Assessment of sampling frequency,
5. Consideration of technical, human, and financial resources,
6. Development of a plan for quality control, data handling, data analysis, reporting, and dissemination.

Defining the monitoring objectives requires a careful consensus on the key management or research question that the program aims to address. This process should involve a formal decision-making framework (see Discussion). The analyses presented in this report focus on points two through four, evaluating sample site similarity, water quality parameter similarity, and sampling frequency. An overview of practical considerations, including program resources and planning, are included in the Discussion.

Methods

Sampling Design

Everglades National Park water quality monitoring sites were dispersed through a subset of basins in the bay, defined by Cosby et al. (2004). There was generally one sample site per basin for basins which have not typically experienced algal bloom conditions (e.g., Northeast Florida Bay), and typically more than one site in basins more likely to experience bloom conditions (**Figure 2**). Sampling usually occurred on a bimonthly basis but varied through time with extended gaps in sample collection between 2018–2019 and 2022–2023 (maximum gap 270 days between sample collection; **Figure 3**). Sampling can take multiple days to complete, so one "sampling event" may occur over two calendar months if it was started at the end of the month. Thus, a unique sampling event is herein defined as the collection of water samples over the course of 7 days or less. A total of 40 sampling events were completed between August 2017 and September 2024 (**Figure 3**). Sites were normally sampled once per sampling event (i.e., not re-sampled within 7 days) except for November 2017, when 13 sites were sampled on two collection dates.

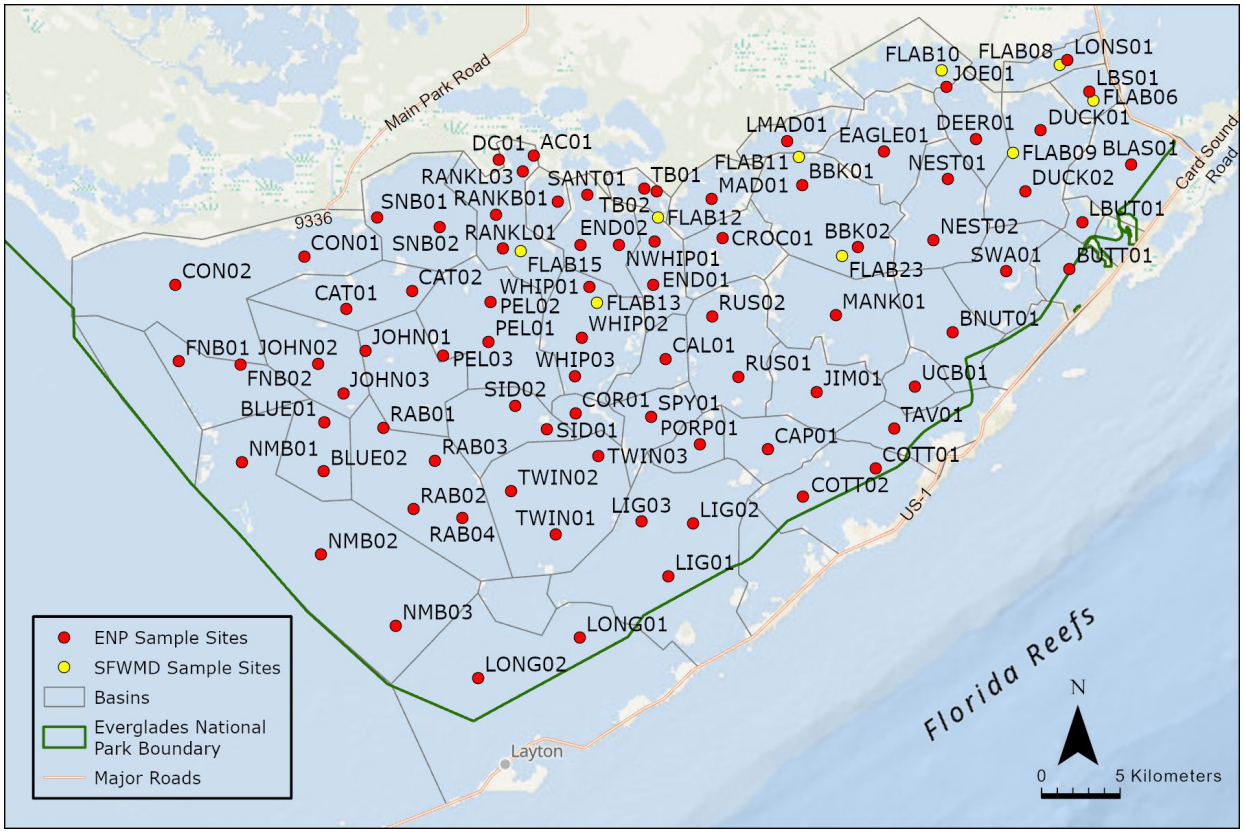


Figure 2. Everglades National Park (ENP) water quality sampling sites, shown with South Florida Water Management District (SFWMD) monthly water quality sampling sites used for the sampling frequency and water quality parameter similarity analysis.

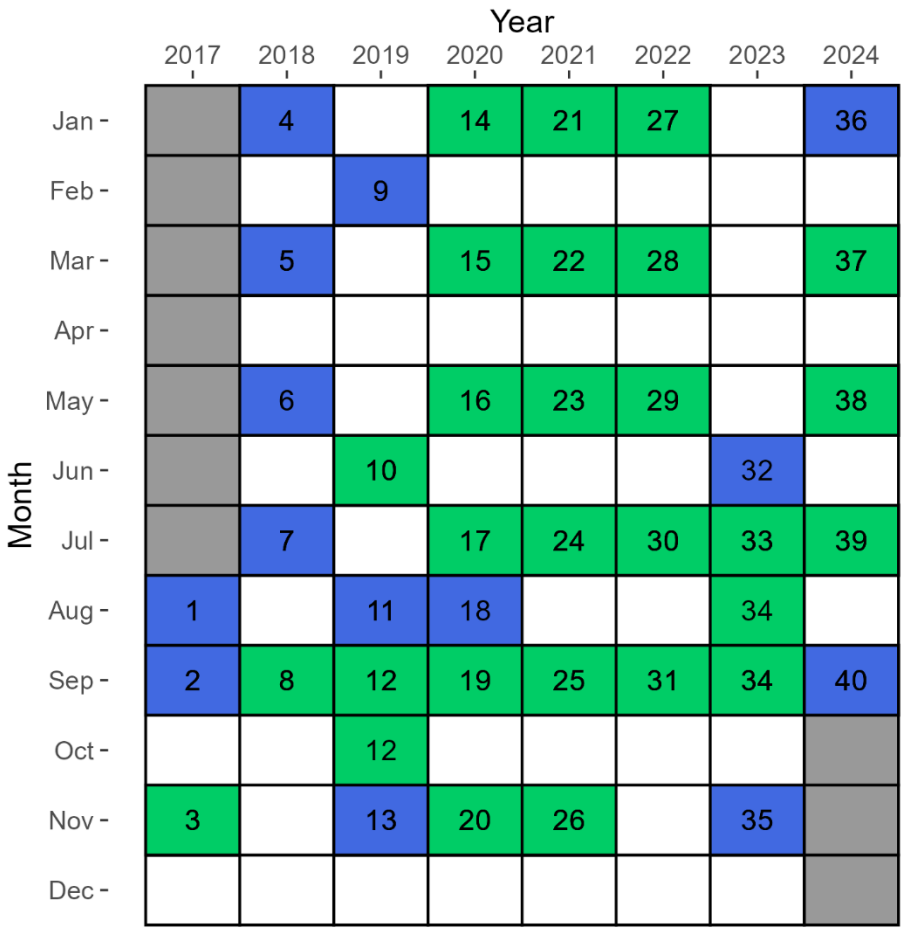


Figure 3. A graphical representation of sample collection between August 2017 and July 2024. Tiles for months during which samples were collected are colored; green months were used for the site similarity analysis, while blue months were not due to incomplete sampling or missing measurements (see Data Analysis). Sampling events are identified as numeric labels on each monthly tile.

There are additionally fixed-station data loggers (sondes), which monitor water quality parameters at more frequent intervals (30 minutes). The sondes are attached to a subset of the hydrologic monitoring network stations in Florida Bay and are maintained once monthly, at which time the water quality parameters measured by the sondes (turbidity, fluorescent dissolve organic matter (FDOM), and chlorophyll-*a*) are calibrated to in-field or grab sample measurements. The parameters measured by these stations include temperature, dissolved oxygen, sonde depth, pH, conductivity (salinity), turbidity, chlorophyll-*a,* and FDOM. Due to inherent differences in data type and methodology, data from the fixed-station sondes are not evaluated here. In the Discussion we briefly outline how they may be incorporated into a planned water quality monitoring network.

Sample Collection and Processing

Simplified versions of field- and laboratory-protocols are summarized here, focusing on the collection and analysis of water samples to measure chlorophyll-*a* concentrations. For more detailed procedures, reference the cited field and laboratory protocols for this project.

At each site, two replicate water samples of approximately 200 mL (250 mL Nalgene with about 50 mL of headspace) were collected at the same time, approximately 0.5 m under the surface of the water. Upon return from the field, 50 mL of water from each sample was passed through a 25 mm filter. If the filter clogged before the full 50 mL of water was used, 25 mL of water was used instead. Filters were then wrapped in aluminum foil and frozen until laboratory analysis could be completed (Fratto 2024). Samples may have been processed and stored this way for up to 3.5 weeks (Arar and Collins 1997).

The frozen filters were ground with 10 mL of acetone to produce a slurry then steeped for approximately 20 hours, but for a maximum of 24 hours, to extract the chlorophyll-*a*. After steeping, the slurry was centrifuged at 3,000 rpm for 10 minutes to clarify the solution. An aliquot of the supernatant was transferred to a glass cuvette and the fluorescence of each sample was measured to determine the concentration of chlorophyll-*a*. The chlorophyll-*a* concentration in each sample was determined following Arar and Collins (1997). Only one reported chlorophyll-*a* was value below the method detection limit (MDL), the minimum concentration of the analyte that can be measured and reported with 99% confidence, was replaced with the MDL (0.025 µg L^{-1}; Turner Designs Inc.).

Water samples were sent to Florida International University (FIU) to measure additional analytes, including nitrate+nitrite (NOX), nitrite (NO2), ammonia/ammonium (NH4), soluble reactive phosphorus (SRP), total phosphorus (TPO4), dissolved silicate (Si), total organic carbon (TOC), and total nitrogen (TN). The standard methods for the laboratory analyses procedures are available in the Appendix, with the method detection limit (MDL) values for each parameter during the period of analysis (**Table A-1**). Non-acidified refrigerated and frozen samples were retained for up to 2 weeks from collection and then transported to FIU. The reports returned by FIU include the analyte measurements and associated MDL values, practical quantitation limits (PQL) and data quality remark codes. The proportion of values below MDL and proportion of missing values of these analytes are summarized in the Appendix (**Table A-2** and **Table A-3**). Only chlorophyll-*a* measurements from the ENP sampling program are used in the following analyses. The proportion of missing values and the proportion of values below laboratory detection limits for all analytes measured by the FIU laboratory are discussed below.

External Data

The South Florida Water Management District (SFWMD) has water quality grab-sampling locations spread throughout Florida Bay, which are sampled on a monthly or bimonthly basis (Brown et al. 2024) with data available from DBHydro Insights (https://apps.sfwmd.gov/dbhydroInsights/#/homepage). Data were obtained on November 18, 2024, for nine sites where water quality grab samples were collected monthly, within the same regions as ENP sampling sites: FLAB06, FLAB08, FLAB09, FLAB10, FLAB 11, FLAB12, FLAB13, FLAB15, and FLAB23 (**Figure 2**). The period of sample collection was between May 1, 2016 and April 30, 2024 to represent water years (WY) 2017 through 2024. Seventeen water quality parameters of interest, which matched those collected by ENP or which were otherwise of interest in relation to algal blooms, were available during this period (**Table 1**). Field-sampling methodology use was similar to that of ENP, but SFWMD uses high-performance liquid chromatography (HPLC)

to determine the concentration of chlorophyll-*a* in their water quality samples (Arar 1997). This difference in methodology may mean that chlorophyll-*a* values are not directly comparable between methods, as measurements obtained using HPLC may be different than those obtained using fluorometric methods (FDEP 2011).

