SEACAR Discrete Water Quality Analysis: Sample Surface Chlorophyll a uncorrected for pheophytin

Last compiled on 19 May, 2022

Contents

Libraries	1
File Import	2
Data Summary Before Filtering	2
Data Filtering and Data Impacted by Specific Value Qualifiers	3
Managed Area Statistics	6
Monitoring Location Statistics	7
Seasonal Kendall Tau Analysis	8
Appendix I: Scatter Plot of Entire Dataset	11
Appendix II: Dataset Summary Box Plots	13
Appendix III: Excluded Managed Areas	19
Appendix IV: Managed Area Trendlines	28
Appendix V: Managed Area Summary Box Plots	44

Libraries

Loads libraries used in the script. The inclusion of scipen option limits how frequently R defaults to scientific notation.

library(knitr)
library(data.table)
library(dplyr)
library(lubridate)
library(ggplot2)

```
library(ggpubr)
library(scales)
library(EnvStats)
library(tidyr)
options(scipen = 999)
```

File Import

 $Imports\ file\ that\ is\ determined\ in\ the\ WC_Discrete_parameter_ReportCompile.R\ script.$

The command fread is used because of its improved speed while handling large data files. Only columns that are used by the script are imported from the file, and are designated in the select input.

The script then gets the name of the parameter as it appears in the data file, units of the parameter, sets the SampleDate as a date object, and creates various scales of the date to be used by plotting functions.

Data Summary Before Filtering

This part is to create a data summary file that looks at the total amount of data, how many pass the initial database filtering, and how many are impacted by various ValueQualifiers. The ValueQualifiers of interest are H, I, Q, S, and U.

```
perc_Q = 100*N_Q/length(data$ValueQualifier),
N_S = length(grep("S", data$ValueQualifier)),
perc_S = 100*N_S/length(data$ValueQualifier),
N_U = length(grep("U", data$ValueQualifier)),
perc_U = 100*N_U/length(data$ValueQualifier))
```

'summarise()' has grouped output by 'AreaID'. You can override using the '.groups' argument.

Data Filtering and Data Impacted by Specific Value Qualifiers

Most data filtering is performed on export from the database, and is indicated by the Include variable. Include values of 1 indicate the data should be used for analysis, values of 0 indicate the data should not be used for analysis. Documentation on the database filtering is provided here: SEACAR Documentation-Analysis Filters and Calculations.docx

The filtering that is performed by the script at this point removes rows that are missing values for ResultValue, and only keeps data that is measured at the relative depth (surface, bottom, etc.) and activity type (field or sample) of interest. This is partly handled on export with the RelativeDepth variable, but there are some measurements that are considered both surface and bottom based on measurement depth and total depth. By default, these are marked as Surface for RelativeDepth and receive a SEACAR_QAQCFlag indicator of 12Q. Data passes the filtering the process if it is from the correct depth and has an Include value of 1. The script also only looks at data of the desired ActivityType which indicates whether it was measured in the field (Field) or in the lab (Sample).

After the initial filtering, a second filter variable is created to determine whether enough time is represented in the managed area, which is that each managed area has 10 year or more of unique year entries for observation that pass the initial filter. If data passes the first set of filtering criteria and the time criteria, they are used in the analysis.

After filtering, the amount of data impacted by the H (for dissolved oxygen & pH in program 476), I, Q, S (for Secchi depth), and U value qualifiers. A variable is also created that determines if scatter plot points should be a different color based on value qualifiers of interest.

```
if(depth=="Bottom"){
   data$RelativeDepth[grep("12Q", data$SEACAR_QAQCFlagCode[
          data$RelativeDepth == "Surface"])] <- "Bottom"
}

data$Include <- as.logical(data$Include)
data <- data[!is.na(data$ResultValue),]
if(param_name!="Secchi_Depth"){
   data <- data[!is.na(data$RelativeDepth),]
data <- data[data$RelativeDepth==depth,]
}

if(length(grep("Blank", data$ActivityType)) >0){
   data <- data[-grep("Blank", data$ActivityType),]
}</pre>
```

```
if(param_name == "Chlorophyll_a_uncorrected_for_pheophytin" |
   param_name == "Salinity" | param_name == "Turbidity"){
   data <- data[grep(activity, data$ActivityType[!is.na(data$ActivityType)]),]</pre>
}
if(param_name == "Water_Temperature"){
   data <- data[data$ResultValue>=-2,]
   data <- data[data$ResultValue>=0,]
data$Include[grep("H", data$ValueQualifier[data$ProgramID==476])] <- TRUE
MA_Years <- data[data$Include == TRUE, ] %>%
   group_by(ManagedAreaName) %>%
   summarize(N = length(unique(Year)))
MA_Years <- as.data.table(MA_Years[order(MA_Years$ManagedAreaName), ])
MA_Years$Enough_Time <- ifelse(MA_Years$N < 10, FALSE, TRUE)
data <- merge.data.frame(data, MA_Years[,c("ManagedAreaName", "Enough_Time")],
                          by = "ManagedAreaName")
data$Use_In_Analysis <- ifelse(data$Include == TRUE &</pre>
                                    data$Enough_Time == TRUE,
                                 TRUE, FALSE)
total <- length(data$Include)</pre>
pass_filter <- length(data$Include[data$Include==TRUE])</pre>
count_H <- length(grep("H", data$ValueQualifier[data$ProgramID==476]))</pre>
perc_H <- 100*count_H/length(data$ValueQualifier)</pre>
count_I <- length(grep("I", data$ValueQualifier))</pre>
perc_I <- 100*count_I/length(data$ValueQualifier)</pre>
count_Q <- length(grep("Q", data$ValueQualifier))</pre>
perc_Q <- 100*count_Q/length(data$ValueQualifier)</pre>
count_S <- length(grep("S", data$ValueQualifier))</pre>
perc_S <- 100*count_S/length(data$ValueQualifier)</pre>
count_U <- length(grep("U", data$ValueQualifier))</pre>
perc_U <- 100*count_U/length(data$ValueQualifier)</pre>
data$VQ Plot <- data$ValueQualifier</pre>
inc_H <- ifelse(param_name=="pH" | param_name=="Dissolved_Oxygen" |</pre>
                    param_name=="Dissolved_Oxygen_Saturation", TRUE, FALSE)
if (inc_H==TRUE){
   data$VQ_Plot <- gsub("[^HU]+", "", data$VQ_Plot)</pre>
   data$VQ_Plot <- gsub("UH", "HU", data$VQ_Plot)</pre>
   data$VQ_Plot[data$ProgramID!=476] <- gsub("[^U]+", "",</pre>
                                                data$VQ_Plot[data$ProgramID!=476])
```

```
data$VQ_Plot[data$VQ_Plot==""] <- NA</pre>
   cat(paste0("Number of Measurements: ", total,
              ", Number Passed Filter: ", pass_filter, "\n",
              "Program 476 H Codes: ", count_H, " (", round(perc_H, 6), "%)\n",
              "I Codes: ", count_I, " (", round(perc_I, 6), "%)\n",
              "Q Codes: ", count Q, " (", round(perc Q, 6), "%)\n",
              "U Codes: ", count_U, " (", round(perc_U, 6), "%)"))
} else if (param_name=="Secchi_Depth") {
   count_S <- length(grep("S", data$ValueQualifier))</pre>
  perc_S <- 100*count_S/length(data$ValueQualifier)</pre>
   data$VQ_Plot <- gsub("[^SU]+", "", data$VQ_Plot)</pre>
   data$VQ_Plot <- gsub("US", "SU", data$VQ_Plot)</pre>
   data$VQ_Plot[data$VQ_Plot==""] <- NA</pre>
   cat(paste0("Number of Measurements: ", total,
              ", Number Passed Filter: ", pass_filter, "\n",
              "I Codes: ", count_I, " (", round(perc_I, 6), "%)\n",
              "Q Codes: ", count_Q, " (", round(perc_Q, 6), "%)\n",
              "S Codes: ", count_S, " (", round(perc_S, 6), "%)\n",
              "U Codes: ", count_U, " (", round(perc_U, 6), "%)"))
} else{
   data$VQ_Plot <- gsub("[^U]+", "", data$VQ_Plot)</pre>
   data$VQ_Plot[data$VQ_Plot==""] <- NA</pre>
   cat(paste0("Number of Measurements: ", total,
              ", Number Passed Filter: ", pass_filter, "\n",
              "I Codes: ", count_I, " (", round(perc_I, 6), "%)\n",
              "Q Codes: ", count_Q, " (", round(perc_Q, 6), "%)\n",
              "U Codes: ", count_U, " (", round(perc_U, 6), "%)"))
}
## Number of Measurements: 39992, Number Passed Filter: 39885
## I Codes: 1192 (2.980596%)
## Q Codes: 91 (0.227546%)
## U Codes: 1132 (2.830566%)
data summ <- data %>%
   group_by(AreaID, ManagedAreaName) %>%
   summarize(ParameterName = parameter,
             RelativeDepth = depth,
             ActivityType = activity,
             N_Total = length(ResultValue),
             N_AnalysisUse = length(ResultValue[Use_In_Analysis==TRUE]),
             N_H = length(grep("H", data$ValueQualifier[data$ProgramID==476])),
             perc_H = 100*N_H/length(data$ValueQualifier),
             N_I = length(grep("I", data$ValueQualifier)),
             perc_I = 100*N_I/length(data$ValueQualifier),
             N_Q = length(grep("Q", data$ValueQualifier)),
             perc_Q = 100*N_Q/length(data$ValueQualifier),
             N_S = length(grep("S", data$ValueQualifier)),
             perc_S = 100*N_S/length(data$ValueQualifier),
             N U = length(grep("U", data$ValueQualifier)),
             perc U = 100*N U/length(data$ValueQualifier))
```

'summarise()' has grouped output by 'AreaID'. You can override using the '.groups' argument.

Managed Area Statistics

Gets summary statistics for each managed area. Excluded managed areas are not included into whether the data should be used or not. Uses piping from dplyr package to feed into subsequent steps. The following steps are performed:

- 1. Take the data variable and only include rows that have a Use_In_Analysis value of TRUE
- 2. Group data that have the same ManagedAreaName, Year, and Month.
 - Second summary statistics do not use the Month grouping and are only for ManagedAreaName and Year.
 - Third summary statistics do not use Year grouping and are only for ManagedAreaName and Month
- 3. For each group, provide the following information: Number of Entries (N), Lowest Value (Min), Largest Value (Max), Median, Mean, Standard Deviation, and a list of all Program IDs included in these measurements.
- 4. Sort the data in ascending (A to Z and 0 to 9) order based on ManagedAreaName then Year then Month
- 5. Write summary stats to a pipe-delimited .txt file in the output directory
 - Click this text to open Git directory with output files

```
MA_names <- unique(data$ManagedAreaName[data$Use_In_Analysis == TRUE])
MA_names <- MA_names[order(MA_names)]</pre>
n <- length(MA_names)</pre>
MA_YM_Stats <- data[data$Use_In_Analysis == TRUE, ] %>%
  group_by(AreaID, ManagedAreaName, Year, Month) %>%
   summarize(ParameterName = parameter,
             RelativeDepth = depth,
             ActivityType = activity,
             N = length(ResultValue),
             Min = min(ResultValue),
             Max = max(ResultValue),
             Median = median(ResultValue),
             Mean = mean(ResultValue),
             StandardDeviation = sd(ResultValue),
             ProgramIDs = paste(sort(unique(ProgramID), decreasing = FALSE),
                                 collapse = ', '))
MA_YM_Stats <- as.data.table(MA_YM_Stats[order(MA_YM_Stats$ManagedAreaName,
                                                MA_YM_Stats$Year,
                                                MA_YM_Stats$Month), ])
fwrite(MA_YM_Stats, paste0(out_dir,"/", param_name, "_", activity, "_", depth,
                            "_ManagedArea_YearMonth_Stats.txt"), sep = "|")
MA_Y_Stats <- data[data$Use_In_Analysis == TRUE, ] %>%
  group by (AreaID, ManagedAreaName, Year) %>%
  summarize(ParameterName = parameter,
```

```
RelativeDepth = depth,
             ActivityType = activity,
             N = length(ResultValue),
             Min = min(ResultValue),
             Max = max(ResultValue),
             Median = median(ResultValue),
             Mean = mean(ResultValue),
             StandardDeviation = sd(ResultValue),
             ProgramIDs = paste(sort(unique(ProgramID), decreasing = FALSE),
                                collapse = ', '))
MA_Y_Stats <- as.data.table(MA_Y_Stats[order(MA_Y_Stats$ManagedAreaName,
                                             MA_Y_Stats$Year), ])
fwrite(MA_Y_Stats, paste0(out_dir,"/", param_name, "_", activity, "_", depth,
                          "_ManagedArea_Year_Stats.txt"), sep = "|")
MA_M_Stats <- data[data$Use_In_Analysis == TRUE, ] %>%
   group_by(AreaID, ManagedAreaName, Month) %>%
   summarize(ParameterName = parameter,
             RelativeDepth = depth,
             ActivityType = activity,
             N = length(ResultValue),
             Min = min(ResultValue),
             Max = max(ResultValue),
             Median = median(ResultValue),
             Mean = mean(ResultValue),
             StandardDeviation = sd(ResultValue),
             ProgramIDs = paste(sort(unique(ProgramID), decreasing = FALSE),
                                collapse = ', '))
MA_M_Stats <- as.data.table(MA_M_Stats[order(MA_M_Stats$ManagedAreaName,
                                             MA_M_Stats$Month), ])
fwrite(MA_M_Stats, paste0(out_dir,"/", param_name, "_", activity, "_", depth,
                          "_ManagedArea_Month_Stats.txt"), sep = "|")
```

Monitoring Location Statistics

Gets monitoring location statistics, which is defined as a unique combination of ManagedAreaName, ProgramID, ProgramAreaName, and ProgramLocationID, using piping from dplyr package. The following steps are performed:

- 1. Take the data variable and only include rows that have a Use_In_Analysis value of TRUE
- 2. Group data that have the same ManagedAreaName, ProgramID, ProgramName, and ProgramLocationID.
- 3. For each group, provide the following information: Earliest Sample Date (EarliestSampleDate), Latest Sample Date (LastSampleDate), Number of Entries (N), Lowest Value (Min), Largest Value (Max), Median, Mean, and Standard Deviation.
- 4. Sort the data in ascending (A to Z and 0 to 9) order based on ManagedAreaName then ProgramLocationID then ProgramLocationID
- 5. Write summary stats to a pipe-delimited .txt file in the output directory
 - Click this text to open Git directory with output files

```
Mon_Stats <- data[data$Use_In_Analysis == TRUE, ] %>%
   group by(AreaID, ManagedAreaName, ProgramID, ProgramName, ProgramLocationID) %>%
   summarize(ParameterName = parameter,
             RelativeDepth = depth,
             ActivityType = activity,
             EarliestSampleDate = min(SampleDate),
             LastSampleDate = max(SampleDate),
             N = length(ResultValue),
             Min = min(ResultValue),
             Max = max(ResultValue),
             Median = median(ResultValue),
             Mean = mean(ResultValue),
             StandardDeviation = sd(ResultValue))
Mon_Stats <- as.data.table(Mon_Stats[order(Mon_Stats$ManagedAreaName,</pre>
                                               Mon_Stats$ProgramName,
                                               Mon_Stats$ProgramID,
                                               Mon_Stats$ProgramLocationID), ])
fwrite(Mon_Stats, paste0(out_dir,"/", param_name, "_", activity, "_", depth,
                          "_MonitoringLoc_Stats.txt"), sep = "|")
```

Seasonal Kendall Tau Analysis

Gets seasonal Kendall Tau statistics using the kendallSeasonalTrendTest from the EnvStats package. The Trend parameter is determined from a user-defined function based on the median, Senn slope, and p values from the data. Analysis modified from code created by Jason Scolaro that performed at The Water Atlas: https://sarasota.wateratlas.usf.edu/water-quality-trends/#analysis-overview

The following steps are performed:

- 1. Define the functions used in the analysis
- 2. Check to see if there are any groups to run analysis on.
- 3. Take the data variable and only include rows that have a Use_In_Analysis value of TRUE
- 4. Group data that have the same ManagedAreaName.
- 5. For each group, provides the following information: Earliest Sample Date (EarliestSampleDate), Latest Sample Date (LastSampleDate), Number of Entries (N), Lowest Value (Min), Largest Value (Max), Median, Mean, Standard Deviation, tau, Senn Slope (SennSlope), Senn Intercept (SennIntercept), and p.
 - The analysis is run with the kendallSeasonalTrendTest function using the Year values for year, and Month as the seasonal qualifier, and Trend.
 - An independent obs value of TRUE indicates that the data should be treated as not being serially auto-correlated. An independent obs value of FALSE indicates that it is treated as being serially auto-correlated, but also requires one observation per season per year for the full time of observation.
- 6. Reformat columns in the data frame from export.
- 7. Write summary stats to a pipe-delimited .txt file in the output directory
 - Click this text to open Git directory with output files

```
tau <- NULL
   tryCatch({
      ken <-
         kendallSeasonalTrendTest(
             y = data$ResultValue,
             season = data$Month,
             year = data$Year,
             independent.obs = independent
         )
      tau <- ken$estimate[1]</pre>
      p <- ken$p.value[2]</pre>
      slope <- ken$estimate[2]</pre>
      intercept <- ken$estimate[3]</pre>
      trend <- trend_calculator(slope, stats.median, p)</pre>
   }, warning = function(w) {
      print(w)
   }, error = function(e) {
      print(e)
   }, finally = {
      if (!exists("tau")) {
          tau <- NULL
      if (!exists("p")) {
         p <- NULL
      if (!exists("slope")) {
          slope <- NULL</pre>
      if (!exists("intercept")) {
          intercept <- NULL</pre>
      if (!exists("trend")) {
          trend <- NULL
   })
   KT <-c(unique(data$AreaID),</pre>
           unique(data$ManagedAreaName),
           parameter,
           depth,
           activity,
           independent,
           stats.median,
           nrow(data),
           stats.minYear,
           stats.maxYear,
           tau,
           p,
           slope,
           intercept,
           trend)
   return(KT)
runStats <- function(data) {</pre>
```