Data quality remark codes are associated with SFWMD water quality data, but no data were excluded from these analyses based on remark codes. Water quality parameters with measurements below their associated MDL values were already replaced with the value of the MDL. The MDL values for all available water quality parameters are summarized in **Table 1**. The percentage of values equal to the MDL of the measurements is summarized in the Appendix (**Table A-4**).

Duplicate water quality samples were collected in September 2016 on the 8th and 13th. Because it was unknown why duplicate samples were collected during this month, the average value for each parameter was used. The percentage of missing monthly measurements at each site during the period analyzed is summarized in the Appendix (**Table A-5**).

Data Analysis

Sample Site Similarity

Chlorophyll-*a* measurements from the ENP water quality grab-sampling program were evaluated to determine if there were any outliers which should be excluded from analysis. Three measurements from samples take on September 27th and 28th in 2017 at RANKL03, TB01, and NWHIP01 were extraordinarily high (145.07 – 219.17 µg L^{-1}), likely due to turbidity and debris in the water from Hurricane Irma, which had recently passed over the region. These measurements were considered unreliable and were excluded from all analyses.

Hierarchical agglomerative clustering requires a single observation per site and sampling event, to calculate pairwise dissimilarity metrics between measurements in each timeseries. Absolute difference in replicate samples taken on a given day ranged between 0 and 13.22 µg L^{-1} (median 0.08 µg L^{-1}). In the study period, 95% of sample replicates had an absolute difference of less than 1.31 µg L^{-1}. Larger differences between replicates were moderately associated with larger values of the average of the chlorophyll-*a* concentration of both replicate samples (Kendall's $\tau = 0.5$, $p < 0.001$). Because the difference between replicate samples was generally small, the average of the replicates was used to represent the chlorophyll-*a* concentration at each site on each collection date. Sites with larger differences between replicates were retained for this analysis, as the reason and significance of larger difference between replicate samples could not be determined, and program managers agree that using the average of replicate samples was an adequate representation of the "true" site conditions.

The sites sampled twice in November 2017, on the 13th and 21st of the month (8 days apart). The difference in chlorophyll-*a* concentration between the first and second visit was between -3.59 and 14.51 µg L^{-1} (median 0.10 µg L^{-1}). Because it was suspected that sites may have been sampled twice due to a weather event, which may have included winds which could have pushed any active algal blooms to different locations, only data collected during the first sampling event on the 13th of

November was used for these sites. The resulting dataset had one value for chlorophyll-*a* per site during each sampling event (**Figure 3**).

Hierarchical agglomerative clustering of timeseries also requires that there are no missing values. The number of sampling events per site was evaluated to determine if any sites had a high number of missing values. Sites SANT01, DC01, AC01, and TB02 were missing measurements for between 30 and 33 sampling events, because they were more recently added to the sampling program (between November 2021 and July 2023) to collect data for a cooperating university. These sites were removed from the dataset to avoid excluding a large proportion of sampling events. Site TB02 was also eliminated after the data collection effort for the so cooperating university ended. With these sites excluded, the number of observations per sampling event was evaluated. Earlier sampling events (1, 2, 4, 5, 6, and 13) were missing a large proportion of measurements because only a limited area of the bay was sampled, excluding sites in Northeastern Florida Bay. Retaining only sampling events during which all sites were sampled resulted in a final dataset of 77 sites and 27 sampling events.

Following Soares et al. (2018), two different dissimilarity metrics were used for clustering, to account for different aspects of similarity: the difference in magnitude of values across the time series (Euclidean distance; **Equation 1**) and changes in concentration over time (Kendall's correlation distance; **Equation 2**). The Euclidean distance (d_{euc}) of timeseries x and y was found using Equation 1, where x_i and y_i represent chlorophyll-*a* concentrations during sampling event i, and n represents the total number of sampling events. A d_{euc} of 0 would mean that both timeseries had the same value at every sampling event.

$$d_{euc}(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \quad (1)$$

Kendall's correlation distance was used as a non-parametric alternative to Pearson's correlation distance and was found using Equation 2. This metric was chosen because the data did not meet parametric assumptions, and so the relative order of the rank assigned to chlorophyll-*a* values in the ordered timeseries could be evaluated, rather than the correlation of the exact rank assigned to the values (Kassambara 2017). The values of timeseries x and timeseries y are ordered and decomposed into $\frac{1}{2}n(n-1)$ ordered pairs. The number of concordant (n_c) and discordant (n_d) pairs in both timeseries are counted, and then normalized by the total number of pairs available to produce a correlation distance value (d_{kend}) between 0 and 1. A perfectly correlated pair of timeseries, with values in the same relative order, would have $d_{kend} = 0$.

$$d_{kend}(x, y) = 1 - \frac{2(n_c - n_d)}{n(n-1)} \quad (2)$$

To determine if each distance metric could be clustered, the clustering tendency of each was compared to a randomly generated dataset, using an ordered dissimilarity image (Bezdek and Hathaway 2002). If there was evidence that they could be clustered, hierarchical agglomerative clustering was performed on the distance metrics using three different linkage algorithms to define distances between clusters: single, average, and complete linkage. Complete linkage sets the distance between clusters as equal to the largest pairwise dissimilarity value between all sites in the two clusters. Single linkage sets the distance as the smallest pairwise dissimilarity value, and average linkage is the average of all pairwise dissimilarity values of sites in the clusters. The performance of the linkage methods was assessed by evaluating the cophenetic distance (the height of the node at which two observations are first joined) of the resulting dendrogram produced by each method, to the corresponding Euclidean or Kendall's correlation distance values (Kassambara 2017). The results of the best-performing linkage type are presented here.

To determine if similar sites were close to each other in space, the distance between each pairwise combination of sites was found using the site's coordinates and compared to the Euclidean and Kendall's correlation distance metrics. All cluster analyses were completed using the R packages "cluster" and "factoextra" (Kassambara 2017; Maechler et al. 2023).

Water Quality Parameter Similarity

Because the SFWMD monthly sampling schedule offered a longer, more-complete time series of measurements (**Tables A-2–A-5**), SFWMD data was used for the water quality parameter similarity analysis. Water quality parameter measures from the SFWMD dataset may not be directly comparable to ENP measurements due to analysis techniques or measurement units, but we expected them to exhibit the same general trends and relationships. Hierarchical agglomerative clustering was used to evaluate the similarity of water quality parameters. Kendall's correlation distance was used as a measure of dissimilarity because the data did not meet parametric assumptions, and to reduce the effect of censored values (values < MDL) on correlation test results (Helsel et al. 2020). Water quality parameter measures from the SFWMD dataset may not be directly comparable to ENP measurements due to analysis techniques or measurement units, but we expected them to exhibit the same general trends and relationships. Because the SFWMD monthly sampling schedule offered a longer, more-complete time series of measurements (**Table A-2–Table A-5**), SFWMD data was used for the water quality parameter similarity analysis.

Only water quality parameters associated with algal blooms, eutrophication, or water clarity were included in this analysis (NH4, NOX, TURB, TOC, TN, CHLA, TPO4, SECCHI; **Table 1**). While other available parameters were also of interest (CHLB, NO2, NO3, PHEO, OPO4), they were excluded either due to a high proportion of missing values (**Table A-4**) or a high proportion of measurements below the MDL (**Table A-5**). The similarity of water quality parameters was evaluated at each site individually, as it was expected that there may be environmental differences between sites. The dataset for each site was limited to months with a measurement for each parameter, which resulted in 91–95 sampling dates included for each site. Kendall's correlation distance with average linkage was used to identify parameters which may be considered redundant to

each other. The clustering tendency of parameters at each site was evaluated, following the same procedure as above, prior to constructing dendrograms.

Table 1. Water quality parameters obtained for South Florida Water Management District (SFWMD) sampling sites. The units, abbreviations, and method detection limits (MDLs) of the laboratory analyses methods are provided. For parameters which had varying MDLs through time, the MDL is presented as a range capturing the minimum and maximum MDL value during the period of analysis. Parameters with no MDL or associated unit are marked N/A.

Water Quality Parameter (SFWMD)	Units	Abbreviation	Method Detection Limit(s)
AMMONIA-N	mg L^{-1}	NH4	0.005 mg L^{-1}
CARBON, TOTAL ORGANIC	mg L^{-1}	TOC	0.8 mg L^{-1}
CHLOROPHYLL-A(LC)	µg L^{-1}	CHLA	0.016–0.16 µg L^{-1}
CHLOROPHYLL-B (LC)	µg L^{-1}	CHLB	0.016–0.08 µg L^{-1}
DISSOLVED OXYGEN	mg L^{-1}	DO	N/A
NITRATE+NITRITE-N	mg L^{-1}	NOX	0.005 mg L^{-1}
NITRATE-N	mg L^{-1}	NO3	0.005 mg L^{-1}
NITRITE-N	mg L^{-1}	NO2	0.002 mg L^{-1}
PH, FIELD	N/A	PH	N/A
PHEOPHYTIN-A (LC)	µg L^{-1}	PHEO	0.008–0.16 µg L^{-1}
PHOSPHATE, ORTHO AS P	mg L^{-1}	OPO4	0.002 mg L^{-1}
PHOSPHATE, TOTAL AS P	mg L^{-1}	TPO4	0.002 mg L^{-1}
SALINITY	PSU	SAL	N/A
SECCHI DISK DEPTH	meters	SECCHI	N/A
TOTAL NITROGEN	mg L^{-1}	TN	0.02–0.05 mg L^{-1}
TURBIDITY	NTU	TURB	0.1 NTU
Temperature	°C	TEMP	N/A

Sampling Frequency

Monthly SFWMD measurements of chlorophyll-*a* were used to evaluate the impact of changes to sampling frequency. Missing observations in the time series were retained as a complete time series was not required, and it was assumed that the reason for missingness was due to a cause unrelated to the planned sampling schedule. Analysis was limited to the period between WY17–WY24.

Monthly data from SFWMD sampling sites was sub-sampled to form datasets which simulated if sampling occurred only every two, three, four, or six months at each site. Selection of dates which were sampled always began in May 2016. With monthly and sub-sampled datasets, the annual minimum, median, 25th and 75th percentiles, and maximum values of chlorophyll-*a* were found for each sampling site, and for each region (**Figure 8** and **Figure A-1**). Significant difference in estimates obtained using different sampling frequencies were evaluated with a Mann-Whitney U test.

Results

Sample Site Similarity

The organized dissimilarity index plots for both Euclidean distance (d_{euc}) and Kendall's correlation distance (d_{kend}) indicated that the data may be clustered. For each distance metric, cluster results using the average linkage method was the most representative, producing cophenetic distances which were most similar to the original pairwise distances. Results of hierarchical agglomerative clustering with average linkage using d_{euc} (**Figure 4**) and d_{kend} (**Figure A-2**) are presented here.

A tanglegram was produced to find common subtrees (pairs of sampling sites) which were consistently clustered using most dissimilarity metrics. Nine pairs of sites were consistently grouped in both dendrograms (**Figure A-3**). Of these, five pairs of sites were within the same basin (FNB01:FNB02, NEST01:NEST02, NMB02:NMB03, RUS01:RUS02, SID01:SID02) and were between 2.53–7.74 km from each other. Consistent pairings which were not in the same basin (BLUE01: JOHN02, CAL01:SPY01, CAT02:SNB02, NWHIP02:WHIP01) were in generally north-south arrangements and were between 2.69–4.41 km from each other (**Figure 2**). Pairwise distances between all sites ranged from 0.78–61.97 km. In general, there was a significant but weak association between pairwise site distance and both d_{euc} ($\tau = 0.10$, $p < 0.001$) and d_{kend} ($\tau = 0.27$, $p < 0.001$; **Figure A-4**).

To evaluate clusters generated using d_{euc}, a cut point at a height of 5.10 was used to represent an approximate difference of 1 μg L^{-1} per sampling event across the time series (see **Equation 2, Figure 4**). This cut point was chosen as, while fluorometry values are measured to two decimal places, a difference of whole-number values was deemed as appropriate for consideration of difference between samples by program managers. A larger cut point was not evaluated, as restoration indicator limits previously established using fluorometric measurements of chlorophyll-*a* in Florida Bay depend on median annual changes of less than 1 μg L^{-1} in regions of the bay (Boyer et al. 2009).

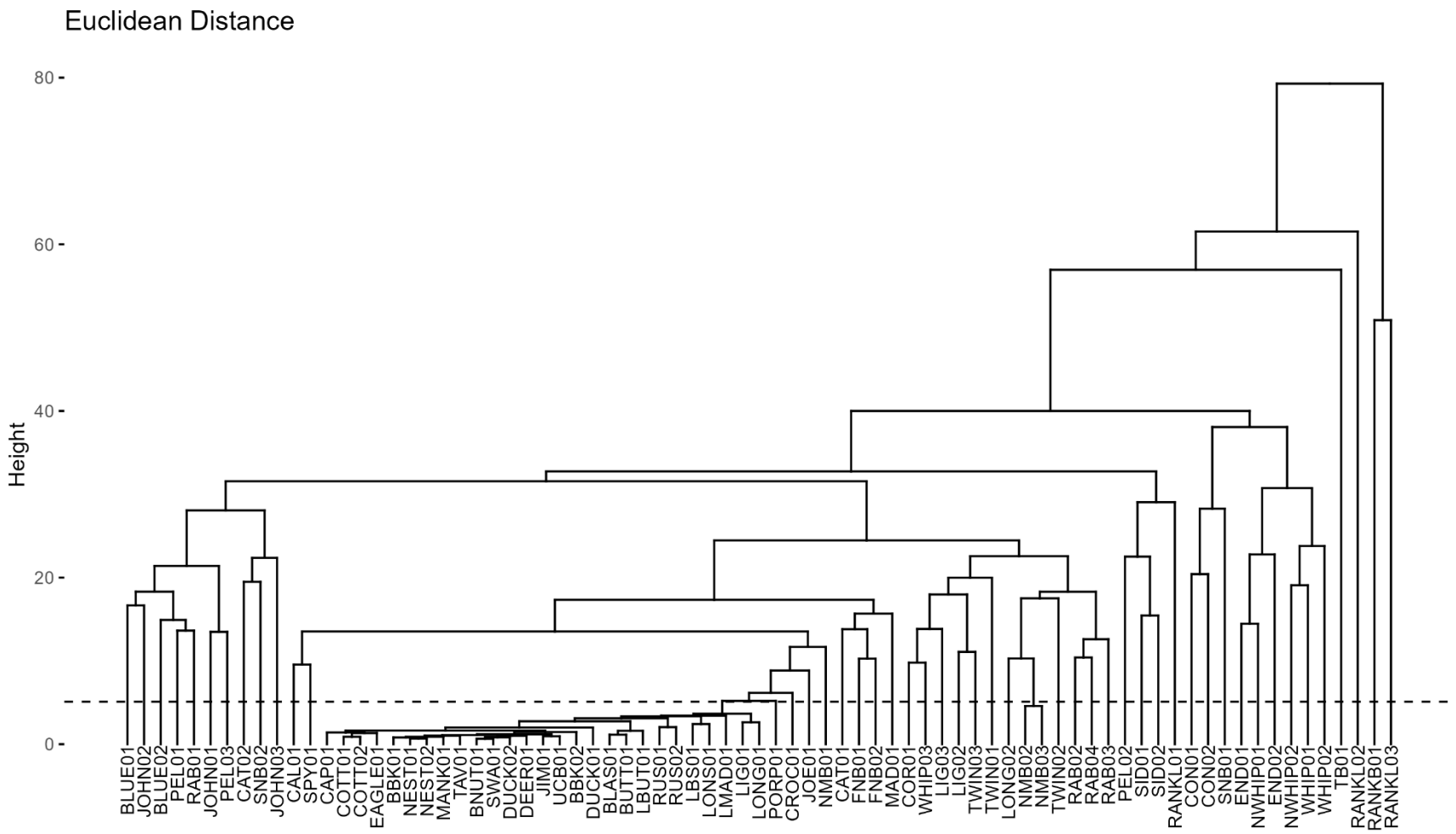


Figure 4. A dendrogram of cluster results, using the Euclidean distance metric (d_{euc}). Increased height at the nodes where cluster join represents sites with a greater magnitude of difference between measurements in the time series. The dashed line represents a chosen cut point, where sites with an approximate difference less than 1 µg L^{-1} chlorophyll-*a* per sampling event over the period were clustered.

One large cluster and one small cluster (of NMB02:NMB03) were formed using this cut point, while all other sites were not grouped. Most clustered sites were within Northeast Florida Bay (**Figure 5**). These sites were all located within a single cluster in the Kendall's correlation distance dendrogram, with some additional sites, if it is cut at a height of approximately 0.75 (**Figure A-2**), a relatively high dissimilarity value. Three site pairs were consistent between distance metrics, were within the same basin, and were generally within 1 µg L^{-1} of each other when considering the time series of sampling events (NEST01:NEST02, NMB02:NMB03, and RUS01:RUS02). The pairs FNB01:FNB02 and SID01:SID02 were consistently clustered when using different distance metrics, and were within the same basins, but the chlorophyll-*a* measurements had a larger magnitude of difference than the chosen cut point (**Figure 6**). The other pairs that were consistently grouped between metrics were not within the same basin and were not clustered using the 1 µg L^{-1} cut point.

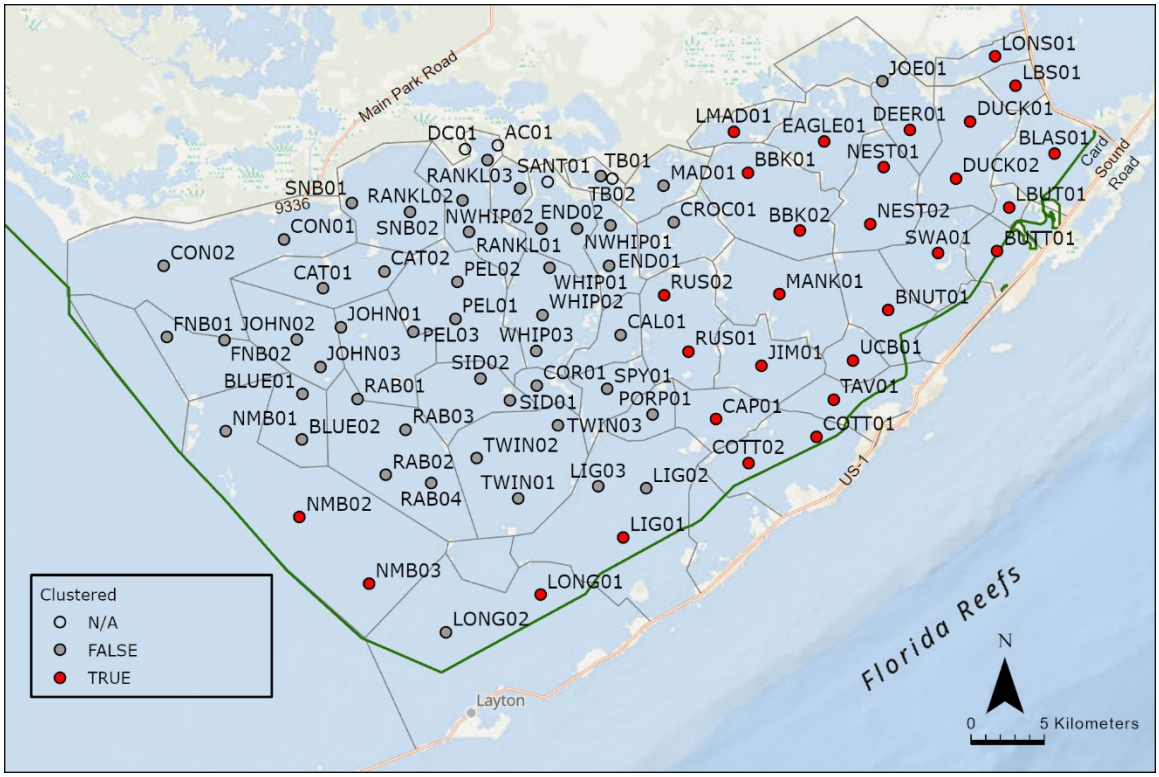


Figure 5. Results of Euclidean distance-based clustering with a cut point of 5.10, which is approximately equivalent to a difference of approximately 1 µg L^{-1} across the time series. Sites which were part of a cluster are red, locations which were not clustered with any other sites are grey, and sites which were not included in the analysis are unfilled circles. All clustered sites except the pair NMB02 and NMB03 were part of one large cluster. Sites SANT01, DC01, AC01, and TB02 were excluded from analysis.

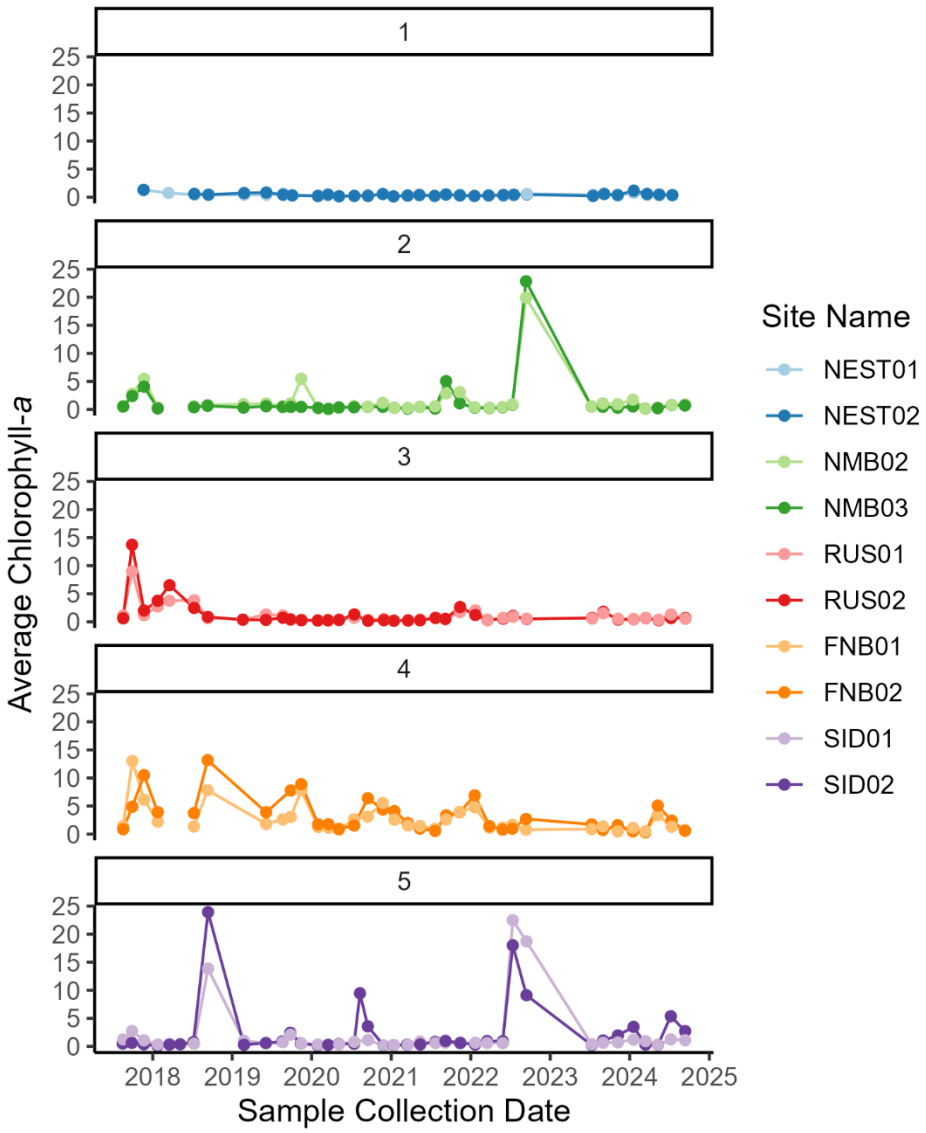


Figure 6. Chlorophyll-*a* concentration, representing an average of replicates per sample collection date, for potentially redundant site pairs (pairs 1–3) and pairs that were within the same basin but did not meet the Euclidean distance criteria (pairs 4 and 5).

Water Quality Parameter Similarity

The dendrograms resulting from the hierarchical agglomerative clustering using d_{kend} revealed that several parameters were consistently more similar to each other than to other parameters across sites (**Figure 7**). Total nitrogen (TN) and total organic carbon (TOC) generally had the lowest dissimilarity index, indicating that they were the most similar. Ammonia/ammonium (NH4) and nitrate+nitrite (NOX) were also consistently grouped across sites. Chlorophyll-*a* (CHLA) was often grouped in a subtree with total phosphate (TPO4), but this was not always consistent. To determine if parameters are suitably similar to consider discontinuing the measurement of a parameter using this approach, an appropriate cut point would have to be chosen based on system knowledge and program objectives (see Khalil et al. (2014) for an example of one such approach).