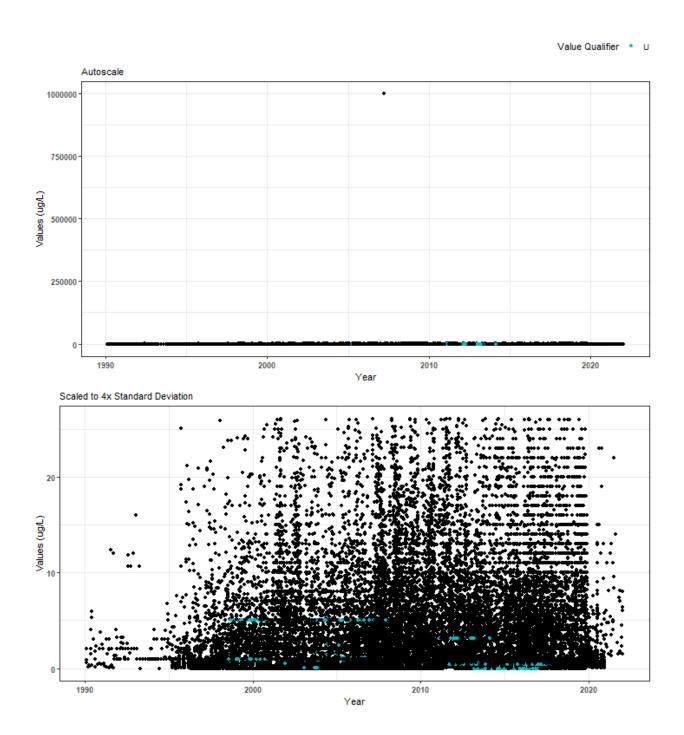
```
data$Index <- as.Date(data$SampleDate) # , "%Y-%m-%d")</pre>
   data$ResultValue <- as.numeric(data$ResultValue)</pre>
   # Calculate basic stats
   stats.median <- median(data$ResultValue, na.rm = TRUE)</pre>
   stats.minYear <- min(data$Year, na.rm = TRUE)</pre>
   stats.maxYear <- max(data$Year, na.rm = TRUE)</pre>
   # Calculate Kendall Tau and Slope stats, then update appropriate columns and table
   KT <- tauSeasonal(data, TRUE, stats.median,</pre>
                      stats.minYear, stats.maxYear)
   if (is.null(KT[11])) {
      KT <- tauSeasonal(data, FALSE, stats.median,</pre>
                          stats.minYear, stats.maxYear)
   if (is.null(KT.Stats) == TRUE) {
      KT.Stats <- KT</pre>
   } else{
      KT.Stats <- rbind(KT.Stats, KT)</pre>
   return(KT.Stats)
}
trend_calculator <- function(slope, median_value, p) {</pre>
   trend <-
      if (p < .05 \& abs(slope) > abs(median_value) / 10.) {
         if (slope > 0) {
             2
         }
         else {
             -2
   else if (p < .05 & abs(slope) < abs(median_value) / 10.) {</pre>
      if (slope > 0) {
         1
      }
      else {
   }
   else
   return(trend)
}
KT.Stats <- NULL
# Loop that goes through each managed area.
# List of managed areas stored in MA_Years$ManagedAreaName
c_names <- c("AreaID", "ManagedAreaName", "ParameterName", "RelativeDepth",</pre>
              "ActivityType", "Independent", "Median", "N", "EarliestYear",
              "LatestYear", "tau", "p", "SennSlope", "SennIntercept", "Trend")
if(n==0)
   KT.Stats <- data.frame(matrix(ncol=15, nrow=0))</pre>
   colnames(KT.Stats) <- c_names</pre>
   fwrite(KT.Stats, paste0(out_dir,"/", param_name, "_", activity, "_", depth,
                             "_KendallTau_Stats.txt"), sep = "|")
```

```
} else{
   for (i in 1:n) {
      values <- data[data$Use_In_Analysis == TRUE &</pre>
                          data$ManagedAreaName == MA_names[i], ]
      if (nrow(values) > 0) {
         KT.Stats <- runStats(values)</pre>
   }
   KT.Stats <- as.data.frame(KT.Stats)</pre>
   c_names <- c("AreaID", "ManagedAreaName", "ParameterName", "RelativeDepth",</pre>
                 "ActivityType", "Independent", "Median", "N", "EarliestYear",
                 "LatestYear", "tau", "p", "SennSlope", "SennIntercept", "Trend")
   if(dim(KT.Stats)[2]==1){
      KT.Stats <- as.data.frame(t(KT.Stats))</pre>
   colnames(KT.Stats) <- c_names</pre>
   rownames(KT.Stats) <- seq(1:nrow(KT.Stats))</pre>
   KT.Stats$Median <- as.numeric(KT.Stats$Median)</pre>
   KT.Stats$N <- as.integer(KT.Stats$N)</pre>
   KT.Stats$EarliestYear <- as.integer(KT.Stats$EarliestYear)</pre>
   KT.Stats$LatestYear <- as.integer(KT.Stats$LatestYear)</pre>
   KT.Stats$tau <- round(as.numeric(KT.Stats$tau), digits=4)</pre>
   KT.Stats$p <- round(as.numeric(KT.Stats$p), digits=4)</pre>
   KT.Stats$SennSlope <- as.numeric(KT.Stats$SennSlope)</pre>
   KT.Stats$SennIntercept <- as.numeric(KT.Stats$SennIntercept)</pre>
   KT.Stats$Trend <- as.integer(KT.Stats$Trend)</pre>
   fwrite(KT.Stats, paste0(out_dir,"/", param_name, "_", activity, "_", depth,
                              "_KendallTau_Stats.txt"), sep = "|")
}
```

Appendix I: Scatter Plot of Entire Dataset

This part will create a scatter plot of the all data that passed initial filtering criteria with points colored based on specific value qualifiers. The values determined at the beginning (year_lower, year_upper, min_RV, mn_RV, x_scale, and y_scale) are solely for use by the plotting functions and are not output as part of the computed statistics.

```
labs(subtitle = "Autoscale",
        x = "Year", y = paste0("Values (", unit, ")"),
        color="Value Qualifier") +
   theme bw() +
   theme(legend.position = "top", legend.box = "horizontal",
         legend.justification = "right",
         axis.text.x = element_text(face = "bold"),
         axis.text.y = element text(face="bold")) +
   scale_x_date(labels = date_format("%Y")) +
   {if(inc H==TRUE){
      scale_color_manual(values = c("H"= "#F8766D", "U"= "#00BFC4",
                                     "HU" = "#7CAE00"), na.value="black")
   } else {
      scale_color_manual(values = c("U"= "#00BFC4"), na.value="black")
   }}
p2 <- ggplot(data = data[data$Include==TRUE,],</pre>
             aes(x = SampleDate, y = ResultValue,
                 color=VQ_Plot)) +
   geom_point(size = 1.5) +
   ylim(min_RV, y_scale) +
   labs(subtitle = "Scaled to 4x Standard Deviation",
        x = "Year", y = paste0("Values (", unit, ")")) +
   theme bw() +
   theme(legend.position = "none",
         axis.text.x = element text(face = "bold"),
         axis.text.y = element_text(face = "bold")) +
   scale_x_date(labels = date_format("%Y")) +
   {if(inc_H==TRUE){
      scale_color_manual(values = c("H"= "#F8766D", "U"= "#00BFC4",
                                     "HU" = "#7CAE00"), na.value="black")
   } else {
      scale_color_manual(values = c("U"= "#00BFC4"), na.value="black")
   }}
leg <- get_legend(p1)</pre>
pset <- ggarrange(leg, p1 + theme(legend.position = "none"), p2,</pre>
                  ncol = 1, heights = c(0.1, 1, 1)
p0 <- ggplot() + labs(title = "Scatter Plot for Entire Dataset") +</pre>
   theme_bw() + theme(plot.title = element_text(face="bold", hjust = 0.5),
                      panel.border = element blank(),
                      panel.grid.major = element_blank(),
                      panel.grid.minor = element_blank(),
                      axis.line = element blank())
ggarrange(p0, pset, ncol = 1, heights = c(0.1, 1))
```



Appendix II: Dataset Summary Box Plots

Box plots are created by using the entire data set and excludes any data that has been previously filtered out. The scripts that create plots follow this format

- 1. Use the data set that only has Use_In_Analysis of TRUE
- 2. Set what values are to be used for the x-axis, y-axis, and the variable that should determine groups for the box plots
- 3. Set the plot type as a box plot with the size of the outlier points
- 4. Create the title, x-axis, y-axis, and color fill labels
- 5. Set the y and x limits
- 6. Make the axis labels bold
- 7. Plot the arrangement as a set of panels

This set of box plots are grouped by year.

```
min_RV <- min(data$ResultValue[data$Include == TRUE])</pre>
mn RV <- mean(data$ResultValue[data$Include == TRUE &</pre>
                                   data$ResultValue <</pre>
                                   quantile(data$ResultValue, 0.98)])
sd_RV <- sd(data$ResultValue[data$Include == TRUE &</pre>
                                 data$ResultValue <</pre>
                                 quantile(data$ResultValue, 0.98)])
y scale \leftarrow mn RV + 4 * sd RV
p1 <- ggplot(data = data[data$Include == TRUE, ],</pre>
             aes(x = Year, y = ResultValue, group = Year)) +
   geom_boxplot(outlier.size = 0.5) +
   labs(subtitle = "Autoscale", x = "Year",
        y = paste0("Values (", unit, ")")) +
   theme bw() +
   theme(axis.text.x = element_text(face = "bold"),
         axis.text.y = element_text(face = "bold"))
p2 <- ggplot(data = data[data$Include == TRUE, ],</pre>
             aes(x = Year, y = ResultValue, group = Year)) +
   geom_boxplot(outlier.size = 0.5) +
   labs(subtitle = "Scaled to 4x Standard Deviation", x = "Year",
        y = paste0("Values (", unit, ")")) +
   ylim(min_RV, y_scale) +
   theme bw() +
   theme(axis.text.x = element text(face = "bold"),
         axis.text.y = element text(face = "bold"))
p3 <- ggplot(data = data[data$Include == TRUE, ],
             aes(x = as.integer(Year), y = ResultValue, group = Year)) +
   geom boxplot(outlier.size = 0.5) +
   labs(subtitle = "Scaled to 4x Standard Deviation, Last 10 Years",
        x = "Year", y = paste0("Values (", unit, ")")) +
   ylim(min_RV, y_scale) +
   scale_x_continuous(limits = c(max(data$Year) - 10.5, max(data$Year)+1),
                      breaks = seq(max(data$Year) - 10, max(data$Year), 2)) +
   theme_bw() +
   theme(axis.text.x = element_text(face = "bold"),
         axis.text.y = element_text(face = "bold"))
set <- ggarrange(p1, p2, p3, ncol = 1)</pre>
p0 <- ggplot() + labs(title = "Summary Box Plots for Entire Data",
```

```
subtitle = "By Year") + theme_bw() +
theme(plot.title = element_text(face="bold", hjust = 0.5),
    panel.border = element_blank(), panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(), axis.line = element_blank())
Yset <- ggarrange(p0, set, ncol=1, heights = c(0.07, 1))
```

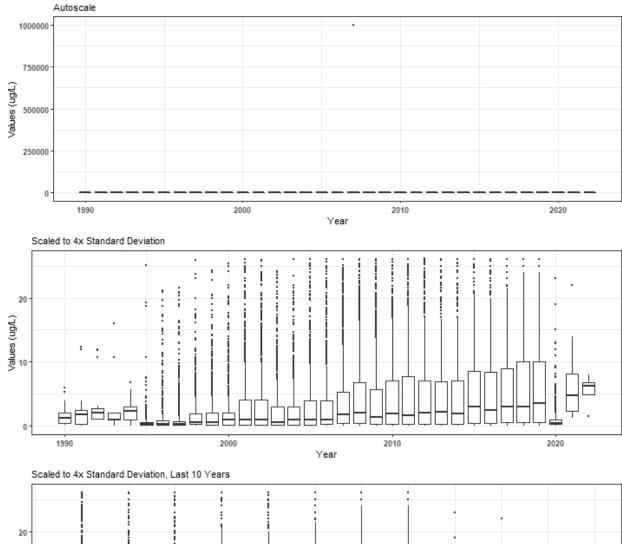
This set of box plots are grouped by year and month with the color being related to the month.

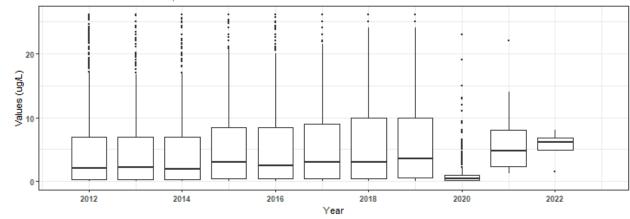
```
p1 <- ggplot(data = data[data$Include == TRUE, ],</pre>
             aes(x = YearMonthDec, y = ResultValue,
                 group = YearMonth, color = as.factor(Month))) +
   geom_boxplot(outlier.size = 0.5) +
   labs(subtitle = "Autoscale", x = "Year",
        y = paste0("Values (", unit, ")"), color="Month") +
   theme bw() +
   theme(legend.position = "top", legend.box = "horizontal",
         axis.text.x = element text(face = "bold"),
         axis.text.y = element_text(face = "bold")) +
   guides(color = guide_legend(nrow = 1))
p2 <- ggplot(data = data[data$Include == TRUE, ],</pre>
             aes(x = YearMonthDec, y = ResultValue,
                 group = YearMonth, color = as.factor(Month))) +
   geom_boxplot(outlier.size = 0.5) +
   labs(subtitle = "Scaled to 4x Standard Deviation",
        x = "Year", y = paste0("Values (", unit, ")")) +
  ylim(min_RV, y_scale) +
   theme bw() +
   theme(legend.position = "none", axis.text.x = element_text(face = "bold"),
         axis.text.y = element_text(face = "bold"))
p3 <- ggplot(data = data[data$Include == TRUE, ],
             aes(x = YearMonthDec, y = ResultValue,
                 group = YearMonth, color = as.factor(Month))) +
   geom boxplot(outlier.size = 0.5) +
   labs(subtitle = "Scaled to 4x Standard Deviation, Last 10 Years",
        x = "Year", y = paste0("Values (", unit, ")")) +
   ylim(min_RV, y_scale) +
   scale_x_continuous(limits = c(max(data$Year) - 10.5, max(data$Year)+1),
                      breaks = seq(max(data$Year) - 10, max(data$Year), 2)) +
   theme_bw() +
   theme(legend.position = "none", axis.text.x = element_text(face = "bold"),
         axis.text.y = element_text(face = "bold"))
leg <- get_legend(p1)</pre>
set <- ggarrange(leg, p1 + theme(legend.position = "none"), p2, p3, ncol = 1,</pre>
                 heights = c(0.1, 1, 1, 1)
p0 <- ggplot() + labs(title = "Summary Box Plots for Entire Data",
                      subtitle = "By Year & Month") + theme_bw() +
   theme(plot.title = element_text(face="bold", hjust = 0.5),
         panel.border = element blank(), panel.grid.major = element blank(),
         panel.grid.minor = element_blank(), axis.line = element_blank())
```

```
YMset <- ggarrange(p0, set, ncol=1, heights = c(0.07, 1))
```

The following box plots are grouped by month with fill color being related to the month. This is designed to view potential seasonal trends.

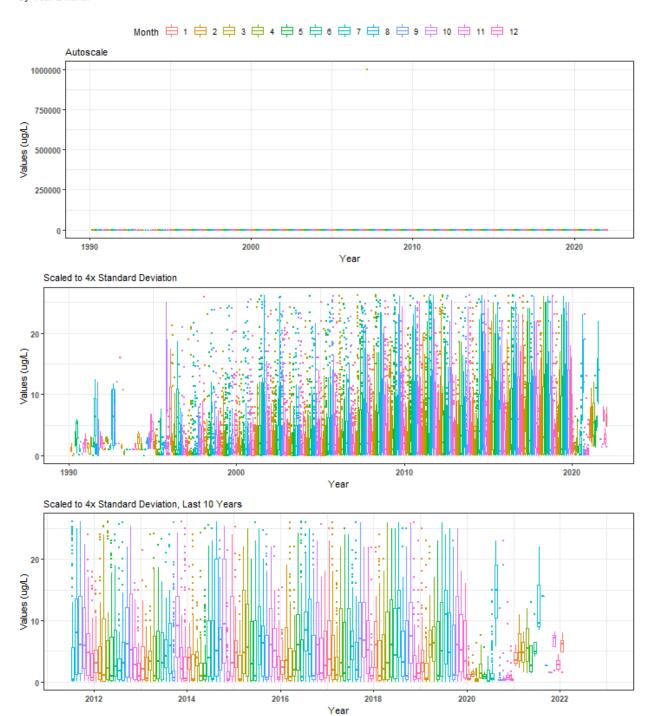
```
p1 <- ggplot(data = data[data$Include == TRUE, ],
             aes(x = Month, y = ResultValue,
                 group = Month, fill = as.factor(Month))) +
   geom_boxplot(outlier.size = 0.5) +
   labs(subtitle = "Autoscale", x = "Month",
        y = paste0("Values (", unit, ")"), fill="Month") +
   scale_x_continuous(limits = c(0, 13), breaks = seq(3, 12, 3)) +
   theme_bw() +
   theme(legend.position = "top", legend.box = "horizontal",
         axis.text.x = element_text(face = "bold"),
         axis.text.y = element_text(face = "bold")) +
   guides(fill = guide_legend(nrow = 1))
p2 <- ggplot(data = data[data$Include == TRUE, ],</pre>
             aes(x = Month, y = ResultValue,
                 group = Month, fill = as.factor(Month))) +
   geom_boxplot(outlier.size = 0.5) +
   labs(subtitle = "Scaled to 4x Standard Deviation",
        x = "Month", y = paste0("Values (", unit, ")")) +
   ylim(min_RV, y_scale) +
   scale_x_continuous(limits = c(0, 13), breaks = seq(3, 12, 3)) +
   theme_bw() +
   theme(legend.position = "none", axis.text.x = element_text(face = "bold"),
         axis.text.y = element_text(face = "bold"))
p3 <- ggplot(data = data[data$Include == TRUE &
                            data$Year >= max(data$Year) - 10, ],
             aes(x = Month, y = ResultValue,
                 group = Month, fill = as.factor(Month))) +
   geom boxplot(outlier.size = 0.5) +
   labs(subtitle = "Scaled to 4x Standard Deviation, Last 10 Years",
        x = "Month", y = paste0("Values (", unit, ")")) +
   ylim(min_RV, y_scale) +
   scale_x_continuous(limits = c(0, 13), breaks = seq(3, 12, 3)) +
   theme bw() +
   theme(legend.position = "none", axis.text.x = element_text(face = "bold"),
         axis.text.y = element_text(face = "bold"))
leg <- get_legend(p1)</pre>
set <- ggarrange(leg, p1 + theme(legend.position = "none"), p2, p3, ncol = 1,</pre>
                 heights = c(0.1, 1, 1, 1)
p0 <- ggplot() + labs(title = "Summary Box Plots for Entire Data",
                      subtitle = "By Month") + theme_bw() +
   theme(plot.title = element_text(face="bold", hjust = 0.5),
         panel.border = element_blank(), panel.grid.major = element_blank(),
         panel.grid.minor = element_blank(), axis.line = element_blank())
Mset <- ggarrange(p0, set, ncol=1, heights = c(0.07, 1))</pre>
```





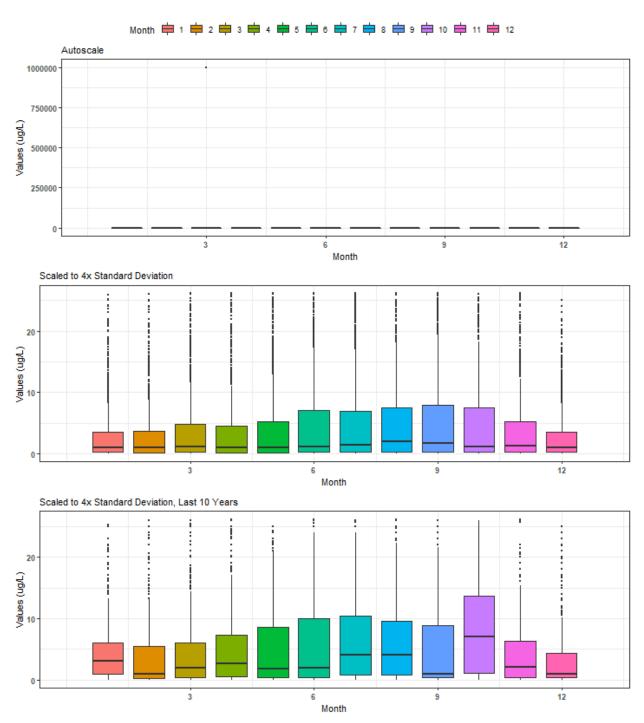
Summary Box Plots for Entire Data

By Year & Month







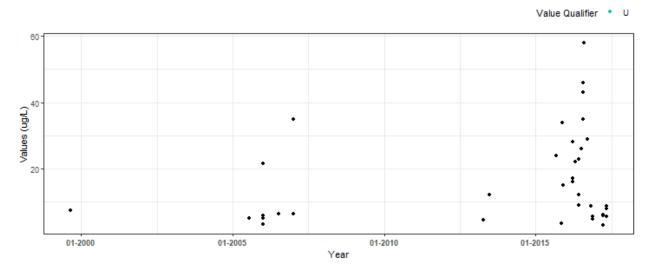


Appendix III: Excluded Managed Areas

Scatter plots of data values are created for managed areas that have fewer than 10 separate years of data entries. Data points are colored based on specific value qualifiers of interest.

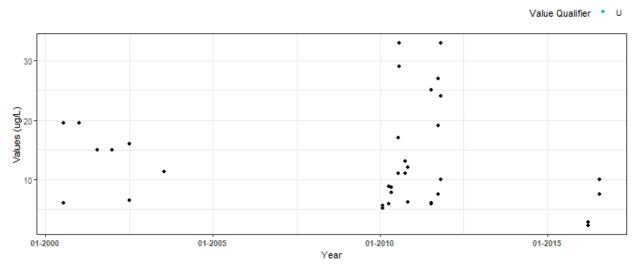
```
MA_Exclude <- MA_Years[MA_Years$Enough_Time==FALSE,]</pre>
MA_Exclude <- MA_Exclude[order(MA_Exclude$ManagedAreaName),]</pre>
z=length(MA_Exclude$ManagedAreaName)
if(z==0){
   print("There are no managed areas that qualify.")
} else {
   for(i in 1:z){
      p1<-ggplot(data=data[data$ManagedAreaName==MA_Exclude$ManagedAreaName[i]&
                              data$Include == TRUE, ],
                 aes(x = SampleDate, y = ResultValue, color=VQ_Plot)) +
         geom_point() +
         labs(title = paste0("Scatter Plot of Excluded Managed Area\n",
                             MA_Exclude$ManagedAreaName[i], " (",
                             MA_Exclude$N[i], " Unique Years)"),
              subtitle="Autoscale", x = "Year",
              y = paste0("Values (", unit, ")"), color="Value Qualifier") +
         theme_bw() +
         theme(plot.title = element_text(face="bold", hjust = 0.5),
               legend.position = "top", legend.box = "horizontal",
               legend.justification = "right",
               axis.text.x = element_text(face = "bold")) +
         scale_x_date(labels = date_format("%m-%Y")) +
         {if(inc H==TRUE){
            scale_color_manual(values = c("H"= "#F8766D", "U"= "#00BFC4",
                                           "HU" = "#7CAE00"), na.value="black")
         } else {
            scale_color_manual(values = c("U"= "#00BFC4"), na.value="black")
         }}
      print(p1)
   }
}
```

Scatter Plot of Excluded Managed Area Banana River Aquatic Preserve (7 Unique Years)

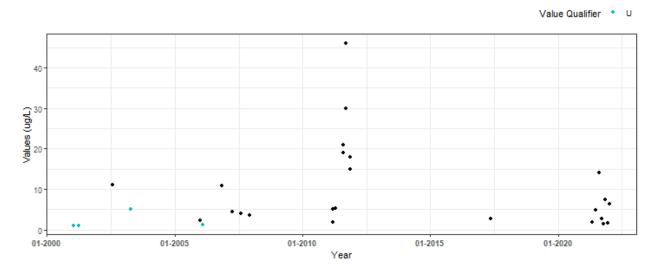


Scatter Plot of Excluded Managed Area Boca Ciega Bay Aquatic Preserve (7 Unique Years)

Autoscale

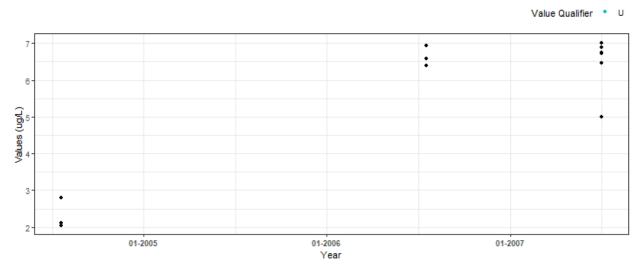


Scatter Plot of Excluded Managed Area Cape Haze Aquatic Preserve (9 Unique Years)

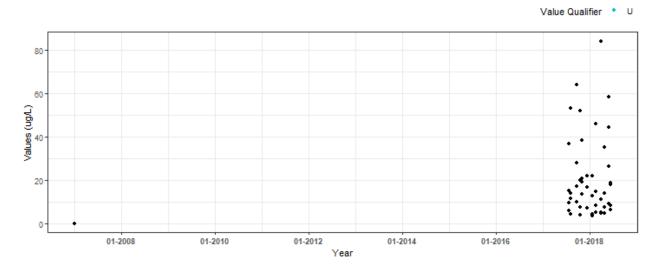


Scatter Plot of Excluded Managed Area Fort Clinch State Park Aquatic Preserve (3 Unique Years)

Autoscale

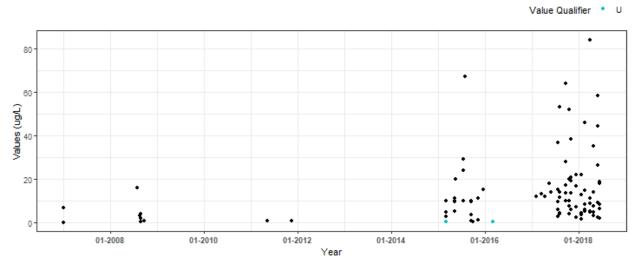


Scatter Plot of Excluded Managed Area Guana River Marsh Aquatic Preserve (3 Unique Years)

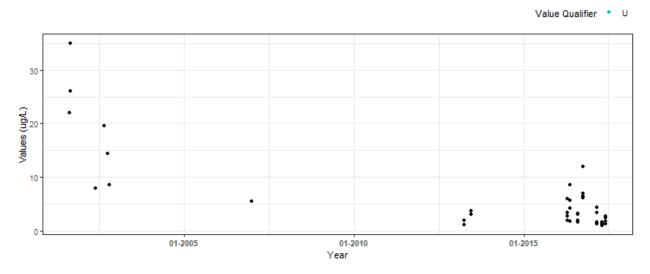


Scatter Plot of Excluded Managed Area Guana Tolomato Matanzas National Estuarine Research Reserve (7 Unique Years)

Autoscale

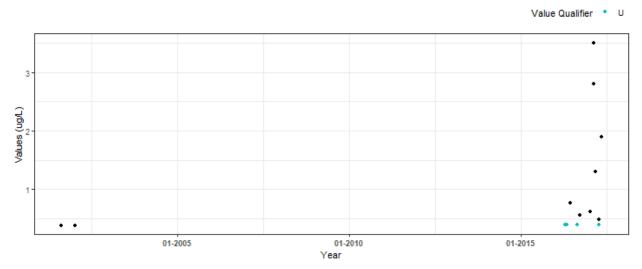


Scatter Plot of Excluded Managed Area Indian River-Vero Beach to Ft. Pierce Aquatic Preserve (6 Unique Years)

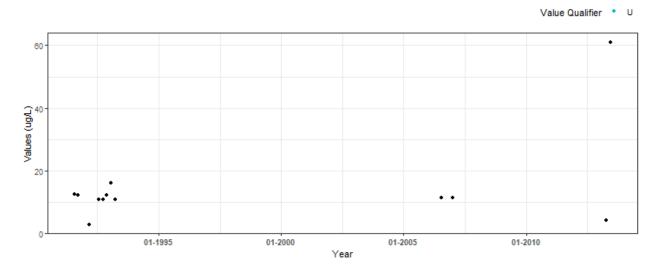


Scatter Plot of Excluded Managed Area Lignumvitae Key Aquatic Preserve (3 Unique Years)

Autoscale

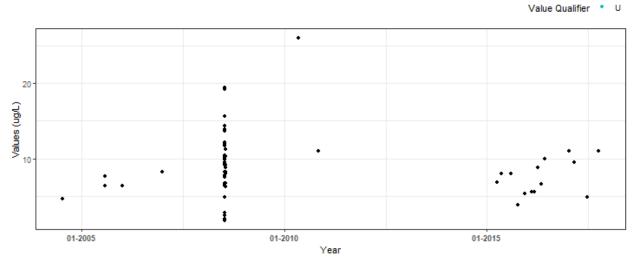


Scatter Plot of Excluded Managed Area Mosquito Lagoon Aquatic Preserve (5 Unique Years)

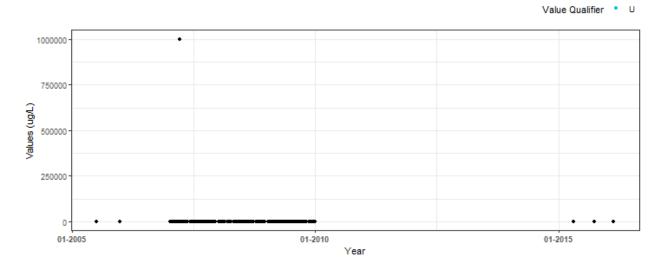


Scatter Plot of Excluded Managed Area Nassau River-St. Johns River Marshes Aquatic Preserve (8 Unique Years)

Autoscale

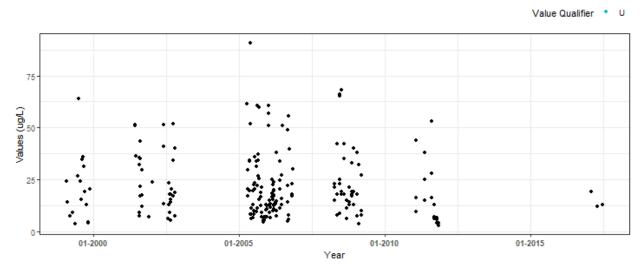


Scatter Plot of Excluded Managed Area Nature Coast Aquatic Preserve (6 Unique Years)

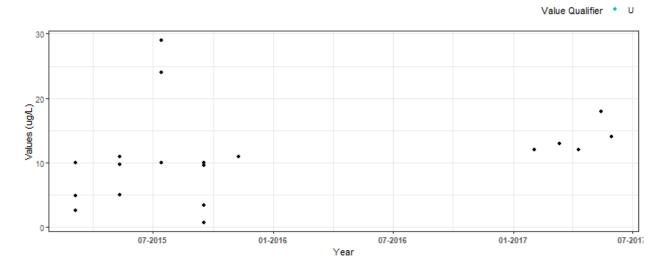


Scatter Plot of Excluded Managed Area North Fork St. Lucie Aquatic Preserve (9 Unique Years)

Autoscale



Scatter Plot of Excluded Managed Area Pellicer Creek Aquatic Preserve (2 Unique Years)



Scatter Plot of Excluded Managed Area Southeast Florida Coral Reef Ecosystem Conservation Area (4 Unique Years)

Autoscale

Value Qualifier • U

2.0

1.5

0.5

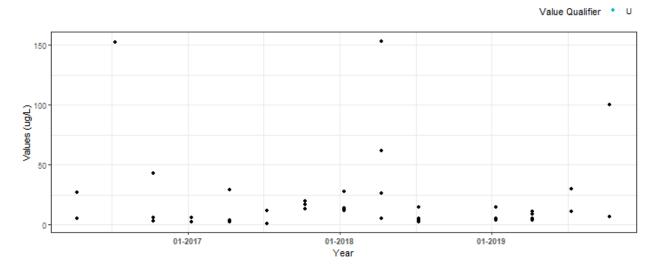
0.1-2010

01-2011

01-2012

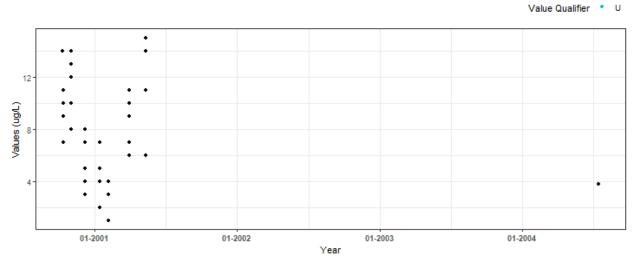
Year

Scatter Plot of Excluded Managed Area St. Joseph Bay State Buffer Preserve (4 Unique Years)