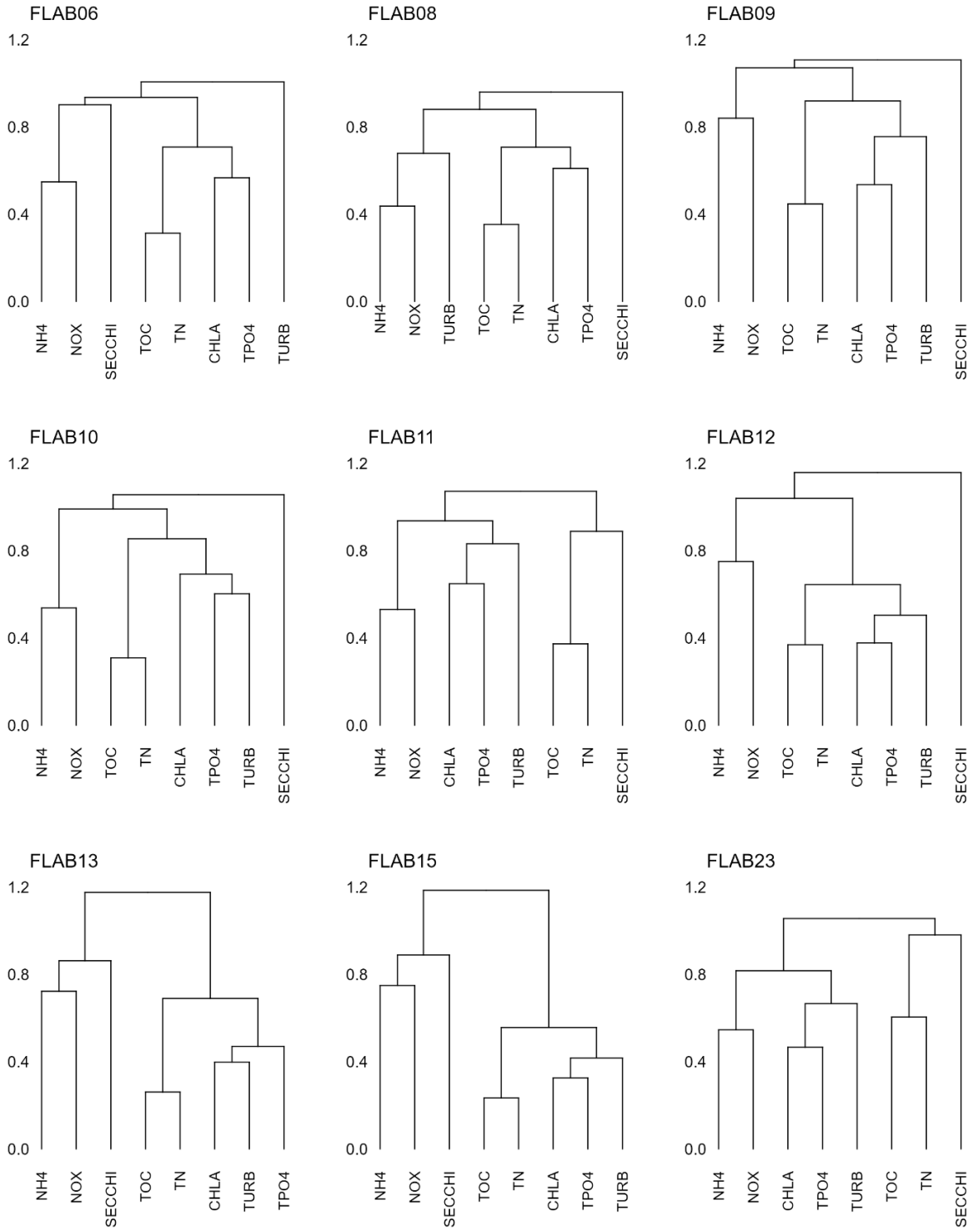


Figure 7. A dendrogram of cluster results of water quality parameters measured at South Florida Water Management District (SFWMD) sampling sites, using the Kendall's correlation distance. Increased height at the nodes where clusters join represents sites with a greater magnitude of difference between measurements in the time series.

Sampling Frequency

There were not significant differences in annual median predictions by site or region when assessed using Mann-Whitney U tests (**Figure 8** and **Figure A-1**). Notably, the maximum values obtained for each region and site often differed greatly depending on the sampling periodicity. The estimated annual median value also differed by ≥ 1 $\mu g\ L^{-1}$ in some cases.

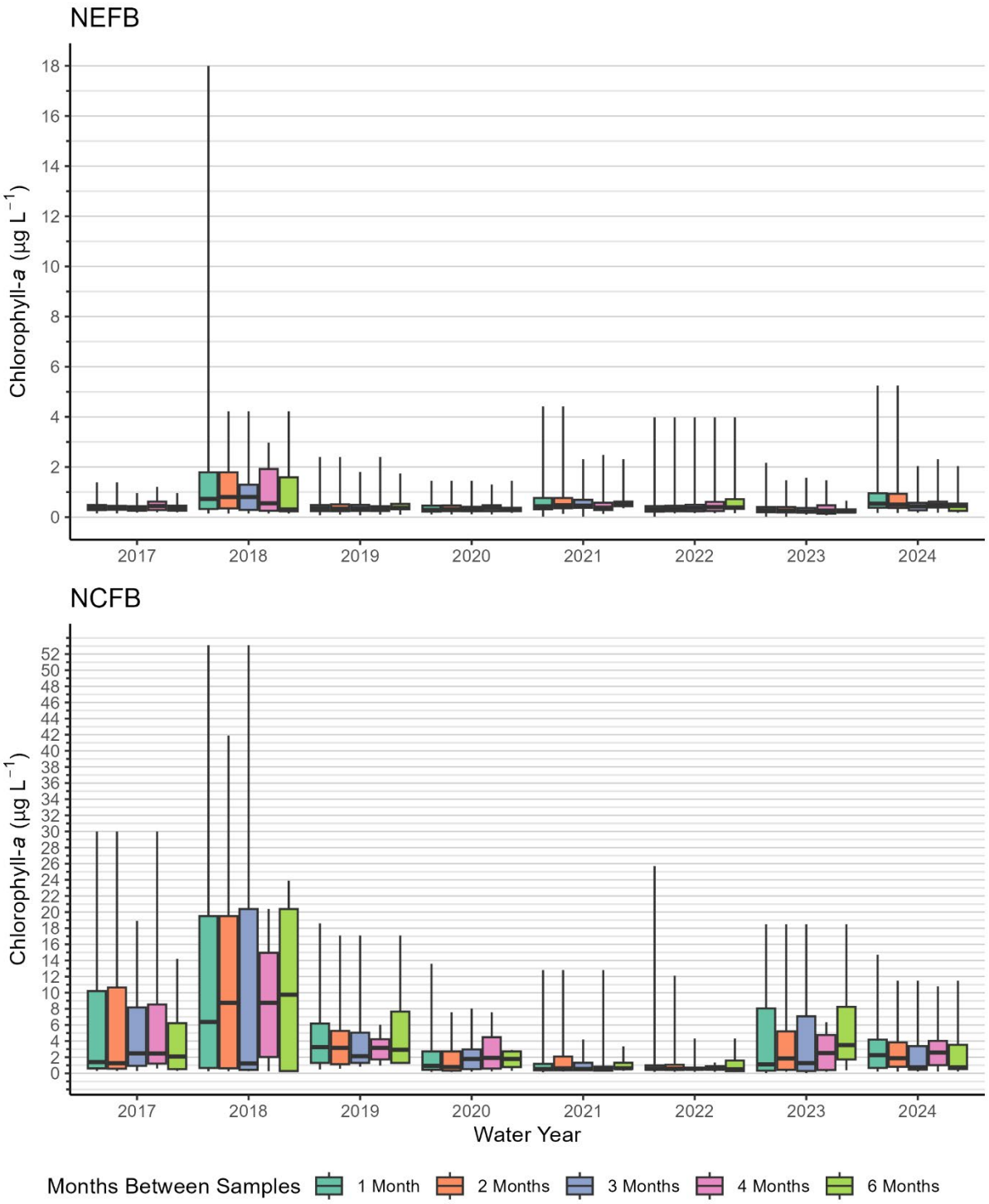


Figure 8. Chlorophyll-*a* concentrations, summarized by water year for Northeast Florida Bay (NEFB) and North Central Florida Bay (NCFB) which had two or more South Florida Water Management District (SFWMD) monthly sampling sites. Box and whisker plots represent the annual median, 25th and 75th percentiles, and the minimum and maximum observed values for each water year. The original monthly data was sub-sampled to simulate sample collection with periodicities of one, two, three, four, and six months.

Discussion

Sample Site Similarity

There were three pairs of sampling sites (NEST01:NEST02, NMB02:NMB03, and RUS01:RUS02) which were consistently paired with hierarchical agglomerative clustering using both distance metrics, were located within the same basin, and were clustered based on the chosen Euclidean distance cut point representing an approximate difference of 1 $\mu g\ L^{-1}$ per sampling event. Other sites which were consistently paired using both distance metrics were either not within the same basin or were not clustered based on the chosen Euclidean distance cut point. Based on these results, a conservative conclusion is that sites within the three pairs identified as similar (NEST01:NEST02, NMB02:NMB03, and RUS01:RUS02) appear to be collecting redundant chlorophyll-*a* data, and **one site of each pair may be considered for removal or replacement**.

Many sites in eastern Florida Bay formed a large cluster using the chosen cut point for Euclidean distance but were only clustered using the Kendall's correlation distance metric if a cut point of approximately 0.75 was used, a relatively high dissimilarity value. If further reduction of sampling effort is of interest, the cluster of sites within the eastern portion of Florida Bay may be candidates. However, removal of sites in eastern Florida Bay, where there may be only one site per basin, should be considered in the context of the desired spatial scale of sampling. Additionally, the raw measurements at any site being considered for reduction should be carefully evaluated, to determine if any one site is capturing critical information which other sites are not.

Because sites within the same basin were often not paired together in clustering results, sites within basins defined by Cosby et al. (2004) should not be assumed to be more similar to each other than to sites in other basins. The large cluster of sites identified using the Euclidean distance metric was similar in spatial coverage to an "Eastern Florida Bay" region identified in a previous multivariate PCA clustering analysis that used water quality samples collected between 1989-1995 (Boyer et al. 1997). It is likely that identified regions of similarity within the bay depend on the water quality parameters, or combination of parameters, being considered. Prior to altering the spatial sampling design, **a site similarity analysis should be repeated with any other water quality parameters** determined to be of key interest to meet program objectives.

There were extended gaps between some ENP sampling events in this analysis, due to missing data from several or many sites, resulting in a final dataset of 26 sampling events at 77 sites. The confidence in conclusions drawn from these results would be greater with a larger sample size, although sites which were most similar when considering Euclidean distance in the eastern area of Florida Bay will likely remain similar, unless the region is impact by an event (e.g., hurricane) which will have variability in spatial effects.

The current spatial sampling design appears to represent targeted, or "judgement" sampling: a subjective selection of what are thought to be representative population units, based on system knowledge by the program designer(s). If the underlying knowledge is sufficient and good choices of sampling units were made, parameters estimated from sample data may be accurate. However, the

accuracy of this type of design is difficult to assess (Gilbert 1987). Initial site selection was not random and site revisit selection for each sampling period is not randomized (follows an "always revisit" design). Because site revisit selection is not randomized each sampling period, the design violates assumptions of many probability-based condition assessments (Kopp and Neckles 2009), but the current "always revisit" spatial design is a suitable strategy for assessment of seasonality and trend. A probabilistic sampling plan such as stratified random sampling, spatially-balanced sampling, or systematic sampling with a randomized start point, would provide more accurate assessments of condition (Tetra Tech Inc. 1997; Kopp and Neckles 2009). **The sampling design should reflect program objectives, and it is likely the design will need to be altered to meet identified program objectives.** Water quality data obtained from the current program may be useful to establish sampling strata for use in a stratified random sampling or spatially-balanced sampling design.

Water Quality Parameter Similarity

Hierarchical agglomerative clustering of water quality parameters measured at SFWMD sites revealed that some were consistently more similar to each other than to other parameters, across sampling sites (TN:TOC and NH4:NOX). However, a large proportion of NOX values were below laboratory detection limits (22–95% of measurements per site), which may have influenced correlation results. To determine if parameters are suitably similar to each other to consider discontinuing the measurement of any, **an acceptable level of similarity between parameters must be chosen based on system knowledge and program objectives** (Khalil et al. 2014).