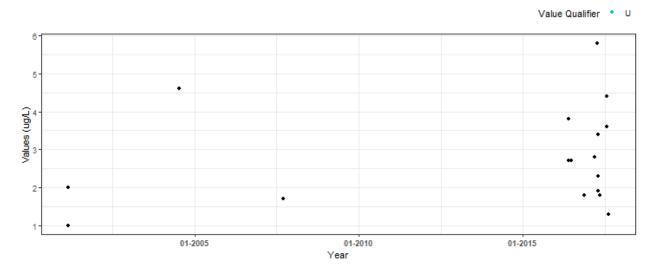
Scatter Plot of Excluded Managed Area Tomoka Marsh Aquatic Preserve (3 Unique Years)





Scatter Plot of Excluded Managed Area Yellow River Marsh Aquatic Preserve (5 Unique Years)





Appendix IV: Managed Area Trendlines

The plots created in this section are designed to show the general trend of the data. Data is taken and grouped by ManagedAreaName. The trendlines on the plots are created using the Senn slope and intercept from the seasonal Kendall Tau analysis. The scripts that create plots follow this format

- 1. Use the data set that only has Use_In_Analysis of TRUE for the desired managed area
- 2. Determine the earliest and latest year of the data to create x-axis scale and intervals
- 3. Determine the minimum, mean, and standard deviation for the data to be used for y-axis scales
 - \bullet Excludes the top 2% of values to reduce the impact of extreme outliers on the y-axis scale
- 4. Set what values are to be used for the x-axis, y-axis, and the variable that should determine groups for the plots

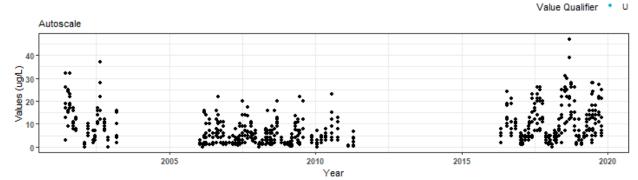
- 5. Set the plot type as a point plot with the size of the points
- 6. Add the linear trend
- 7. Create the title, x-axis, y-axis, and color fill labels
- 8. Set the y and x limits
- 9. Make the axis labels bold
- 10. Plot the arrangement as a set of panels

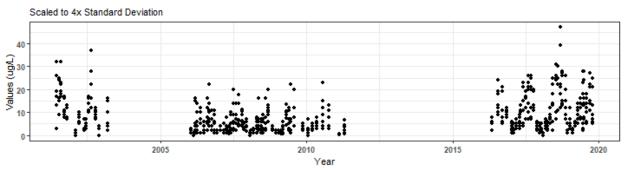
```
if(n==0){
   print("There are no managed areas that qualify.")
} else {
   for (i in 1:n) {
      plot_data <- data[data$Use_In_Analysis == TRUE &</pre>
                                       data$ManagedAreaName == MA_names[i],]
      year_lower <- min(plot_data$Year)</pre>
      year_upper <- max(plot_data$Year)</pre>
      min_RV <- min(plot_data$ResultValue)</pre>
      mn_RV <- mean(plot_data$ResultValue[data$ResultValue <</pre>
                                          quantile(data$ResultValue, 0.98)])
      sd RV <- sd(plot data$ResultValue[data$ResultValue <</pre>
                                        quantile(data$ResultValue, 0.98)])
      x_scale <- ifelse(year_upper - year_lower > 30, 10, 5)
      y scale \leftarrow mn RV + 4 * sd RV
      tau <- KT.Stats$tau[KT.Stats$ManagedAreaName==MA_names[i]]</pre>
      s_slope <- KT.Stats$SennSlope[KT.Stats$ManagedAreaName==MA_names[i]]</pre>
      s_int <- KT.Stats$SennIntercept[KT.Stats$ManagedAreaName==MA_names[i]]</pre>
      trend <- KT.Stats$Trend[KT.Stats$ManagedAreaName==MA_names[i]]</pre>
      p <- KT.Stats$p[KT.Stats$ManagedAreaName==MA_names[i]]</pre>
      model <- lm(ResultValue ~ DecDate,</pre>
                   data = plot_data)
      m_int <- coef(model)[[1]]</pre>
      m_slope <- coef(model)[[2]]</pre>
      p1 <- ggplot(data = plot_data,
                    aes(x = DecDate, y = ResultValue,
                        color=VQ Plot)) +
         geom_point(size = 1.5) +
         geom_abline(aes(slope=s_slope, intercept=s_int),
                      color="red", size=1.5) +
         labs(subtitle = "Autoscale",
               x = "Year", y = paste0("Values (", unit, ")"),
               color="Value Qualifier") +
         theme_bw() +
         theme(legend.position = "top", legend.box = "horizontal",
                legend.justification = "right",
                axis.text.x = element_text(face = "bold"),
               axis.text.y = element_text(face="bold")) +
         {if(inc_H==TRUE){
            scale_color_manual(values = c("H"= "#F8766D", "U"= "#00BFC4",
                                             "HU" = "#7CAE00"), na.value="black")
         } else {
            scale_color_manual(values = c("U"= "#00BFC4"), na.value="black")
         }}
```

```
p2 <- ggplot(data = plot_data,</pre>
                   aes(x = DecDate, y = ResultValue,
                       color=VQ_Plot)) +
         geom_point(size = 1.5) +
         geom_abline(aes(slope=s_slope, intercept=s_int),
                     color="red", size=1.5) +
         ylim(min_RV, y_scale) +
         labs(subtitle = "Scaled to 4x Standard Deviation",
              x = "Year", y = paste0("Values (", unit, ")")) +
         theme bw() +
         theme(legend.position = "none",
               axis.text.x = element_text(face = "bold"),
               axis.text.y = element_text(face="bold")) +
         {if(inc_H==TRUE){
            scale_color_manual(values = c("H"= "#F8766D", "U"= "#00BFC4",
                                           "HU" = "#7CAE00"), na.value="black")
            scale_color_manual(values = c("U"= "#00BFC4"), na.value="black")
         }}
      leg <- get_legend(p1)</pre>
      KTset <- ggarrange(leg, p1 + theme(legend.position = "none"), p2,</pre>
                         ncol = 1, heights = c(0.1, 1, 1))
      p0 <- ggplot() + labs(title = paste0("Data Points with Trendlines for ",
                                            MA_names[i]),
                            subtitle =paste0("Senn Slope = ", s_slope,
                                              ", Senn Intercept = ", s_int,
                                              "\nTrend = ", trend,
                                              ", tau = ", tau,
                                                   p = ", p,
                                              "\nLinear Trendline: ",
                                              "y = ", m_slope, "x + ", m_int)) +
         theme_bw() + theme(plot.title = element_text(face="bold", hjust = 0.5),
                            panel.border = element_blank(),
                            panel.grid.major = element_blank(),
                            panel.grid.minor = element_blank(),
                            axis.line = element_blank())
      print(ggarrange(p0, KTset, ncol = 1, heights = c(0.15, 1)))
      rm(plot_data)
   }
}
```

Data Points with Trendlines for Alligator Harbor Aquatic Preserve

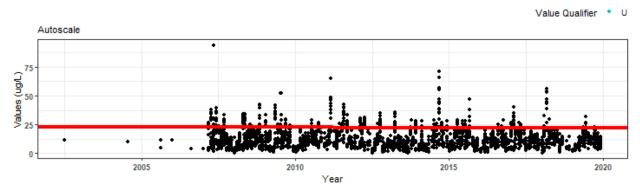
Senn Slope = 0.153846153846154, Senn Intercept = -339.1 Trend = 1, tau = 0.1885, p = 0 Linear Trendline: y = 0.244995307604981x + <math>-485.052006562353

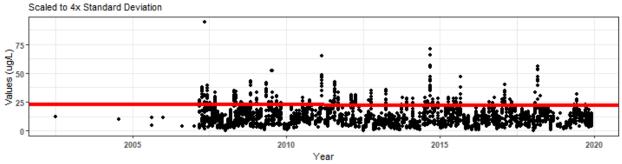




Data Points with Trendlines for Apalachicola Bay Aquatic Preserve

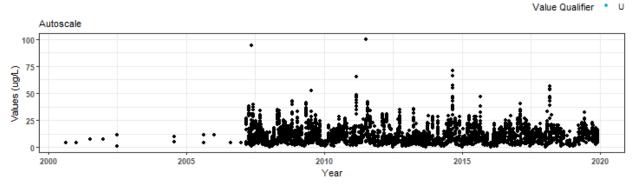
Senn Slope = -0.05183333333333334, Senn Intercept = 127.296 Trend = 0, tau = -0.0178, p = 0.0527 Linear Trendline: y = -0.157022402081742x + 328.88865400802

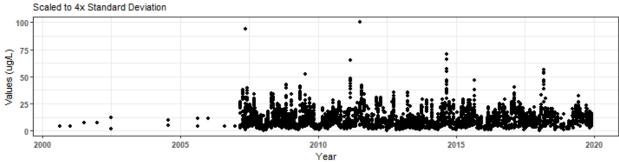




Data Points with Trendlines for Apalachicola National Estuarine Research Reserve

Senn Slope = 0, Senn Intercept = -14.4 Trend = 0, tau = 0.0125, p = 0.5574 Linear Trendline: y = -0.0648928345665301x + 142.296717728925

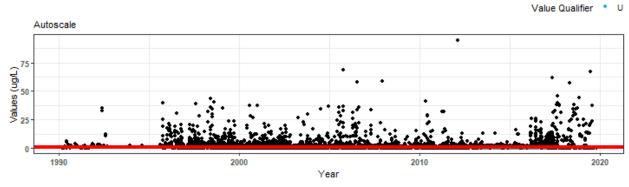


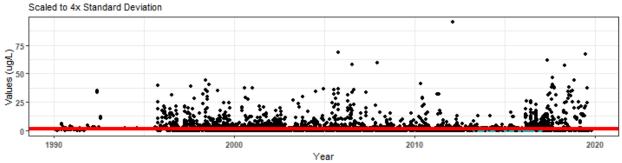


Data Points with Trendlines for Big Bend Seagrasses Aquatic Preserve

Senn Slope = 0, Senn Intercept = 1

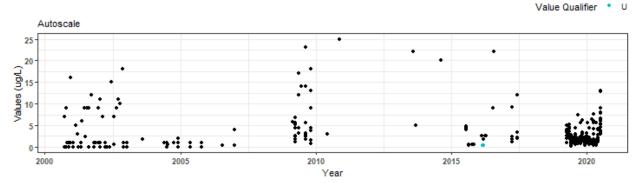
Trend = -1, tau = -0.068, p = 0 Linear Trendline: y = -0.020422123816274x + 43.9759402161498

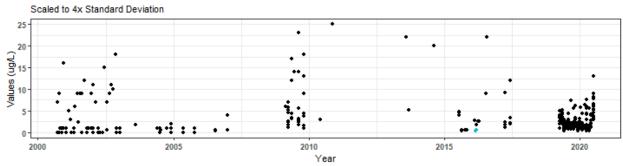




Data Points with Trendlines for Biscayne Bay Aquatic Preserve

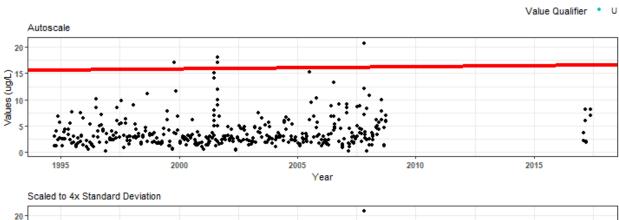
Senn Slope = 0.01, Senn Intercept = 24.4306060606061 Trend = 0, tau = 0.0485, p = 0.1212 Linear Trendline: y = -0.0536176282615864x + 110.600107388218

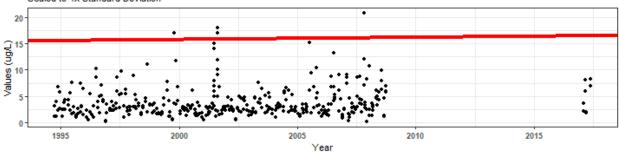




Data Points with Trendlines for Cape Romano-Ten Thousand Islands Aquatic Preserve

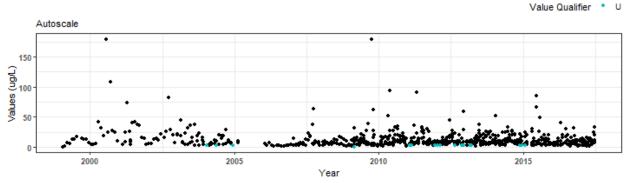
Senn Slope = 0.04125, Senn Intercept = -66.6063325Trend = 1, tau = 0.075, p = 0.0475Linear Trendline: y = 0.0797432737942312x + -155.85791227074

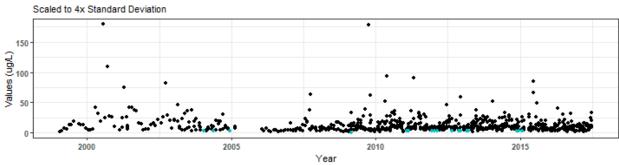




Data Points with Trendlines for Cockroach Bay Aquatic Preserve

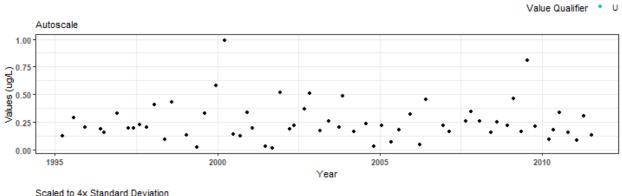
Senn Slope = -0.03333333333333333. Senn Intercept = 308.221875 Trend = 0, tau = -0.0204, p = 0.4646 Linear Trendline: y = -0.450939780539659x + 919.449699853098

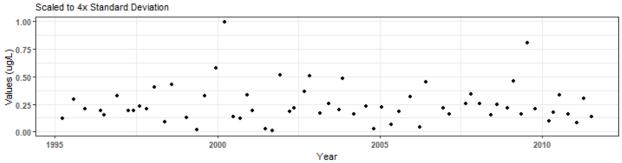




Data Points with Trendlines for Coupon Bight Aquatic Preserve

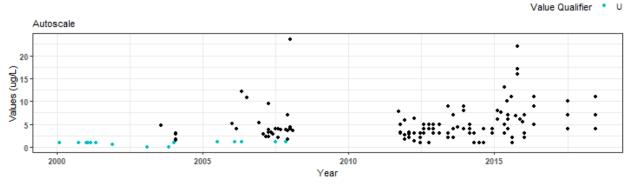
Senn Slope = -0.00366114923076923, Senn Intercept = 1.23763791741477 Trend = 0, tau = -0.0294, p = 0.3537 Linear Trendline: y = -0.000371266297656604x + 0.994330662583425

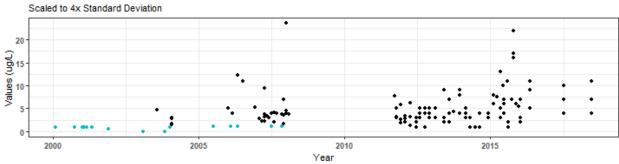




Data Points with Trendlines for Estero Bay Aquatic Preserve

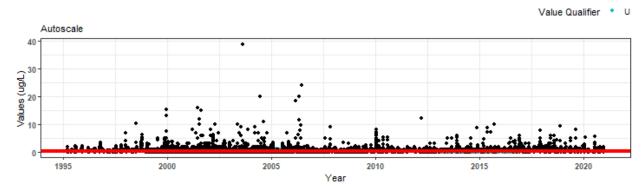
Senn Slope = 0.307692307692308, Senn Intercept = -697.630625 Trend = 1, tau = 0.3399, p = 0 Linear Trendline: y = 0.300326823529334x + -599.663741921542

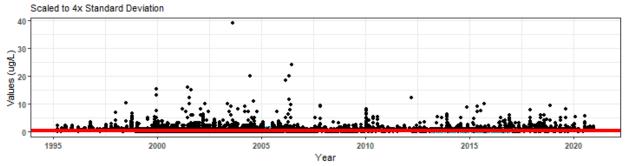




Data Points with Trendlines for Florida Keys National Marine Sanctuary

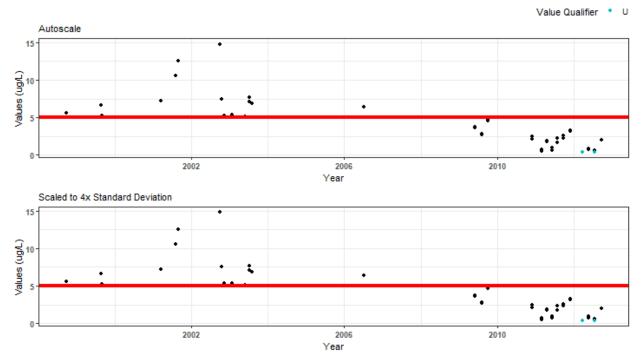
Senn Slope = 0.003066666666666667, Senn Intercept = -5.85250926 Trend = 1, tau = 0.0669, p = 0 Linear Trendline: y = 0.00126862724454288x + <math>-2.10814297092954





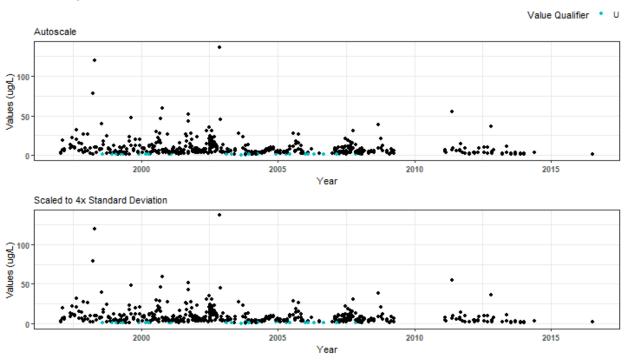
Data Points with Trendlines for Fort Pickens State Park Aquatic Preserve

Senn Slope = 0, Senn Intercept = 5 Trend = -1, tau = -0.2417, p = 0Linear Trendline: y = -0.254683153160981x + 514.967681594146



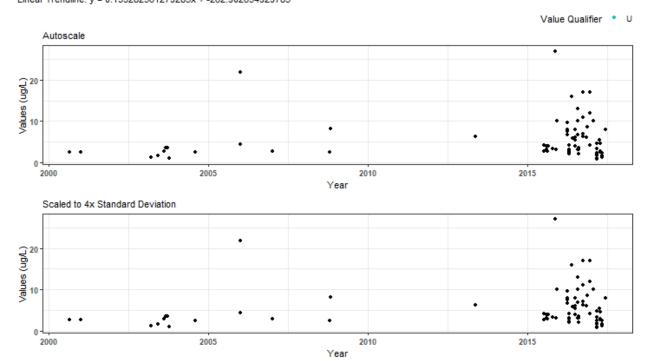
Data Points with Trendlines for Gasparilla Sound-Charlotte Harbor Aquatic Preserve

Senn Slope = -0.25, Senn Intercept = 312.8835 Trend = -1, tau = -0.1581, p = 0 Linear Trendline: y = -0.478984749284008x + 968.282586040035



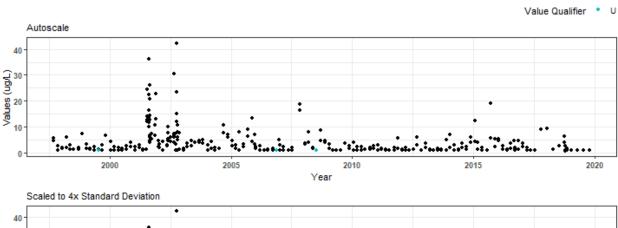
Data Points with Trendlines for Indian River-Malabar to Vero Beach Aquatic Preserve

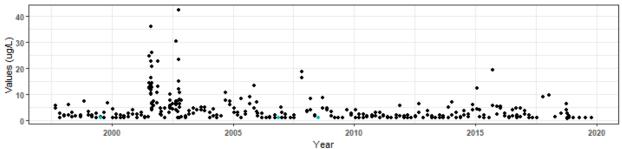
Senn Slope = 0.0384615384615385, Senn Intercept = -234.179345285714 Trend = 0, tau = 0.1367, p = 0.2513 Linear Trendline: y = 0.133282561279285x + <math>-262.902654929785



Data Points with Trendlines for Jensen Beach to Jupiter Inlet Aquatic Preserve

Senn Slope = -0.0285714285714286, Senn Intercept = 164.465 Trend = -1, tau = -0.1403, p = 0.0019 Linear Trendline: y = -0.222981685154151x + 451.988735116039





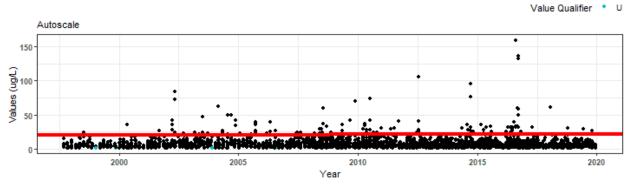
Data Points with Trendlines for Lemon Bay Aquatic Preserve

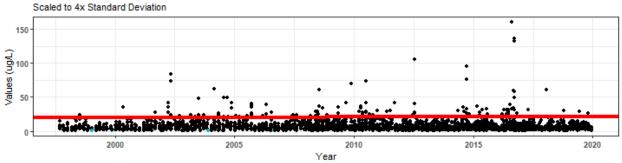
Senn Slope = 0.131, Senn Intercept = -201.23125 Trend = 1, tau = 0.236, p = 0.0012 Linear Trendline: y = 0.246022329932319x + -489.536280148704



Data Points with Trendlines for Loxahatchee River-Lake Worth Creek Aquatic Preserve

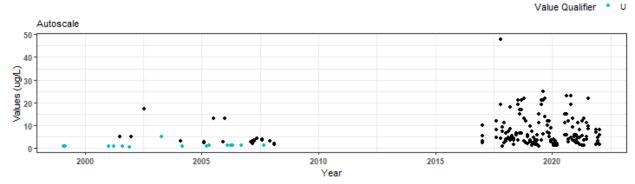
Senn Slope = 0.03750000000000001, Senn Intercept = -54.0250000000001 Trend = 1, tau = 0.0286, p = 0.0008 Linear Trendline: y = 0.061469110034196x + -115.81102173855





Data Points with Trendlines for Matlacha Pass Aquatic Preserve

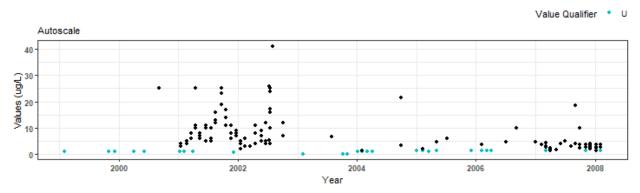
Senn Slope = 0.0941176470588235, Senn Intercept = -97.4225000000001 Trend = 1, tau = 0.1472, p = 0.0072 Linear Trendline: y = 0.285836840816889x + -569.540697255765

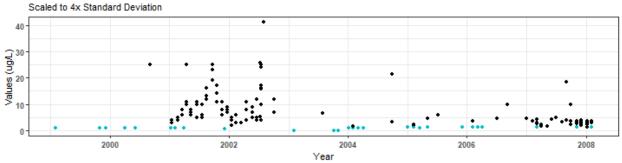




Data Points with Trendlines for Pine Island Sound Aquatic Preserve

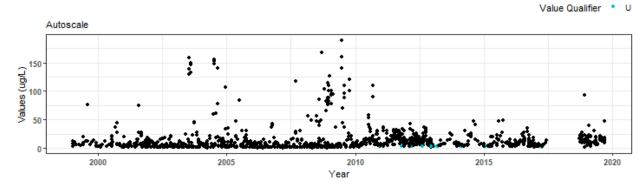
Senn Slope = -0.08, Senn Intercept = 1000.83166666667Trend = 0, tau = -0.1362, p = 0.2533Linear Trendline: y = -0.684258245498316x + 1376.92162261143

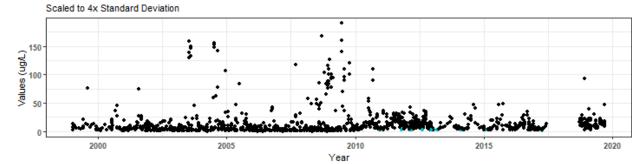




Data Points with Trendlines for Pinellas County Aquatic Preserve

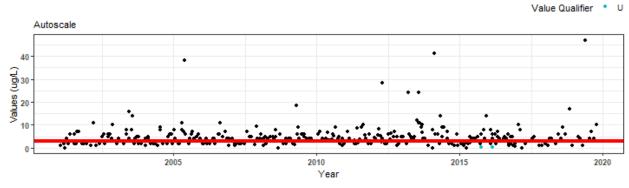
Senn Slope = 0.25, Senn Intercept = -551.498857142857 Trend = 1, tau = 0.1651, p = 0 Linear Trendline: y = 0.060346776244395x + -106.938967937032

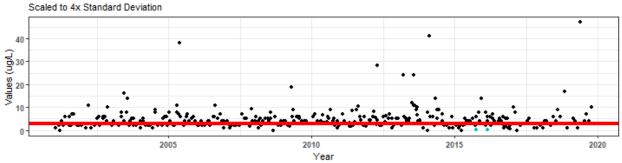




Data Points with Trendlines for Rocky Bayou State Park Aquatic Preserve

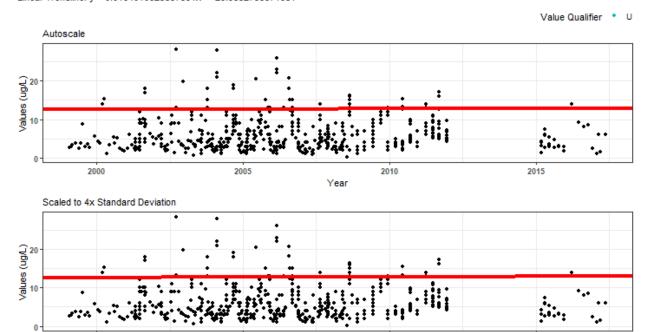
Senn Slope = 0, Senn Intercept = 3 Trend = 0, tau = 0.0235, p = 0.4879Linear Trendline: y = 0.0712738499091406x + -138.657972309424





Data Points with Trendlines for Rookery Bay Aquatic Preserve

Senn Slope = 0.01585, Senn Intercept = -18.9233334615384Trend = 0, tau = 0.023, p = 0.2752Linear Trendline: y = 0.0164010625837591x + -26.9852755571681



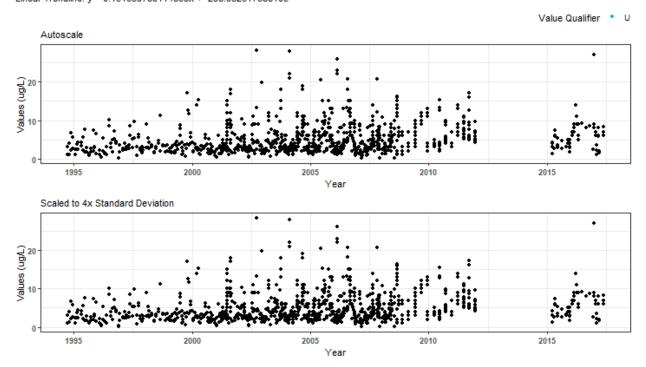
Data Points with Trendlines for Rookery Bay National Estuarine Research Reserve

2005

2010

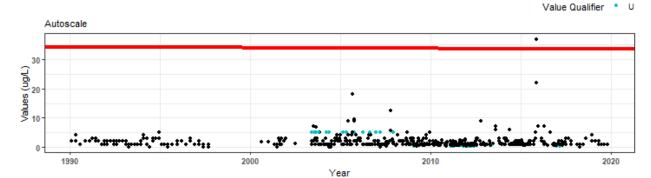
Senn Intercept = -217.0174125

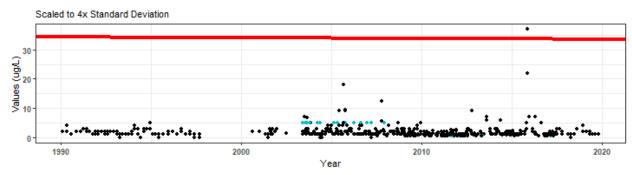
2000



Data Points with Trendlines for St. Andrews State Park Aquatic Preserve

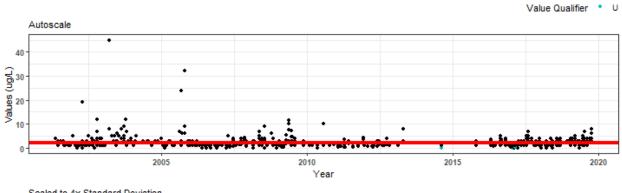
Senn Slope = -0.025, Senn Intercept = 84.1871825396825 Trend = -1, tau = -0.1862, p = 0 Linear Trendline: y = -0.040921682895982x + 84.1810730272178

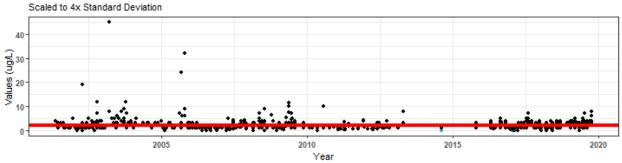




Data Points with Trendlines for St. Joseph Bay Aquatic Preserve

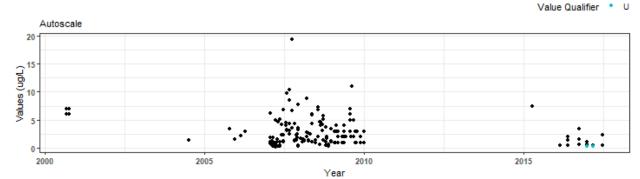
Senn Slope = 0, Senn Intercept = 2 Trend = 0, tau = 0.0321, p = 0.2778Linear Trendline: y = -0.0341425355946756x + 70.9909136605636

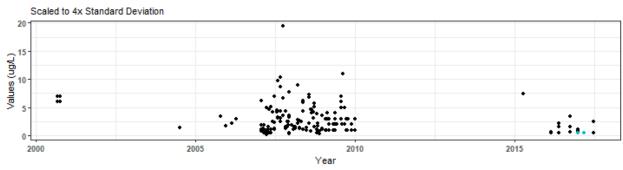




Data Points with Trendlines for St. Martins Marsh Aquatic Preserve

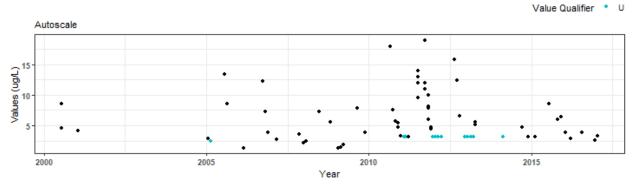
 $Senn\ Slope = -0.111805555555556, \qquad Senn\ Intercept = 134.128$ $Trend = -1, \qquad tau = -0.1281, \qquad p = 0.0086$ $Linear\ Trendline:\ y = -0.21094720498719x + 426.489882811847$

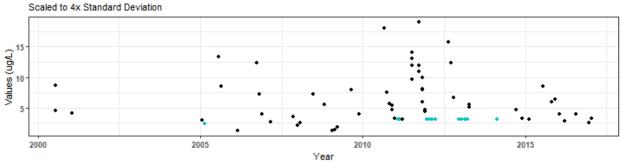




Data Points with Trendlines for Terra Ceia Aquatic Preserve

Senn Slope = 0.0571428571428572, Senn Intercept = 3.1 Trend = 1, tau = 0.1677, p = 0.0407 Linear Trendline: y = -0.023460832157331x + <math>53.2442469973497





Appendix V: Managed Area Summary Box Plots

Data is taken and grouped by ManagedAreaName. The scripts that create plots follow this format

- 1. Use the data set that only has Use_In_Analysis of TRUE for the desired managed area
- 2. Determine the earliest and latest year of the data to create x-axis scale and intervals
- 3. Determine the minimum, mean, and standard deviation for the data to be used for y-axis scales
 - Excludes the top 2% of values to reduce the impact of extreme outliers on the y-axis scale
- 4. Set what values are to be used for the x-axis, y-axis, and the variable that should determine groups for the box plots
- 5. Set the plot type as a box plot with the size of the outlier points
- 6. Create the title, x-axis, y-axis, and color fill labels
- 7. Set the y and x limits
- 8. Make the axis labels bold
- 9. Plot the arrangement as a set of panels

The following plots are arranged by ManagedAreaName with data grouped by Year, then Year and Month, then finally Month only. Each managed area will have 3 sets of plots, each with 3 panels in them. Each panel goes as follows:

- 1. Y-axis autoscaled
- 2. Y-axis set to be mean + 4 times the standard deviation
- 3. Y-axis set to be mean + 4 times the standard deviation for most recent 10 years of data