Additionally, when choosing to continue or discontinue the monitoring of water quality parameters, the analytical detection limits of available methods relative to expected background concentrations should be considered. Chlorophyll-*a* was almost always above detection limits in both datasets and the proportion of censored values which were present would likely not present an obstacle for statistical analyses. TN, TPO4, and TOC were also always above laboratory detection limits for ENP's dataset. However, for some parameters more than 10% of measurements were below detection limits in the ENP dataset: SRP (85%), NOX (52%) and NO2 (15%). This presents a challenge for many statistical techniques as well as for the interpretation of annual status estimates. Highly censored data can produced biased estimates of annual means (Helsel et al. 2020), and less than 25% of data must be below detection limits to know the median and interquartile range of a sample (Helsel and Hirsch 2002). When applying a Mann Kendall tests with a Theil-Sen slope estimate (as in Millette et al. (2019)), highly censored data (> 5% of records) is likely to result in errant estimates of the magnitude of the slope (Helsel and Hirsch 2002). Ideally, 5% or less of records should be below laboratory detection limits. It is recommended that program managers **consider if it is worthwhile to continue measuring parameters whose concentrations are typically too low to detect**. Alternatively, program managers can **explore if there are alternative laboratory methods available** which may be capable of detecting lower concentrations.

At the time of writing MDLs are not included within the structure of the Florida Bay water quality grab sample database. A future update to the database should **introduce columns for laboratory detection limits and data quality flags which may be associated with each measured value**.

Laboratory values should be recorded exactly as returned, rather than replacing values below detection limits with a categorical entry (e.g., "< 5 µmol L^{-1}").

Sampling Frequency

There was no evidence of a statistically significant difference between estimates of annual median values obtained by simulating different sampling frequencies. However, estimates did sometimes differ by ≥ 1 µg L^{-1} and the maximum concentration detected during a particular water year at a site or region was dependent on sampling frequency. **The most appropriate sampling frequency is dependent on program objectives**, alongside decisions about the acceptable levels of precision in annual estimates of central tendency and desired statistical power (Kupper and Hafner 1989; Tetra Tech Inc. 1997; Khalil et al. 2014). It is worth noting that some approaches which may be of interest, such as seasonal and long-term trend analyses, require large sample sizes to draw robust conclusions (Hirsch and Slack 1984). Fragmentation of the data by introducing large gaps reduces the ability to use the data for analyses such as trend assessment, which requires frequent sampling over many years (Lettenmaier et al. 1982; Hirsch and Slack 1984). The current monitoring program to date and as is will not be appropriate for these methodologies.

Most water quality parameters exhibited evidence of seasonality, trend, or both at some or all sites. The frequency and timing of sampling for parameters exhibiting cyclical behavior should also be based upon pre-existing knowledge of underlying population characteristics, as the frequency and timing of sampling can impact estimates such as annual means (Gilbert 1987). The sampling periodicity will also impact whether the peak intensity of algal bloom events is captured. Low frequency sampling may result in the start of algal blooms, or entire ephemeral events, being missed. Determination of adequate sampling frequency will be dependent on program objectives. Based on these points, **it is likely that the sampling frequency should not be reduced, and should possibly be increased, but will ultimately depend on program objectives**.

Program Objectives Inform Program Design

Water monitoring programs may be classified into broad categories based on their objectives: surveillance, model parameterization, cause-effect, trend detection, water quality control, and baseline (Lettenmaier et al. 1982). These broad categories generally require different sampling strategies in terms of sampling frequency and spatial coverage (Lettenmaier et al. 1982; Urquhart et al. 1998). For example, in trend detection programs spatial density may be sacrificed so long sequencies of high frequency observations necessary for statistical techniques may be reasonably obtained (Lettenmaier et al. 1982).

It is crucial that measurable program objectives be well-defined and documented prior to introducing changes to the sampling design (Fancy et al. 2009). The established objectives should be **S**pecific, **M**easurable, **A**chievable, **R**esults-oriented, and **T**ime-fixed (**SMART**; Williams et al. (2009)). To develop SMART objectives it may be useful to create a hierarchy of fundamental objectives, which are the core program outcomes, and means objectives, which are how the core outcomes will be achieved (Reynolds et al. 2016). **It is recommended that a workshop be held with program**

managers and internal stakeholders to outline program objectives and discuss the level of spatial and temporal detail necessary to address them.

A preliminary meeting to discuss program goals converged on several possible fundamental objectives:

1. Determine the drivers of algal blooms in Florida Bay,
2. Use alternative data sources (e.g., sonde data, satellite imagery) to detect bloom formation followed with responsive sampling to monitor the spatial and temporal extent and intensity of the bloom, or
3. Establish baseline water quality criteria at an appropriate spatial scale, followed by less frequent but routine monitoring at set intervals.

Possible objective one is best-served by a time-limited research study alongside a monitoring network designed to assess seasonality and trend. The drivers of cyanobacterial algal blooms in Florida Bay are currently being investigated by partners at FIU under a cooperative agreement which may or may not fully meet this objective in the coming year. The second possible objective generally describes a surveillance program, while the third generally represents a baseline program (Lettenmaier et al. 1982).

If the objective agreed upon is related to surveillance to detect algal bloom formation, followed by responsive sampling, implementing a hybrid approach using remote sensing and in-situ data, such as the program designed by Haakonsson et al. (2024), may be most appropriate. Using initial in-situ grab sampling, sonde measurements, and modeling, the authors developed a continuous modeling approach using sonde data (conductivity, temperature, depth) to predict favorable conditions for harmful cyanobacterial bloom formation. The formation and spatial extent of algal blooms was concurrently monitoring using satellite chlorophyll-*a* imagery. By combining the two approaches, field evaluations of algal blooms are initiated when certain conditions are met (Haakonsson et al. 2024). This design is ideal for early warning and responsive sampling but would not be appropriate for probabilistic condition assessments or evaluations of seasonality and long-term trends of water quality.

If the goal is more related to tracking water quality status, the United States Environmental Protection Agency (U.S. EPA) National Aquatic Resources Surveys (NARS) National Coastal Condition Assessment (NCCA) surveys is an example of a spatially-representative sampling design (U.S. EPA 2020). These are grab sampling surveys, conducted once every 5 years following a probabilistic, spatially-balanced sampling design. This design is capable of providing information on the regional status of coastal resources, at the time of sampling, with a known level of confidence (U.S. EPA 2020). Conclusions may be drawn about changes in water quality at a regional level, and compared to previous sampling periods, but long-term trends and seasonality of monitored parameters cannot be assessed.

If investigating long-term trends or seasonality is of interest, a fixed-site sampling approach with higher-frequency sampling is likely appropriate. The National Oceanic Atmospheric Administration (NOAA) National Estuarine Research Reserve System (NERRS) is an example of one such program (NERRS 2011). The goal of the NERRS is to assess short-term variability and long-term changes in estuarine systems to support management. Time series data such as this may also be used to characterize bloom and non-bloom conditions at a locality, by characterizing conditions which represent a departure from normal seasonal conditions. However, long time series of frequent measurements are required to draw accurate conclusions concerning seasonality and long-term trends (Hirsch and Slack 1984; Carstensen et al. 2007; Millette et al. 2019).

A hybrid program design, incorporating multiple sampling designs, may be best to address multiple program objectives. The Southeast Coast Network (SECN) parks combine two sampling approaches to satisfy multiple objectives. Fixed-site monitoring based on the NOAA NERRS protocols, using sondes and grab samples, provide temporally representative data for determining diel, monthly, and seasonal water-quality patterns. A probabilistic, spatially-balanced sampling design with a fixed return interval based on the U.S. EPA NCCA protocol provides a spatially representative data. The number of sites selected for spatially representative sampling fulfills both statistical and practical considerations, while selection of fixed-site sites was based on both environmental factors (e.g., habitat type, resource concerns) and practical considerations (e.g., accessibility; Starkey et al. (2023)).

Additionally, it is noted that to design an optimal sampling network, the SECN parks met with local and national collaborators and stakeholders who provided input on the monitoring program design, so the program data would be useful for multiple parties and as compatible as possible with national programs (Starkey et al. 2023). We suggest that **a local workshop or coordination meeting may lead to fruitful collaborations and enhance the research and management of algal blooms in Florida Bay**. At minimum, it is recommended that the water quality sampling activities of other agencies in Florida Bay (e.g., SFWMD, U.S. EPA) be considered. If data collected by partner agencies is comparable to ENP data, or may be used to meet program objectives, it may be that additional sampling by ENP would be redundant.

Program Documentation

A monitoring protocol should be created as program goals and objectives are enacted. The formatting of this protocol should follow best-practices as established by National Park Service inventory and monitoring programs (Oakley et al. 2003; Fancy et al. 2009) and current organizational guidelines for the development of project documentation. At a minimum, the documentation should include:

- Background information,
- Questions or goals addressed by the program,
- Measurable (SMART) objectives,
- Sampling design (spatial and temporal),

- Field and laboratory methods,
- Data management strategy,
- Data analysis and reporting plan, and
- Plan for periodic re-evaluation to assess if programmatic goals are being met.

The periodicity of program re-evaluation should be determined in part from program objectives. For example, if a program goal is to estimate long-term change in water quality parameters, adjusting the location of sampling sites or introducing temporal gaps in sampling within a short 5-year time span would be disadvantageous.

Other Considerations

Several practical and technical considerations arose during this evaluation. If not already performed, continual checks of the quality of fluorometric data (e.g., routine checks of detection limit, precision analysis) should be performed as outlined in (Arar and Collins 1997). The high variability between replicates of both water quality and nutrient parameters collected by ENP are of concern, as samples were collected at the same time and place (field duplicates).

ENP has several sonde sensors, measuring various parameters at fixed hydro sites in Florida Bay. These may be incorporated into a water quality monitoring network design but have not been evaluated here. They may be best served as complements to other monitoring efforts in a hybrid design.

Finally, **practical components of the monitoring program (e.g., staffing, workload, logistics, cost…) must be considered during the discussion of program objectives and sampling design** to ensure any updates to the program's sampling design are reasonable. However, such practical limitations are not covered in the scope of this document.

Literature Cited

Arar, E. J. 1997. Method 447.0: Determination of chlorophylls *a* and *b* and identification of other pigments of interest in marine and freshwater algae using high performance liquid chromatography with visible wavelength detection. National Exposure Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Cincinnati, Ohio.

Arar, E. J., and G. B. Collins 1997. Method 445.0 *In vitro* determination of chlorophyll *a* and pheophytin *a* in marine and freshwater algae by fluorescence. National Exposure Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Cincinnati, Ohio.

Behmel, S., M. Damour, R. Ludwig, and M. J. Rodriguez. 2016. Water quality monitoring strategies - A review and future perspectives. Science of the Total Environment 571:1312-1329. https://doi.org/10.1016/j.scitotenv.2016.06.235

Bezdek, J. C., and R. J. Hathaway. 2002. VAT: A tool for visual assessment of (cluster) tendency. Proceedings of the 2002 International Joint Conference on Neural Networks (Cat. No. 02CH37290). IEEE. 2225-2230.

Boyer, J. N., J. W. Fourqurean, and R. D. Jones. 1997. Spatial characterization of water quality in Florida Bay and Whitewater Bay by multivariate analyses: Zones of similar influence. Estuaries 20(4):743-758.

Boyer, J. N., C. R. Kelble, P. B. Ortner, and D. T. Rudnick. 2009. Phytoplankton bloom status: Chlorophyll *a* biomass as an indicator of water quality condition in the southern estuaries of Florida, USA. Ecological Indicators 9S:S56-S67. https://doi.org/10.1016/j.ecolind.2008.11.013

Brown, J., D. Gierhart, J. Moorman, and D. Francisco 2024. Operational Project Monitoring Plan for Florida Bay Water Quality Monitoring Network (FLAB). SFWMD-FIELD-MP-079-08, South Florida Water Management District.

Carstensen, J., P. Henriksen, and A.-S. Heiskanen. 2007. Summer algal blooms in shallow estuaries: Definition, mechanisms, and link to eutrophication. Limnology and Oceanography 52(1):370-384. https://doi.org/10.4319/lo.2007.52.1.0370

Cosby, B., W. Nuttle, and J. Fourqurean 2004. FATHOM - Flux accounting and tidal hydrology at the ocean margin: Florida Bay simulation 1991-2002 updated model and preliminary results. Interim project report for Everglades National Park. Florida International University.