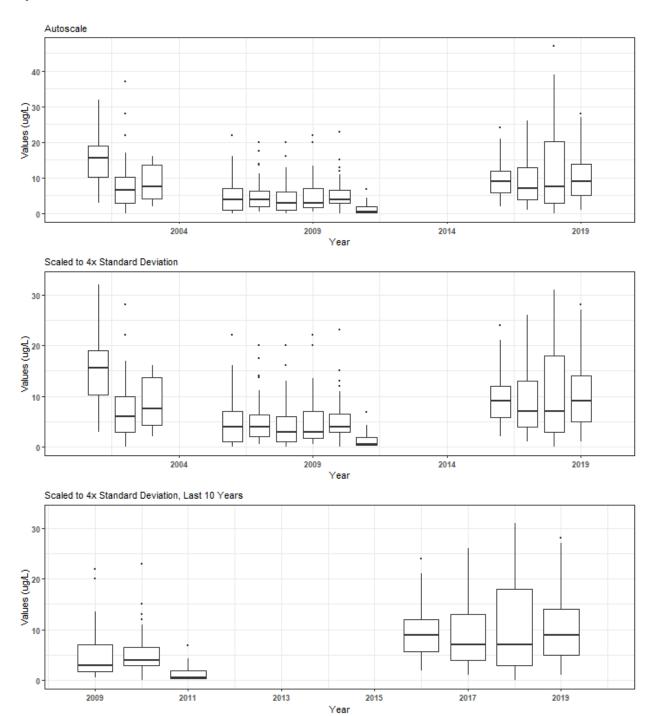
```
if(n==0){
   print("There are no managed areas that qualify.")
} else {
   for (i in 1:n) {
      year_lower <- min(data$Year[data$Use_In_Analysis == TRUE &</pre>
                                       data$ManagedAreaName == MA_names[i]])
      year_upper <- max(data$Year[data$Use_In_Analysis == TRUE &</pre>
                                       data$ManagedAreaName == MA_names[i]])
      min_RV \leftarrow min(data\$ResultValue[data\$Use_In_Analysis == TRUE &
                                          data$ManagedAreaName == MA names[i]])
      mn RV <- mean(data$ResultValue[data$Use In Analysis == TRUE &
                                          data$ManagedAreaName == MA names[i] &
                                          data$ResultValue <</pre>
                                          quantile(data$ResultValue, 0.98)])
      sd_RV <- sd(data$ResultValue[data$Use_In_Analysis == TRUE &</pre>
                                        data$ManagedAreaName == MA_names[i] &
                                        data$ResultValue <</pre>
                                        quantile(data$ResultValue, 0.98)])
      x_scale <- ifelse(year_upper - year_lower > 30, 10, 5)
      y_scale <- mn_RV + 4 * sd_RV</pre>
      ##Year plots
      p1 <- ggplot(data = data[data$Use_In_Analysis == TRUE &
                                   data$ManagedAreaName == MA_names[i], ],
                    aes(x = Year, y = ResultValue, group = Year)) +
         geom_boxplot(outlier.size = 0.5) +
         labs(subtitle = "Autoscale",
              x = "Year", y = paste0("Values (", unit, ")")) +
```

```
scale_x_continuous(limits = c(year_lower - 1, year_upper + 1),
                      breaks = rev(seq(year_upper,
                                       year_lower, -x_scale))) +
   theme bw() +
   theme(axis.text.x = element_text(face = "bold"),
         axis.text.y = element_text(face = "bold"))
p2 <- ggplot(data = data[data$Use In Analysis == TRUE &
                            data$ManagedAreaName == MA_names[i], ],
             aes(x = Year, y = ResultValue, group = Year)) +
   geom boxplot(outlier.size = 0.5) +
   labs(subtitle = "Scaled to 4x Standard Deviation",
        x = "Year", y = paste0("Values (", unit, ")")) +
   ylim(min_RV, y_scale) +
   scale_x_continuous(limits = c(year_lower - 1, year_upper + 1),
                      breaks = rev(seq(year_upper,
                                       year_lower, -x_scale))) +
   theme_bw() +
   theme(axis.text.x = element_text(face = "bold"),
         axis.text.y = element_text(face = "bold"))
p3 <- ggplot(data = data[data$Use_In_Analysis == TRUE &
                            data$ManagedAreaName == MA_names[i] &
                            data$Year>=year_upper-10, ],
             aes(x = Year, y = ResultValue, group = Year)) +
   geom boxplot(outlier.size = 0.5) +
   labs(subtitle = "Scaled to 4x Standard Deviation, Last 10 Years",
        x = "Year", y = paste0("Values (", unit, ")")) +
   ylim(min_RV, y_scale) +
   scale_x_continuous(limits = c(year_upper - 10.5, year_upper + 1),
                      breaks = rev(seq(year_upper, year_upper - 10,-2))) +
   theme_bw() +
   theme(axis.text.x = element_text(face = "bold"),
         axis.text.y = element_text(face = "bold"))
Yset <- ggarrange(p1, p2, p3, ncol = 1)</pre>
p0 <- ggplot() + labs(title = paste0("Summary Box Plots for ",</pre>
                                     MA_names[i]), subtitle = "By Year") +
   theme_bw() + theme(plot.title = element_text(face="bold", hjust = 0.5),
                      panel.border = element_blank(),
                      panel.grid.major = element_blank(),
                      panel.grid.minor = element blank(),
                      axis.line = element_blank())
## Year & Month Plots
p4 <- ggplot(data = data[data$Use_In_Analysis == TRUE &
                            data$ManagedAreaName == MA_names[i], ],
             aes(x = YearMonthDec, y = ResultValue,
                 group = YearMonth, color = as.factor(Month))) +
   geom_boxplot(outlier.size = 0.5) +
   labs(subtitle = "Autoscale",
```

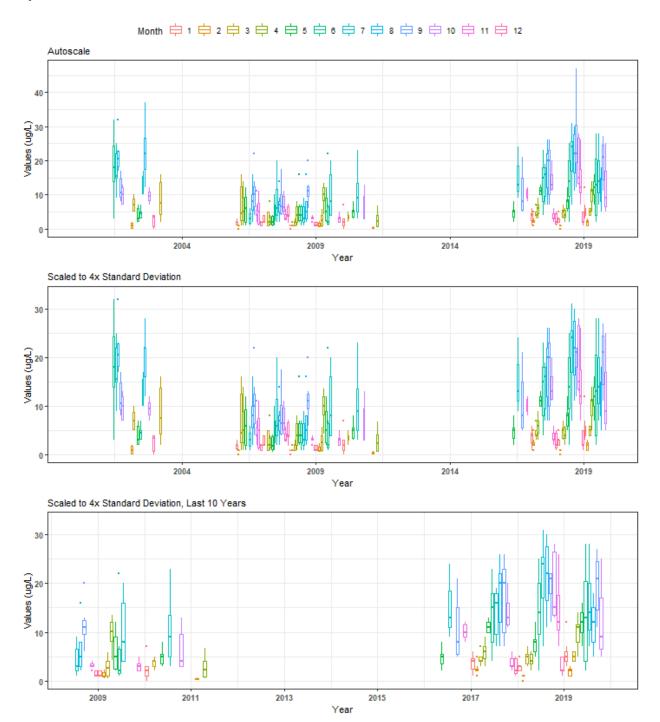
```
x = "Year", y = pasteO("Values (", unit, ")"), color = "Month") +
   scale_x_continuous(limits = c(year_lower - 1, year_upper + 1),
                      breaks = rev(seq(year_upper,
                                       year_lower, -x_scale))) +
   theme_bw() +
   theme(legend.position = "none",
         axis.text.x = element_text(face = "bold"),
         axis.text.y = element text(face = "bold"))
p5 <- ggplot(data = data[data$Use_In_Analysis == TRUE &
                            data$ManagedAreaName == MA_names[i], ],
             aes(x = YearMonthDec, y = ResultValue,
                 group = YearMonth, color = as.factor(Month))) +
   geom_boxplot(outlier.size = 0.5) +
   labs(subtitle = "Scaled to 4x Standard Deviation",
        x = "Year", y = paste0("Values (", unit, ")"), color = "Month") +
   ylim(min_RV, y_scale) +
   scale_x_continuous(limits = c(year_lower - 1, year_upper + 1),
                      breaks = rev(seq(year_upper,
                                       year_lower, -x_scale))) +
   theme_bw() +
   theme(legend.position = "top", legend.box = "horizontal",
         axis.text.x = element_text(face = "bold"),
         axis.text.y = element_text(face = "bold")) +
   guides(color = guide legend(nrow = 1))
p6 <- ggplot(data = data[data$Use_In_Analysis == TRUE &
                            data$ManagedAreaName == MA_names[i], ],
             aes(x = YearMonthDec, y = ResultValue,
                group = YearMonth, color = as.factor(Month)
             )) +
   geom_boxplot(outlier.size = 0.5) +
   labs(subtitle = "Scaled to 4x Standard Deviation, Last 10 Years",
        x = "Year", y = paste0("Values (", unit, ")"), color = "Month") +
   ylim(min_RV, y_scale) +
   scale_x_continuous(limits = c(year_upper - 10.5, year_upper + 1),
                      breaks = rev(seq(year_upper, year_upper - 10,-2))) +
   theme bw() +
   theme(legend.position = "none",
         axis.text.x = element_text(face = "bold"),
         axis.text.y = element_text(face = "bold"))
leg1 <- get_legend(p5)</pre>
YMset <- ggarrange(leg1, p4, p5 + theme(legend.position = "none"), p6,
                   ncol = 1, heights = c(0.1, 1, 1, 1)
p00 <- ggplot() + labs(title = paste0("Summary Box Plots for ",
                                      MA_names[i]),
                       subtitle = "By Year & Month") + theme_bw() +
   theme(plot.title = element_text(face="bold", hjust = 0.5),
         panel.border = element_blank(),
         panel.grid.major = element_blank(),
         panel.grid.minor = element_blank(), axis.line = element_blank())
```

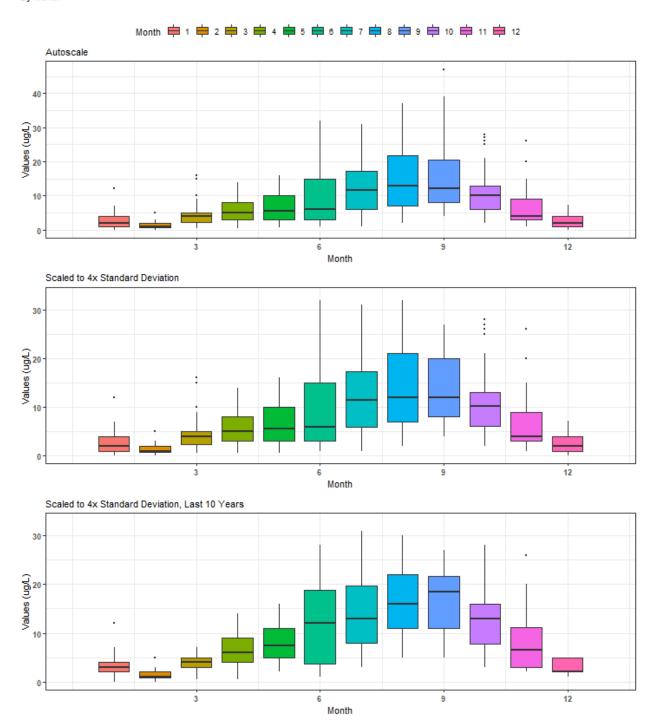
```
## Month Plots
p7 <- ggplot(data = data[data$Use_In_Analysis == TRUE &
                            data$ManagedAreaName == MA_names[i], ],
             aes(x = Month, y = ResultValue,
                 group = Month, fill = as.factor(Month))) +
   geom_boxplot(outlier.size = 0.5) +
   labs(subtitle = "Autoscale",
        x = "Month", y = pasteO("Values (", unit, ")"), fill = "Month") +
   scale x continuous(limits = c(0, 13), breaks = seq(3, 12, 3)) +
   theme_bw() +
   theme(legend.position = "none",
         axis.text.x = element_text(face = "bold"),
         axis.text.y = element_text(face = "bold"))
p8 <- ggplot(data = data[data$Use_In_Analysis == TRUE &
                            data$ManagedAreaName == MA_names[i], ],
             aes(x = Month, y = ResultValue,
                 group = Month, fill = as.factor(Month))) +
   geom_boxplot(outlier.size = 0.5) +
   labs(subtitle = "Scaled to 4x Standard Deviation",
        x = "Month", y = paste0("Values (", unit, ")"), fill = "Month") +
   ylim(min_RV, y_scale) +
   scale_x_continuous(limits = c(0, 13), breaks = seq(3, 12, 3)) +
   theme_bw() +
   theme(legend.position = "top", legend.box = "horizontal",
         axis.text.x = element_text(face = "bold"),
         axis.text.y = element_text(face = "bold")) +
   guides(fill = guide_legend(nrow = 1))
p9 <- ggplot(data = data[data$Use_In_Analysis == TRUE &
                            data$ManagedAreaName == MA_names[i] &
                            data$Year >= year_upper - 10, ],
             aes(x = Month, y = ResultValue,
                 group = Month, fill = as.factor(Month))) +
   geom_boxplot(outlier.size = 0.5) +
   labs(subtitle = "Scaled to 4x Standard Deviation, Last 10 Years",
        x = "Month", y = paste0("Values (", unit, ")"), fill = "Month") +
   ylim(min RV, y scale) +
   scale_x_continuous(limits = c(0, 13), breaks = seq(3, 12, 3)) +
   theme bw() +
   theme(legend.position = "none",
         axis.text.x = element_text(face = "bold"),
         axis.text.y = element_text(face = "bold"))
leg2 <- get_legend(p8)</pre>
Mset <- ggarrange(leg2, p7, p8 + theme(legend.position = "none"), p9,</pre>
                  ncol = 1, heights = c(0.1, 1, 1, 1)
p000 <- ggplot() + labs(title = paste0("Summary Box Plots for ",
                                       MA_names[i]),
                        subtitle = "By Month") + theme_bw() +
   theme(plot.title = element_text(face="bold", hjust = 0.5),
         panel.border = element_blank(),
```

Summary Box Plots for Alligator Harbor Aquatic Preserve

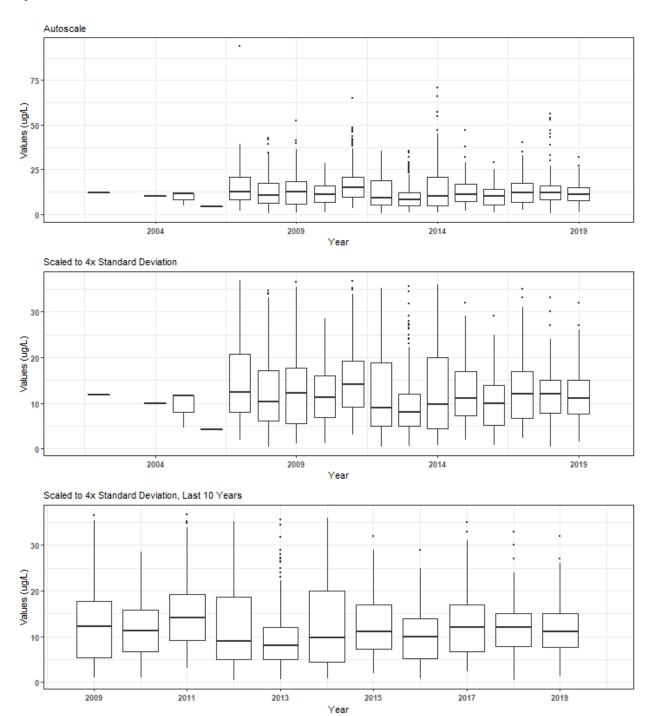


Summary Box Plots for Alligator Harbor Aquatic Preserve

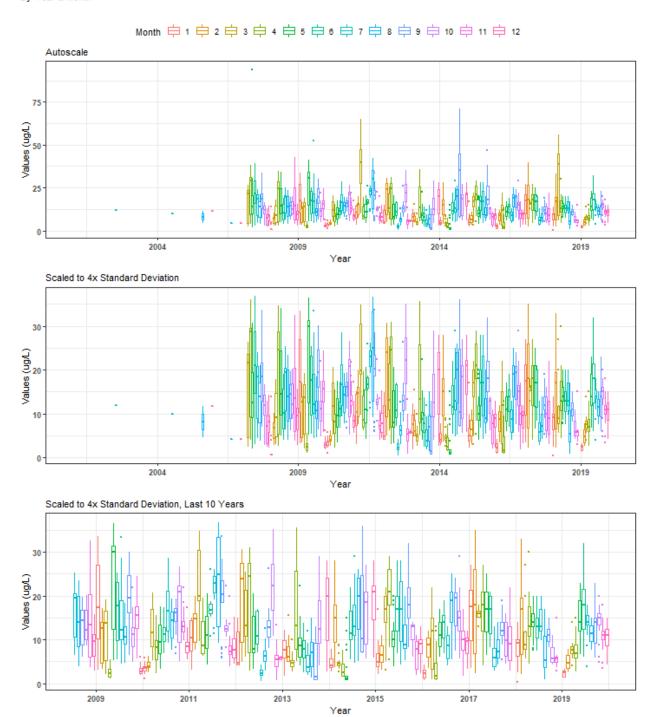


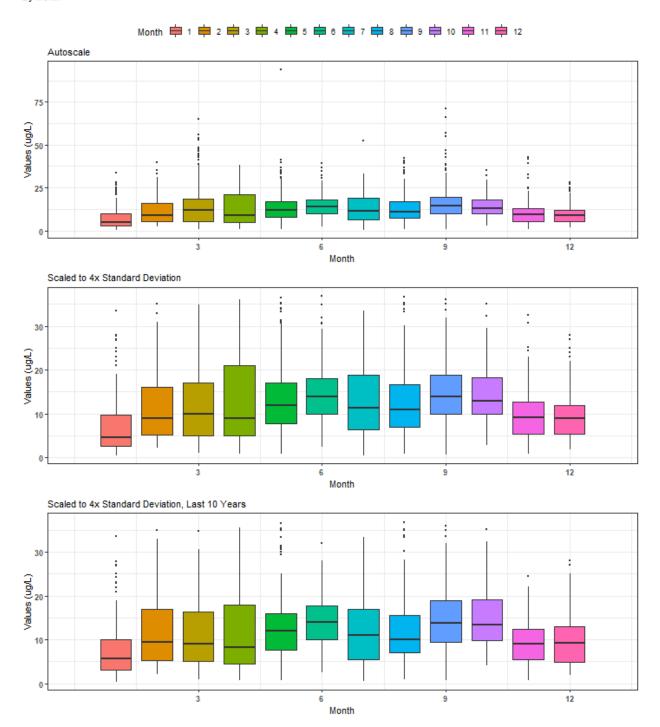


Summary Box Plots for Apalachicola Bay Aquatic Preserve

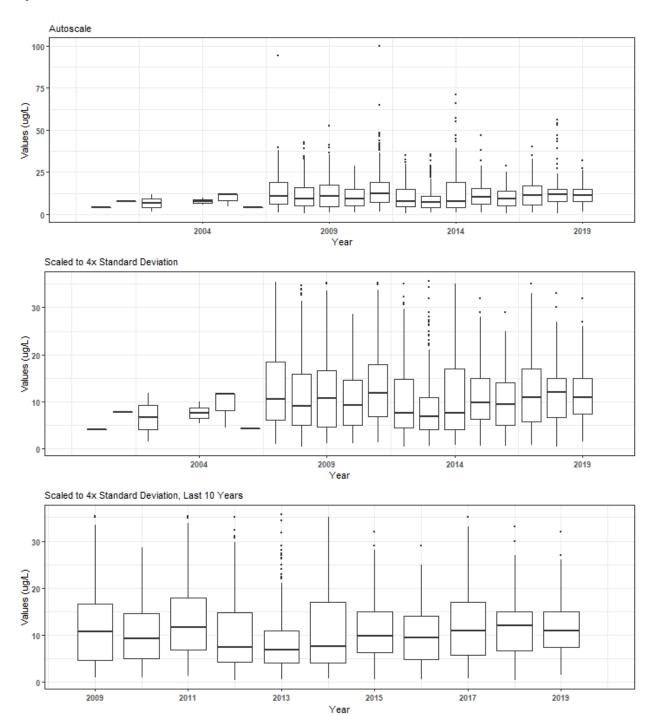


Summary Box Plots for Apalachicola Bay Aquatic Preserve

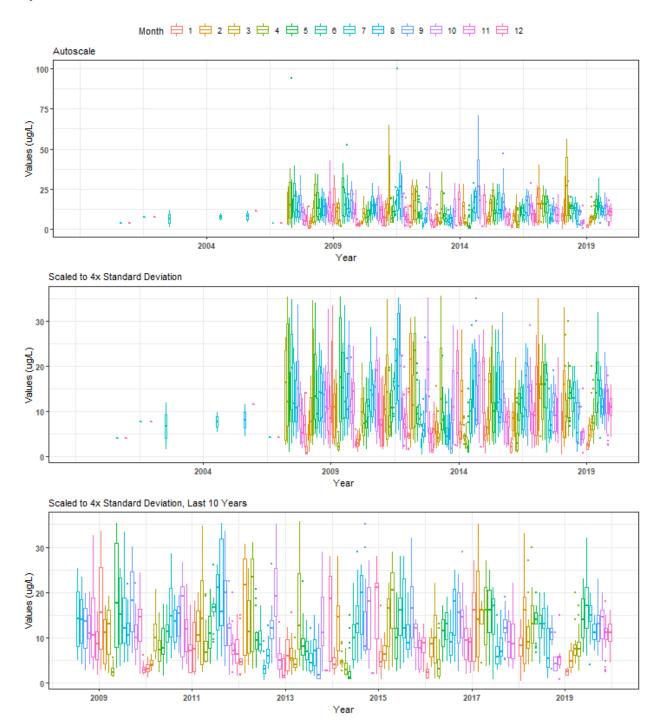




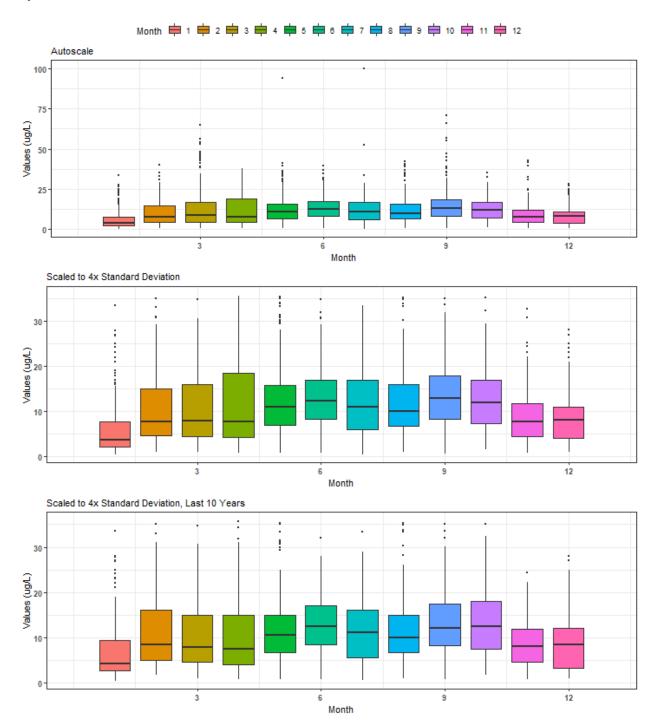
Summary Box Plots for Apalachicola National Estuarine Research Reserve



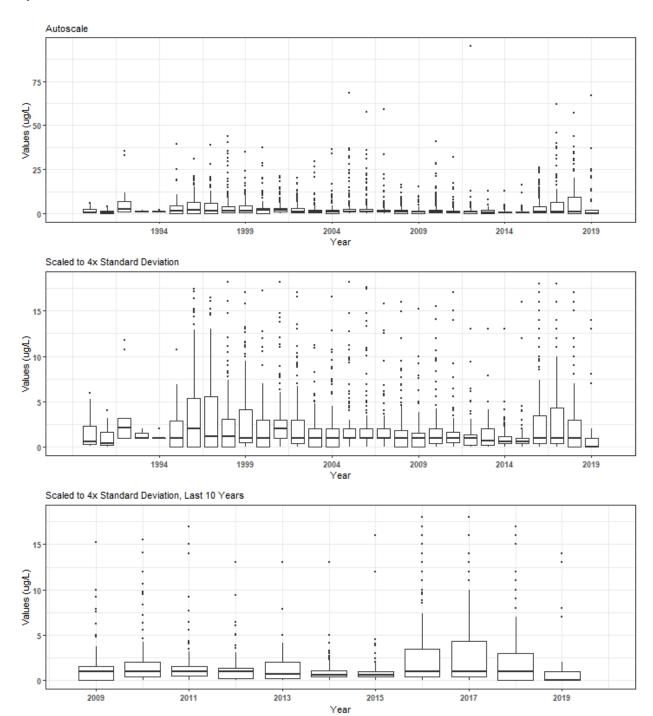
Summary Box Plots for Apalachicola National Estuarine Research Reserve



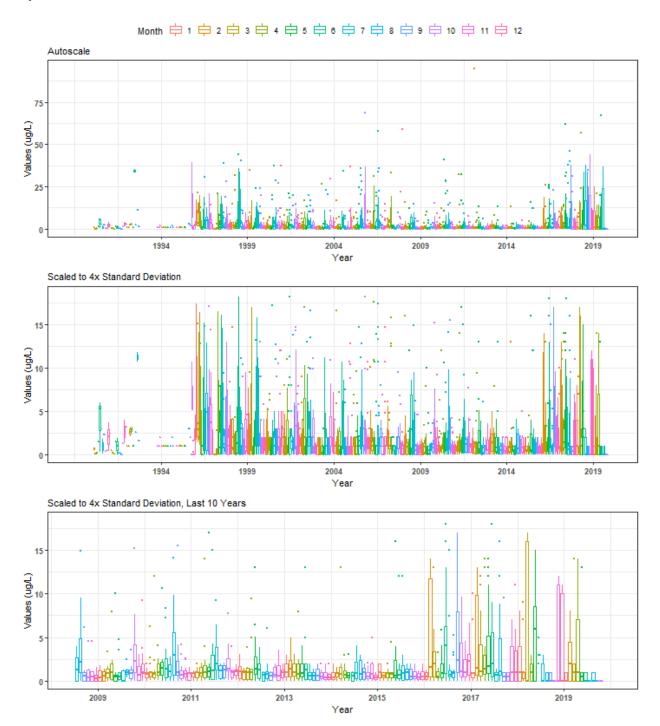
Summary Box Plots for Apalachicola National Estuarine Research Reserve



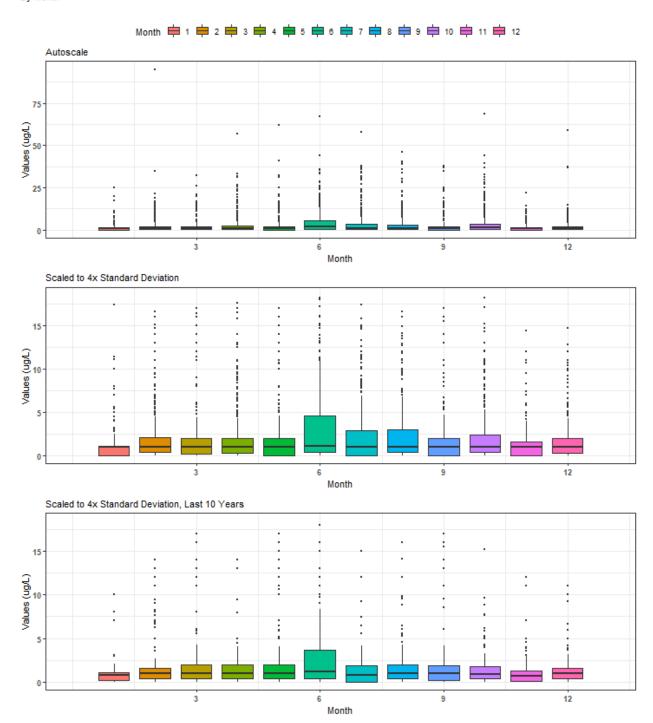
Summary Box Plots for Big Bend Seagrasses Aquatic Preserve



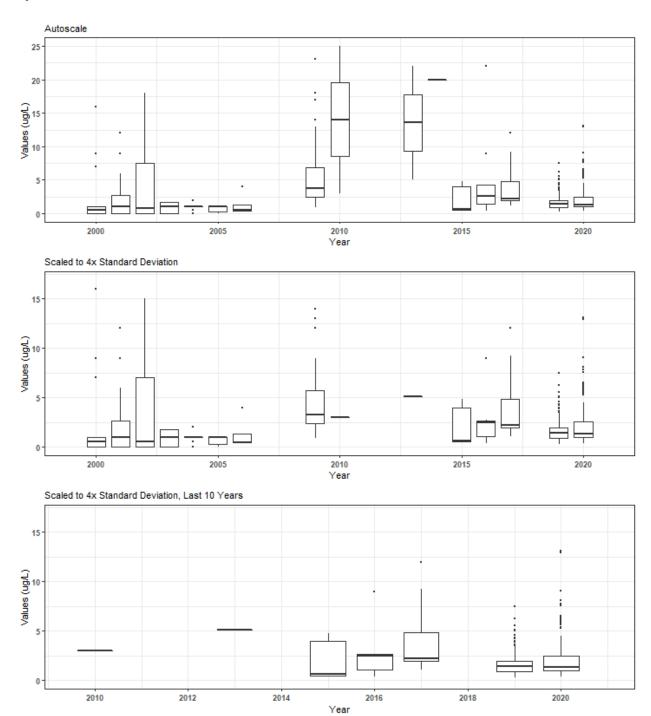
Summary Box Plots for Big Bend Seagrasses Aquatic Preserve

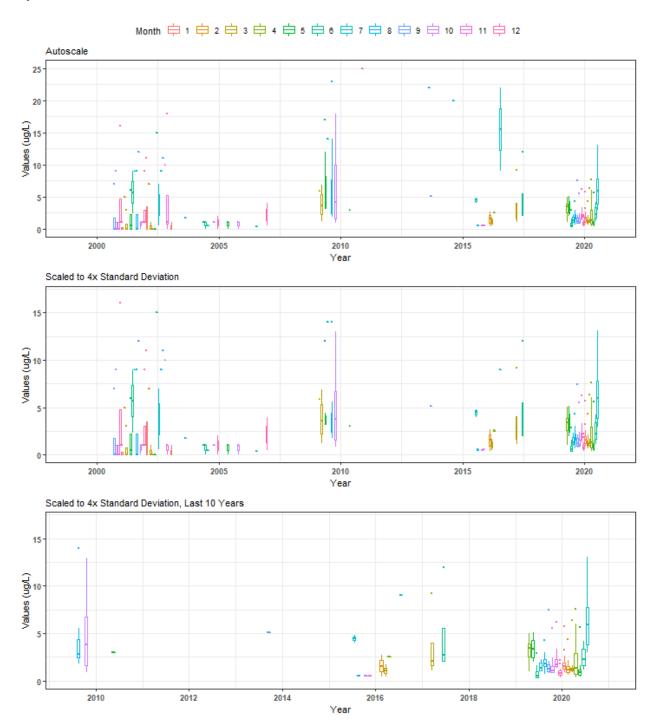


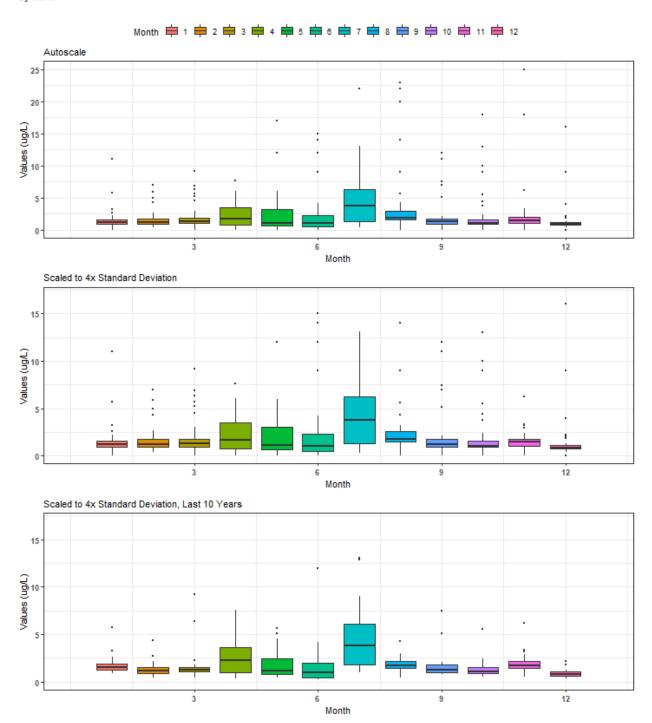
Summary Box Plots for Big Bend Seagrasses Aquatic Preserve



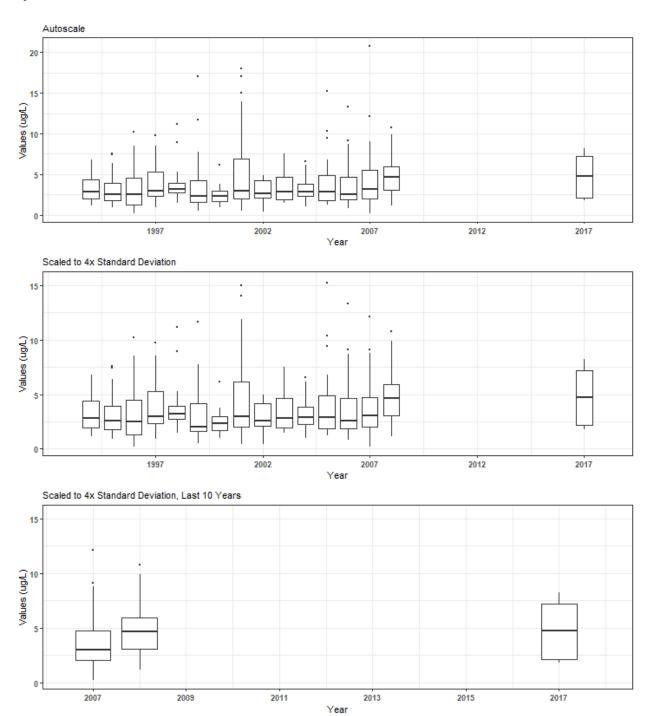
Summary Box Plots for Biscayne Bay Aquatic Preserve