Cosby, B., W. Nuttle, and F. Marshall 2005. Final Report FATHOM enhancements and implementation to support development of minimum flows and levels for Florida Bay. Contract C-C-15975-WO05-05. Environmental Consulting & Technology, Inc., New Symyrna Beach, Florida, USA.

Fancy, S. G., J. E. Gross, and S. L. Carter. 2009. Monitoring the condition of natural resources in US National Parks. Environmental Monitoring and Assessment 151:161-174. https://doi.org/10.1007/s10661-008-0257-y

FDEP (Florida Department of Environmental Protection). 2011. Applicability of chlorophyll *a* methods. DEP-SAS-002/10, Florida Department of Environmental Protection.

Fratto, Z. 2024. Baywide standard operation procedure surface water sampling. South Florida Natural Resources Center, National Park Service, Florida Bay Interagency Science Center.

Gilbert, R. O. 1987. Statistical methods for environmental pollution monitoring. Van Nostrand Reinhold Company, New York.

Haakonsson, S., F. Maciel, M. A. Rodríguez, L. Ponce de León, L. Rodríguez-Gallego, R. Arocena, F. Pedocchi, and S. Bonilla. 2024. Monitoring cyanobacterial blooms: a strategy combining predictive modeling and remote sensing approaches. Environmental Earth Sciences 83(195). https://doi.org/10.1007/s12665-024-11488-3.

Helsel, D. R., and R. M. Hirsch 2002. Statistical methods in water resources. Techniques of Water-Resources Investigations, U.S. Geological Survey, Reston, VA. https://doi.org/10.3133/twri04A3

Helsel, D. R., R. M. Hirsch, K. R. Ryberg, S. A. Archfield, and E. J. Gilroy 2020. Statistical methods in water resources. Techniques and Methods, U.S. Geological Survey, Reston, VA. https://doi.org/10.3133/tm4a3.

Hirsch, R. M., and J. R. Slack. 1984. A nonparametric trend test for seasonal data with serial dependence. Water Resources Research 20(6):727-732. https://doi.org/10.1029/WR020i006p00727

Kassambara, A. 2017. Practical guide to cluster analysis in R: Unsupervised machine learning. 1st edition. STHDA.

Khalil, B., C. Ou, S. Proulx-McInnis, A. St-Hilaire, and E. Zanacic. 2014. Statistical assessment of the surface water quality monitoring network in Saskatchewan. Water, Air, & Soil Pollution 225:2128. https://doi.org/10.1007/s11270-014-2128-1

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. Natural Resource Report NPS/NCBN/NRR—2009/110, U.S. Department of the Interior, National Park Service, Natural Resource Program Center, Fort Collins, Colorado.

Kupper, L. L., and K. B. Hafner. 1989. How appropriate are popular sample size formulas? The American Statistician 43(2):101-105. https://doi.org/10.1080/00031305.1989.10475628

Lee, T. N., N. Melo, N. Smith, E. M. Johns, C. R. Kelble, R. H. Smith, and P. B. Ortner. 2016. Circulation and water renewal of Florida Bay, USA. Bulletin of Marine Science 92(2):153-180. http://dx.doi.org/10.5343/bms.2015.1019

Lettenmaier, D. P., L. L. Conquest, and J. P. Hughes 1982. Routine streams and rivers water quality trend monitoring review. Water Resources Series, Technical Report No. 75, Department of Civil Engineering, Seattle, Washington.

Maechler, M., P. Rousseeuw, A. Struyf, M. Hubert, and K. Hornik. 2023. cluster: Cluster analysis basics and extensions. R package version 2.1.6.

Millette, N. C., C. Kelble, A. Linhoss, S. Ashby, and L. Visser. 2019. Using spatial variability in the rate of change of chlorophyll *a* to improve water quality management in a subtropical oligotrophic estuary. Estuaries and Coasts 42:1792-1803. https://doi.org/10.1007/s12237-019-00610-5

Naja, G. M., M. Collier, J. Nestler, E. Stabenau, T. Dean, D. Surratt, J. Kline, M. Parry, L. Oberhofer, M. Tupaj, J. Osborne, H. Cooley, and B. Falk 2022. Everglades National Park: 2022 State of Conservation. South Florida Natural Resources Center, Everglades National Park, Homestead, FL.

NERRS (National Estuarine Reserach Reserve System). 2011. System-wide monitoring program plan. National Estuarine Research Reserve System, National Oceanic and Atmospheric Administration. www.nerrs.noaa.gov

Oakley, K. L., L. P. Thomas, and S. G. Fancy. 2003. Guidelines for long-term monitoring protocols. Wildlife Society Bulletin 31(4):1000-1003.

Reynolds, J. H., M. G. Knutson, K. B. Newman, E. D. Silverman, and W. L. Thompson. 2016. A road map for designing and implementing a biological monitoring program. Environmental Monitoring and Assessment 188:1-25. https://doi.org/10.1007/s10661-016-5397-x

Soares, J., P. A. Makar, Y. Aklilu, and A. Akingunola. 2018. The use of hierarchical clustering for the design of optimized monitoring networks. Atmospheric Chemistry and Physics 18:6543-6566. https://doi.org/10.5194/acp-18-6543-2018

Starkey, E. N., M. B. Gregory, W. K. Wright, J. C. DeVivo, E. M. DiDonato, C. J. Wright, and E. Thompson 2023. Protocol for monitoring estuarine water and sediment quality in selected Southeast Coast Network parks: Protocol narrative—version 2.0. Natural Resource Report, NPS/SECN/NRR—2023/2528, Fort Collins, Colorado. https://doi.org/10.36967/2299077.

Tetra Tech Inc. 1997. Sampling Design. Pages 1-27 *in* Techniques for tracking, evaluating, and reporting the implementation of nonpoint source control measures - Forestry. U.S. Environmental Protection Agency.

Turner Designs Inc. Trilogy® Laboratory Fluorometer. https://docs.turnerdesigns.com/t2/doc/brochures/S-0068.pdf. (accessed 25 February 2025).

U.S. EPA (United States Environmental Protection Agency). 2020. National Coastal Condition Assessment: Site evaluation guidelines. EPA-841-B-20-001, Washington, DC.,

Urquhart, N. S., S. G. Paulsen, and D. P. Larsen. 1998. Monitoring for policy-relevant regional trends over time. Ecological Applications 8(2):246-257. https://doi.org/10.1890/1051-0761(1998)008[0246:MFPRRO]2.0.CO;2

Williams, B. K., R. C. Szaro, and C. D. Shapiro. 2009. Adaptive management: The U.S. Department of the Interior technical guide. Adaptive Management Working Group, U.S. Department of Interior, Washington, DC.

Appendix A. Supporting Tables and Figures

Table A-1. A list of standard methods used by Florida International University (FIU) to measure nutrient analytes in Everglades National Park water quality grab samples. The method detection limit(s) (MDL) of the laboratory analyses methods during the period of analysis are provided. For variables with varying MDLs through time, the MDL is presented as a range capturing the minimum and maximum value during the period of analysis.

Analyte	Abbreviation	Standard Method	Method Detection Limit(s) in µmols/L
Nitrite+Nitrate	NOX	EPA353.4 Rev. 2.0	0.03–0.13
Nitrite	NO2	EPA353.4 Rev. 2.0	0.01–0.02
Ammonia/Ammonium	NH4	EPA349.0 Rev. 2.0	0.09–0.179
Soluble Reactive Phosphorus	SRP	EPA365.1 Rev. 2.0	0.01–0.046
Total Phosphorus	TPO4	EPA365.1 Rev. 2.0	0.015–0.03
Dissolved Silicate	Si	EPA366.0 Rev. 1.0	0.05–0.20
Total Organic Carbon	TOC	SM5310B-1999	5.0–10.40
Total Nitrogen	TN	ASTM D5176-91	1.43–3.63

Table A-2. Proportion missing (NA) values at each site, and overall, for Everglades National Park Florida Bay water quality nutrient samples collected between August 2017 and September 2024. A sample measurement was considered missing if both replicates for a parameter were missing on a given sampling event. It's noted that some sites with a higher proportion of missing values (e.g., CON02 and FNB02) are sites where nutrient collection started at a later date than many of the other sites, where there were few overall samples.

Site	NOX	NO2	NH4	TN	TPO4	SRP	TOC	Si
BLUE01	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
BLUE02	9.09%	9.09%	6.06%	0.00%	0.00%	30.30%	0.00%	3.03%
CAL01	8.57%	2.86%	2.86%	0.00%	0.00%	20.00%	0.00%	5.71%
CAT02	9.09%	9.09%	3.03%	0.00%	0.00%	18.18%	0.00%	3.03%
CON01	5.88%	2.94%	2.94%	0.00%	0.00%	14.71%	0.00%	5.88%
CON02	14.29%	14.29%	14.29%	0.00%	0.00%	14.29%	0.00%	14.29%
COR01	12.12%	6.06%	3.03%	0.00%	0.00%	18.18%	0.00%	3.03%
CROC01	5.88%	5.88%	2.94%	0.00%	0.00%	23.53%	0.00%	2.94%
END02	2.94%	8.82%	2.94%	0.00%	0.00%	17.65%	0.00%	5.88%
FNB02	16.67%	16.67%	16.67%	16.67%	16.67%	16.67%	16.67%	16.67%
JOE01	2.94%	2.94%	2.94%	0.00%	0.00%	17.65%	0.00%	5.88%
JOHN01	12.12%	9.09%	3.03%	0.00%	0.00%	24.24%	0.00%	3.03%
JOHN02	16.67%	16.67%	16.67%	0.00%	0.00%	16.67%	0.00%	0.00%
JOHN03	2.86%	5.71%	2.86%	0.00%	0.00%	20.00%	0.00%	5.71%
LIG02	8.82%	5.88%	2.94%	0.00%	0.00%	23.53%	0.00%	2.94%
LIG03	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
LMAD01	2.94%	5.88%	2.94%	0.00%	0.00%	20.59%	0.00%	5.88%
LONS01	16.67%	16.67%	16.67%	0.00%	0.00%	16.67%	0.00%	0.00%
MAD01	16.67%	16.67%	16.67%	0.00%	0.00%	16.67%	0.00%	0.00%
NWHIP02	11.76%	5.88%	2.94%	0.00%	0.00%	17.65%	0.00%	5.88%
PEL01	14.29%	14.29%	14.29%	0.00%	0.00%	14.29%	0.00%	0.00%
PEL02	12.12%	6.06%	3.03%	0.00%	0.00%	18.18%	0.00%	3.03%
PEL03	15.15%	6.06%	3.03%	0.00%	0.00%	18.18%	3.03%	3.03%
RAB01	5.88%	8.82%	2.94%	0.00%	0.00%	20.59%	0.00%	2.94%
RAB02	8.82%	8.82%	2.94%	0.00%	0.00%	23.53%	0.00%	2.94%
RANKB01	14.29%	28.57%	14.29%	0.00%	0.00%	28.57%	0.00%	0.00%
RANKL01	14.71%	5.88%	2.94%	0.00%	0.00%	17.65%	0.00%	5.88%
RANKL02	16.67%	16.67%	16.67%	0.00%	0.00%	16.67%	0.00%	0.00%
RANKL03	11.76%	5.88%	2.94%	0.00%	0.00%	14.71%	0.00%	5.88%
SID01	14.29%	14.29%	14.29%	0.00%	0.00%	14.29%	0.00%	14.29%

SID02	3.13%	9.38%	3.13%	0.00%	0.00%	21.88%	0.00%	0.00%
SNB01	16.67%	16.67%	16.67%	0.00%	0.00%	16.67%	0.00%	0.00%
SNB02	19.35%	6.45%	3.23%	0.00%	0.00%	16.13%	0.00%	3.23%
SPY01	8.57%	5.71%	2.86%	0.00%	0.00%	22.86%	0.00%	5.71%
TB01	8.82%	5.88%	2.94%	0.00%	0.00%	8.82%	0.00%	5.88%
TWIN01	16.67%	16.67%	16.67%	0.00%	0.00%	16.67%	0.00%	0.00%
TWIN02	5.88%	8.82%	2.94%	0.00%	0.00%	23.53%	0.00%	2.94%
TWIN03	8.82%	5.88%	2.94%	0.00%	0.00%	20.59%	0.00%	2.94%
WHIP01	16.67%	16.67%	16.67%	0.00%	0.00%	16.67%	0.00%	0.00%
WHIP02	12.50%	9.38%	3.13%	0.00%	0.00%	18.75%	0.00%	3.13%
WHIP03	14.29%	14.29%	14.29%	0.00%	0.00%	14.29%	0.00%	0.00%
All Sites	9.38%	7.51%	4.17%	0.10%	0.10%	19.40%	0.21%	4.17%

Table A-3. Percentage of values at each site, and overall, which were less than or equal to the method detection limit (MDL) of the laboratory analysis methods for Everglades National Park Florida Bay water quality grab samples collected between August 2017 and September 2024. The percentage of values below the MDL for each parameter was not consistent across sites, likely due to underlying environmental differences at each station.

Site	NOX	NO2	NH4	TN	TPO4	SRP	TOC	Si
AC01	83.33%	16.67%	0.00%	0.00%	0.00%	75.00%	0.00%	0.00%
BLUE01	100.00%	0.00%	0.00%	0.00%	0.00%	50.00%	0.00%	0.00%
BLUE02	70.69%	27.59%	11.67%	0.00%	0.00%	89.13%	0.00%	0.00%
CAL01	50.79%	11.94%	0.00%	0.00%	0.00%	87.04%	0.00%	0.00%
CAT02	55.93%	16.67%	18.75%	0.00%	0.00%	92.59%	0.00%	0.00%
CON01	30.65%	21.54%	23.08%	0.00%	0.00%	52.63%	0.00%	0.00%
CON02	36.36%	0.00%	18.18%	0.00%	0.00%	54.55%	0.00%	0.00%
COR01	67.24%	11.29%	6.25%	0.00%	0.00%	92.16%	0.00%	0.00%
CROC01	26.56%	0.00%	0.00%	0.00%	0.00%	86.27%	0.00%	0.00%
DC01	50.00%	20.00%	0.00%	0.00%	0.00%	10.00%	0.00%	0.00%
END02	33.33%	3.28%	0.00%	0.00%	0.00%	89.09%	0.00%	0.00%
FNB02	100.00%	70.00%	40.00%	0.00%	0.00%	100.00%	0.00%	0.00%
JOE01	6.06%	4.55%	0.00%	0.00%	0.00%	77.78%	0.00%	0.00%
JOHN01	79.31%	43.33%	6.25%	0.00%	0.00%	80.00%	0.00%	0.00%
JOHN02	90.00%	50.00%	20.00%	0.00%	0.00%	100.00%	0.00%	0.00%
JOHN03	65.67%	16.92%	4.48%	0.00%	0.00%	87.04%	0.00%	0.00%
LIG02	22.58%	6.25%	1.52%	0.00%	0.00%	94.00%	0.00%	6.06%
LIG03	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	N/A
LMAD01	9.09%	0.00%	0.00%	0.00%	0.00%	77.36%	0.00%	0.00%
LONS01	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%
MAD01	20.00%	20.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%
NWHIP02	47.37%	13.11%	3.13%	0.00%	0.00%	96.15%	0.00%	0.00%
PEL01	27.27%	9.09%	18.18%	0.00%	0.00%	80.00%	0.00%	0.00%
PEL02	63.79%	20.97%	9.38%	0.00%	0.00%	90.74%	0.00%	0.00%
PEL03	75.00%	8.06%	3.13%	0.00%	0.00%	85.19%	0.00%	0.00%
RAB01	70.97%	8.06%	7.58%	0.00%	0.00%	73.58%	0.00%	0.00%
RAB02	70.49%	12.90%	4.55%	0.00%	0.00%	76.47%	0.00%	0.00%
RANKB01	63.64%	50.00%	18.18%	0.00%	0.00%	100.00%	0.00%	0.00%
RANKL01	64.91%	31.75%	10.94%	0.00%	0.00%	92.59%	0.00%	0.00%
RANKL02	80.00%	0.00%	20.00%	0.00%	0.00%	100.00%	0.00%	0.00%
RANKL03	61.02%	20.63%	9.09%	0.00%	0.00%	76.79%	0.00%	0.00%
SANT01	40.00%	30.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%
SID01	83.33%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%
SID02	59.32%	17.24%	1.61%	0.00%	0.00%	92.00%	0.00%	0.00%
SNB01	20.00%	0.00%	20.00%	0.00%	0.00%	90.00%	0.00%	0.00%
SNB02	80.00%	26.32%	17.24%	0.00%	0.00%	90.20%	0.00%	0.00%

SPY01	59.02%	20.31%	1.52%	0.00%	0.00%	90.00%	0.00%	0.00%
TB01	36.67%	15.87%	6.25%	0.00%	0.00%	86.67%	0.00%	0.00%
TB02	30.00%	20.00%	0.00%	0.00%	0.00%	90.00%	0.00%	0.00%
TWIN01	60.00%	0.00%	0.00%	0.00%	0.00%	90.00%	0.00%	0.00%
TWIN02	46.03%	9.68%	0.00%	0.00%	0.00%	87.76%	0.00%	0.00%
TWIN03	58.33%	7.81%	3.03%	0.00%	0.00%	90.20%	0.00%	0.00%
WHIP01	90.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%
WHIP02	69.09%	8.62%	6.45%	0.00%	0.00%	95.83%	0.00%	0.00%
WHIP03	66.67%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%
All Sites	52.06%	14.93%	6.35%	0.00%	0.00%	85.32%	0.00%	0.22%

Table A-4. Percentage of missing (N/A) monthly values at each site, and overall, for South Florida Water Management District Florida Bay water quality grab samples collected between May 2016 and October 2024.

Parameter	FLAB06	FLAB08	FLAB09	FLAB10	FLAB11	FLAB12	FLAB13	FLAB15	FLAB23	All Sites
NH4	2.08%	1.04%	2.08%	4.17%	3.13%	0.00%	1.04%	0.00%	2.08%	1.74%
TOC	1.04%	1.04%	2.08%	4.17%	2.08%	1.04%	1.04%	0.00%	1.04%	1.50%
CHLA	2.08%	2.08%	3.13%	4.17%	3.13%	1.04%	0.00%	0.00%	1.04%	1.85%
CHLB	52.08%	53.13%	53.13%	55.21%	53.13%	53.13%	52.08%	52.08%	52.08%	52.89%
DO	1.04%	1.04%	2.08%	4.17%	1.04%	0.00%	0.00%	0.00%	1.04%	1.16%
NOX	2.08%	1.04%	2.08%	4.17%	3.13%	0.00%	1.04%	0.00%	2.08%	1.74%
NO3	56.25%	56.25%	57.29%	56.25%	58.33%	55.21%	55.21%	56.25%	56.25%	56.37%
NO2	2.08%	3.13%	3.13%	5.21%	3.13%	0.00%	1.04%	0.00%	2.08%	2.20%
PHEO	2.08%	2.08%	3.13%	4.17%	3.13%	1.04%	0.00%	0.00%	1.04%	1.85%
OPO4	2.08%	3.13%	3.13%	5.21%	3.13%	0.00%	1.04%	0.00%	2.08%	2.20%
TPO4	1.04%	1.04%	2.08%	4.17%	2.08%	1.04%	1.04%	0.00%	1.04%	1.50%
SAL	1.04%	1.04%	2.08%	4.17%	1.04%	0.00%	0.00%	0.00%	1.04%	1.16%
SECCHI	2.08%	2.08%	3.13%	5.21%	2.08%	1.04%	1.04%	1.04%	2.08%	2.20%
TN	1.04%	1.04%	2.08%	4.17%	3.13%	1.04%	1.04%	0.00%	1.04%	1.62%
TURB	1.04%	1.04%	2.08%	4.17%	3.13%	0.00%	1.04%	0.00%	1.04%	1.50%
TEMP	1.04%	1.04%	2.08%	4.17%	1.04%	0.00%	0.00%	0.00%	1.04%	1.16%

Table A-5. Percentage of values at each station, and overall, which were less than or equal to the method detection limit (MDL) of the laboratory analysis method for South Florida Water Management District Florida Bay water quality samples collected between May 2016 and October 2024.

Parameter	FLAB06	FLAB08	FLAB09	FLAB10	FLAB11	FLAB12	FLAB13	FLAB15	FLAB23	All Sites
NH4	0.00%	0.00%	0.00%	1.08%	0.00%	0.00%	4.21%	1.04%	0.00%	0.70%
TOC	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
CHLA	0.00%	0.00%	1.06%	0.00%	1.06%	0.00%	0.00%	0.00%	1.04%	0.35%
CHLB	91.30%	95.56%	95.56%	95.35%	100.00%	84.44%	97.83%	76.09%	100.00%	92.87%
DO	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
NOX	49.47%	44.79%	22.11%	19.35%	59.57%	84.38%	94.74%	89.58%	23.16%	54.27%
NO3	67.44%	48.84%	42.86%	18.60%	87.80%	93.02%	100.00%	90.48%	25.58%	63.71%
NO2	56.84%	52.13%	34.04%	42.39%	48.94%	75.00%	91.58%	83.33%	41.05%	58.52%
PHEO	95.79%	83.16%	86.17%	45.16%	87.23%	47.37%	48.96%	12.50%	92.71%	66.51%
OPO4	98.95%	97.87%	97.87%	97.83%	98.94%	98.96%	96.84%	97.92%	93.68%	97.65%
TPO4	5.21%	6.25%	25.26%	0.00%	9.47%	0.00%	0.00%	0.00%	43.75%	10.04%
SAL	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
SECCHI	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
TN	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
TURB	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
TEMP	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

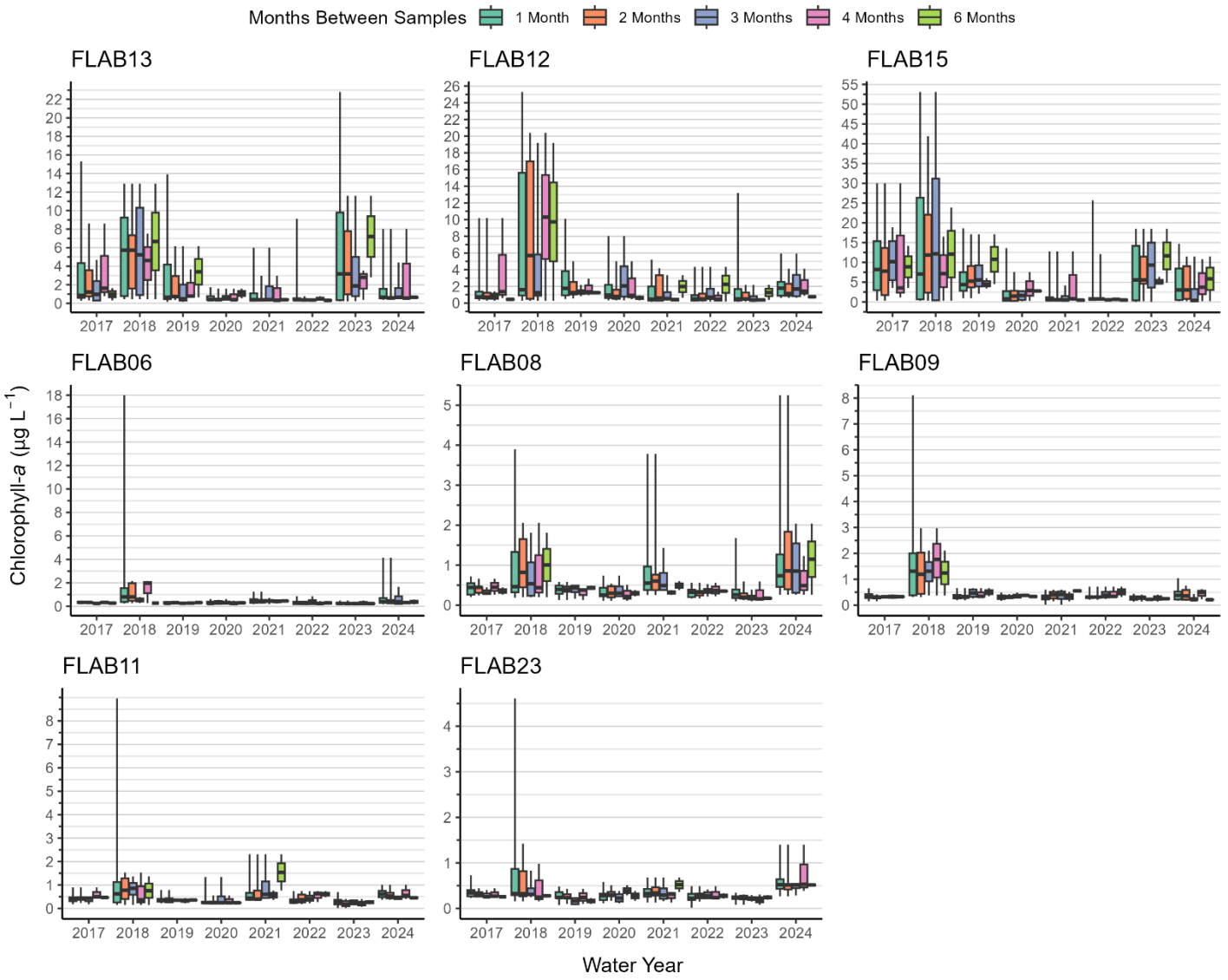


Figure A-1. Chlorophyll-a concentrations, summarized by water year for South Florida Water Management District (SFWMD) monthly sampling sites. Box and whisker plots represent the annual median, 25th and 75th percentiles, and the minimum and maximum observed values. The original monthly data was sub-sampled to simulate sample collection with periodicities of one, two, three, four, and six months.

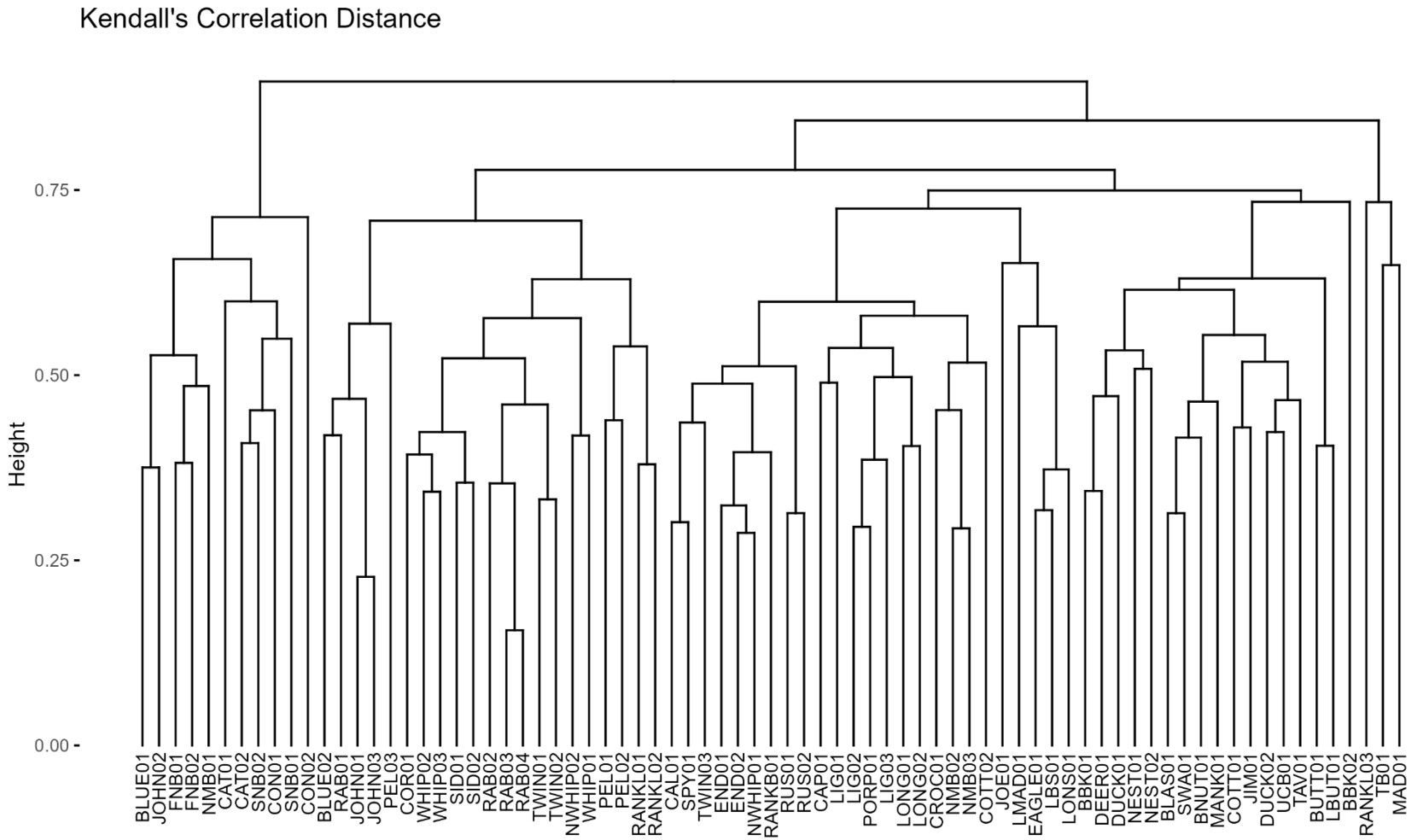


Figure A-2. A dendrogram of cluster results, using the Kendall's correlation distance metric. Increased height at the nodes where clusters join represents sites whose measurements are less correlated across the timeseries.

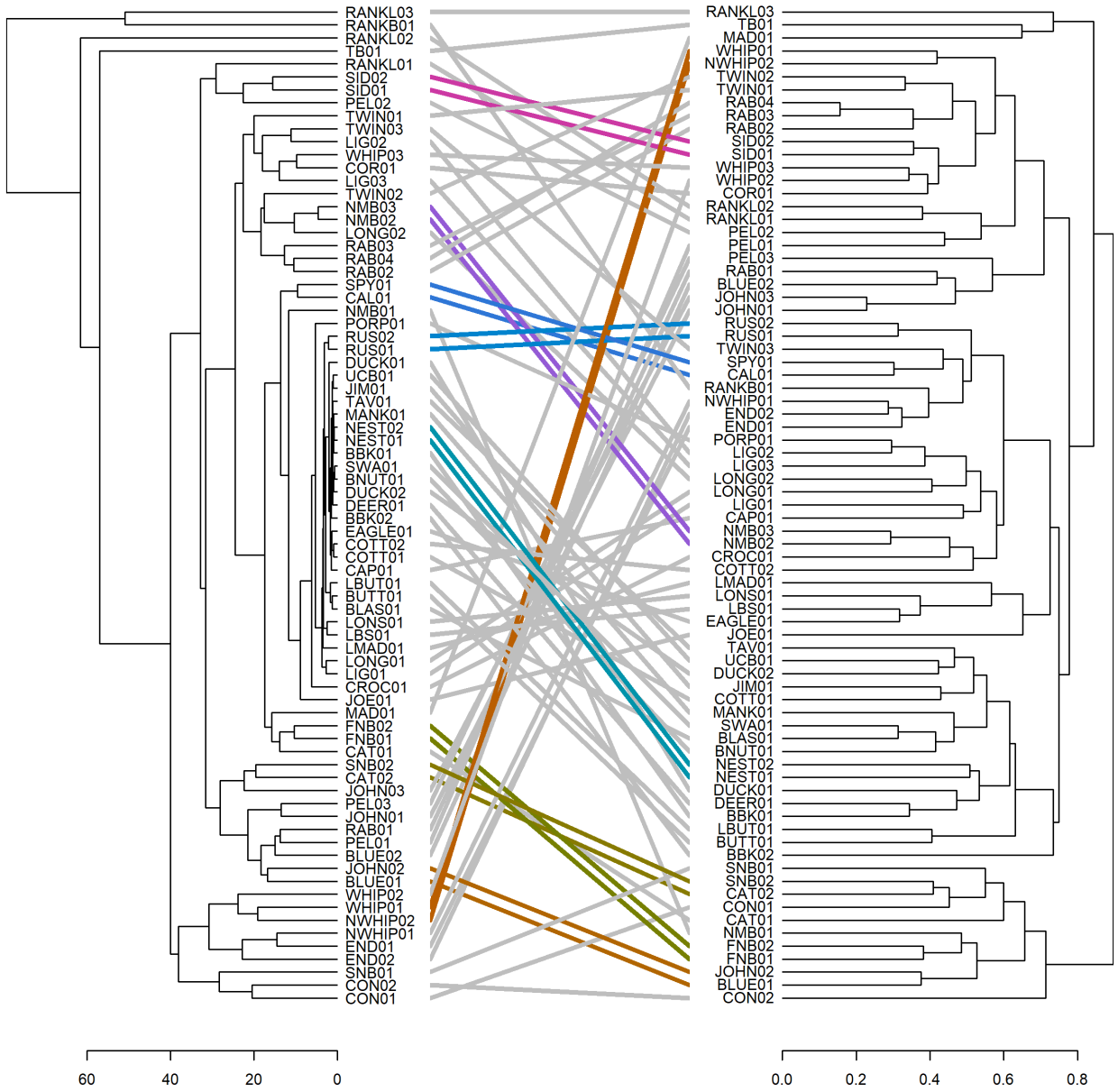


Figure A-3. A tanglegram of cluster results, showing similar subtrees (site pairs) between dendrograms created using the Euclidean distance metric (left) and Kendall's correlation distance metric (right). Similar subtrees are highlighted with colored lines linking them between dendrograms.

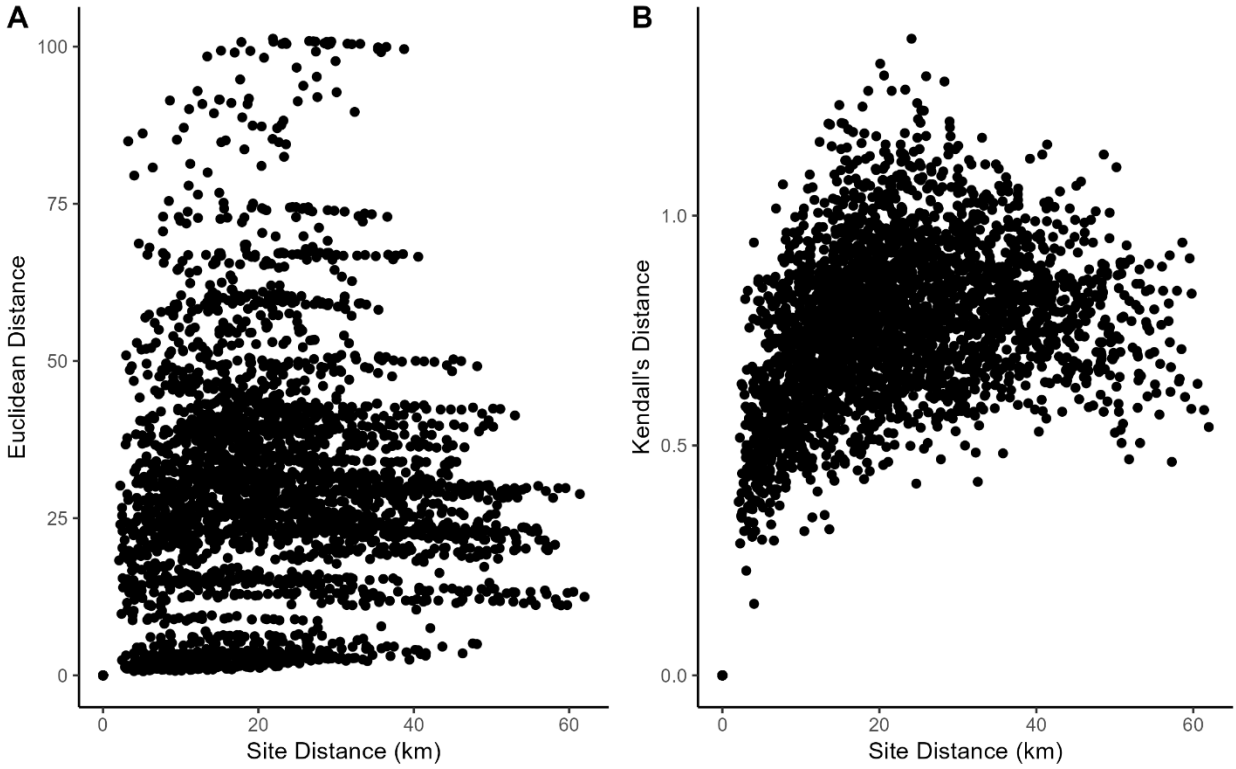


Figure A-4. Scatterplots of pairwise distance measurements of the distance between sites in kilometers (km) and the Euclidean distance and Kendall's distance metrics.