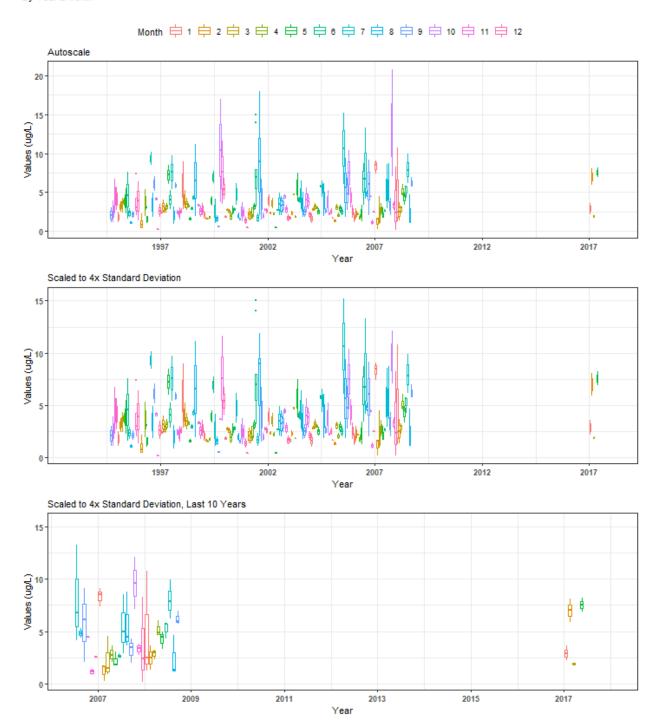




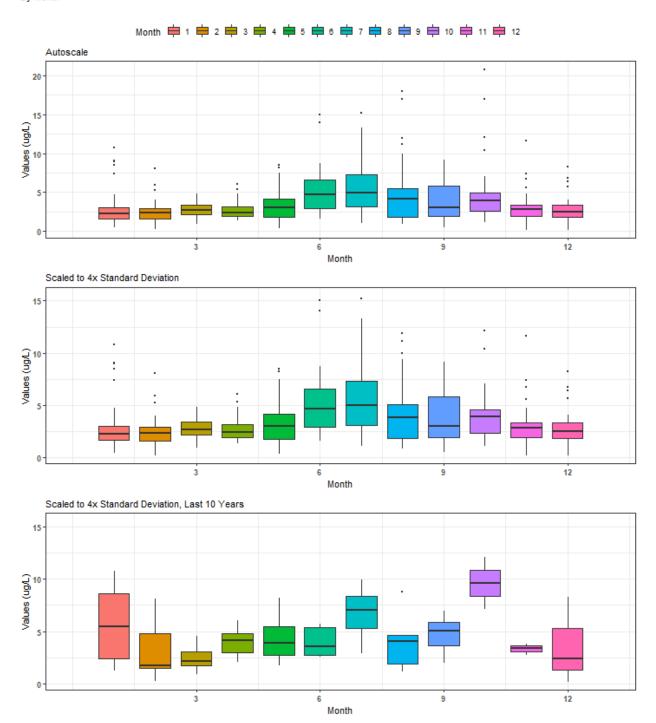
Summary Box Plots for Cape Romano-Ten Thousand Islands Aquatic Preserve



Summary Box Plots for Cape Romano-Ten Thousand Islands Aquatic Preserve

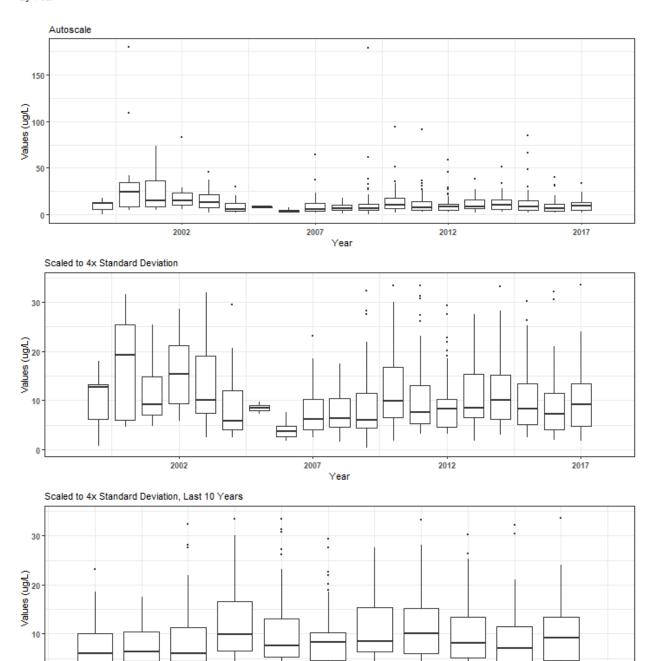


Summary Box Plots for Cape Romano-Ten Thousand Islands Aquatic Preserve



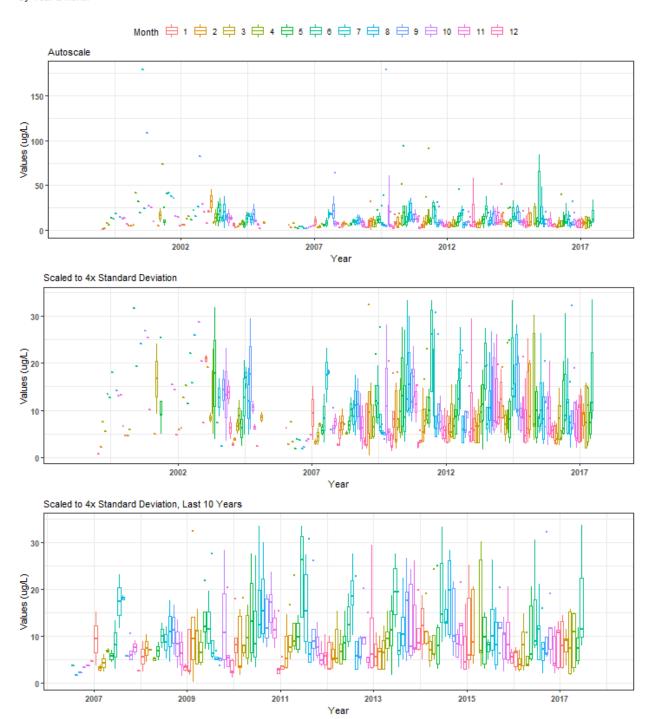
Summary Box Plots for Cockroach Bay Aquatic Preserve

By Year

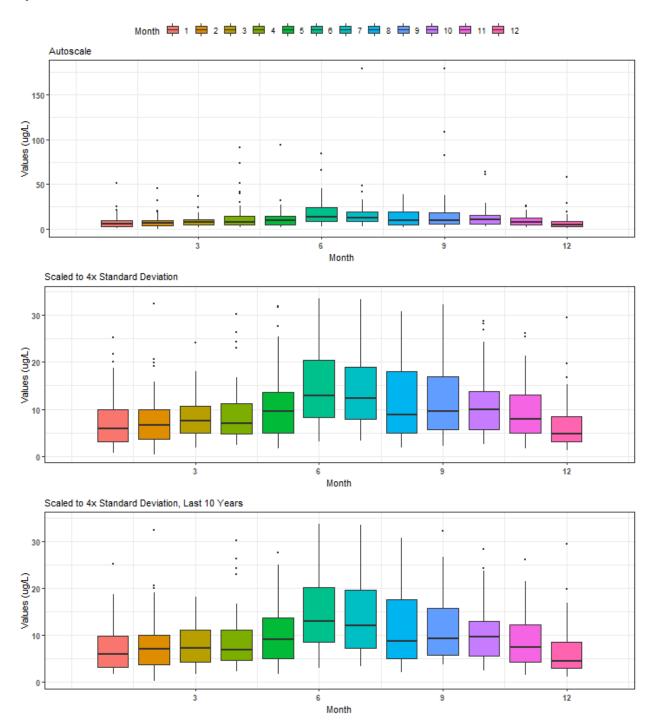


Year

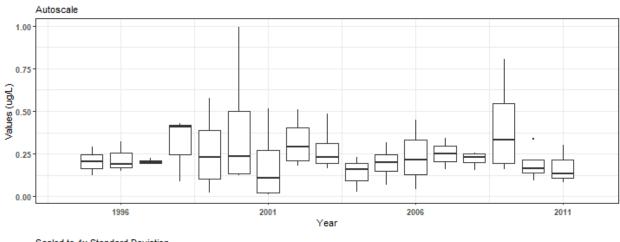
Summary Box Plots for Cockroach Bay Aquatic Preserve

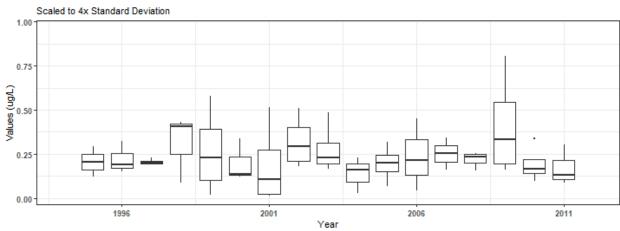


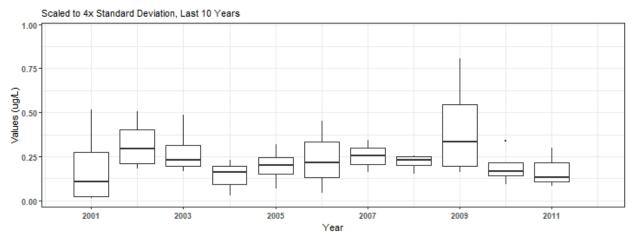
Summary Box Plots for Cockroach Bay Aquatic Preserve



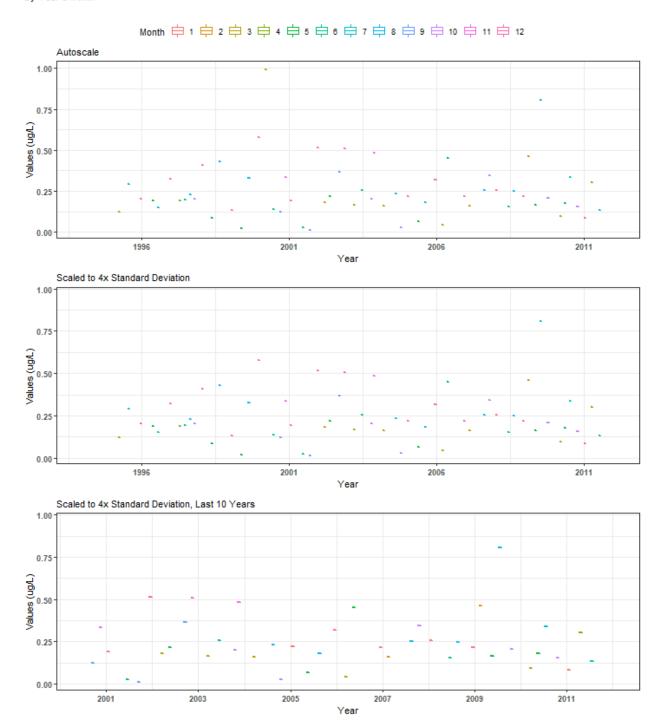
Summary Box Plots for Coupon Bight Aquatic Preserve

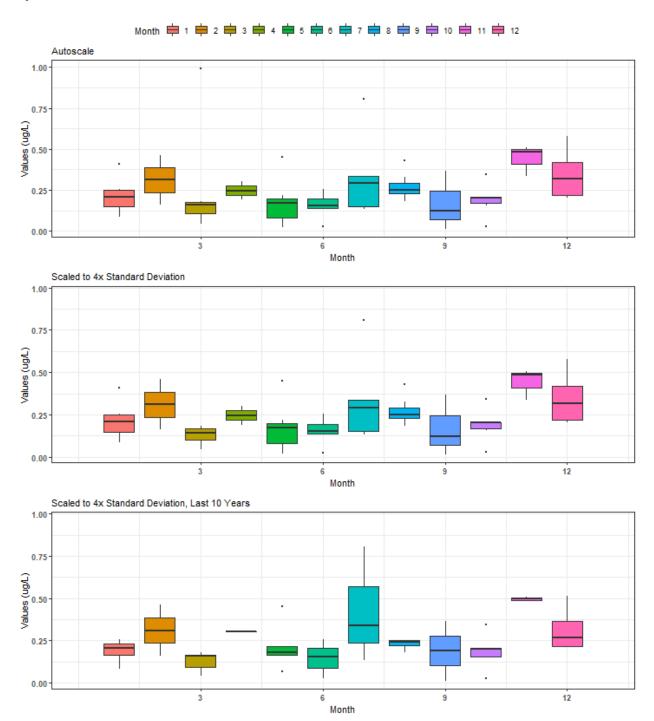






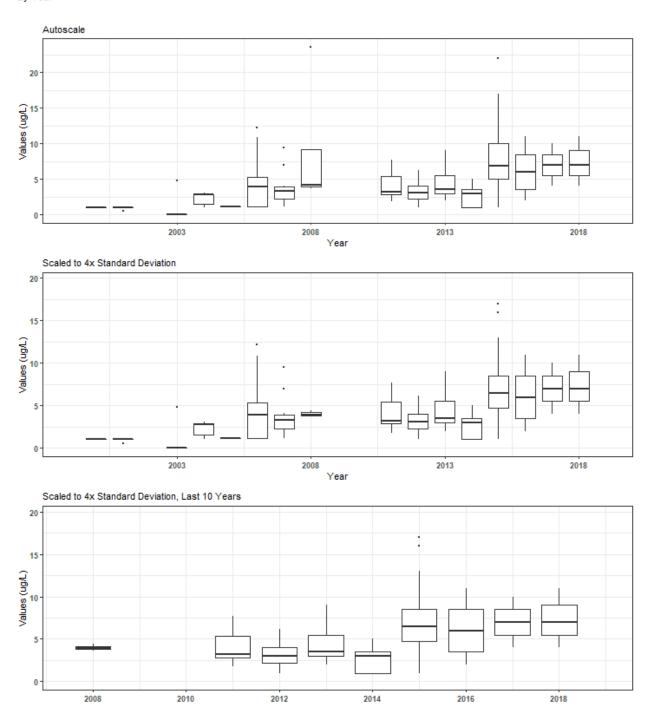
Summary Box Plots for Coupon Bight Aquatic Preserve





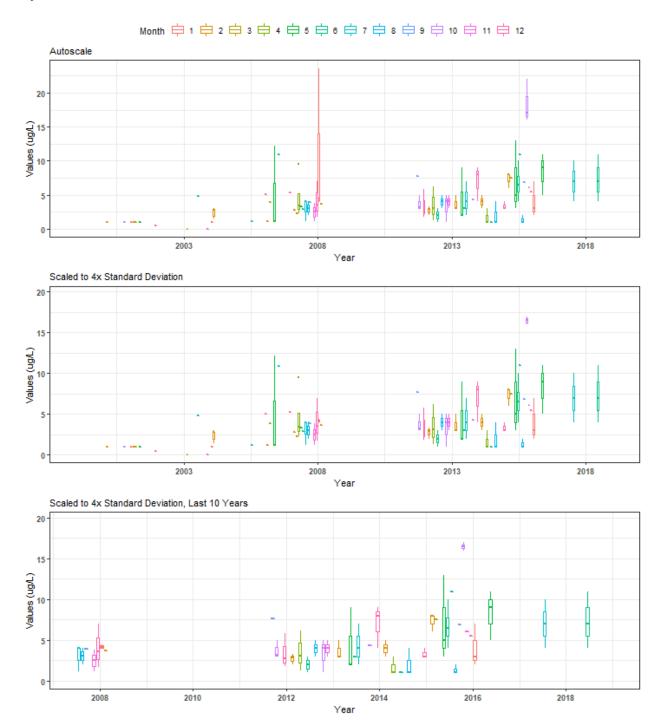
Summary Box Plots for Estero Bay Aquatic Preserve

By Year

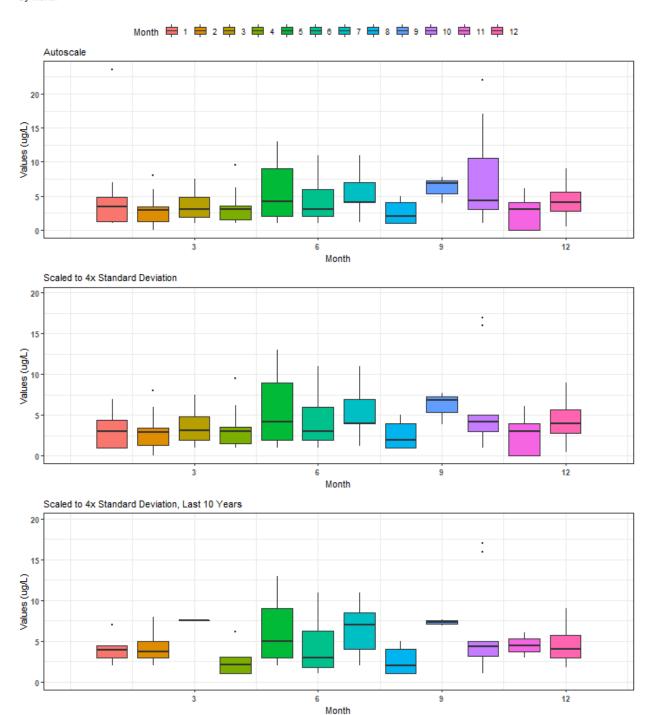


Year

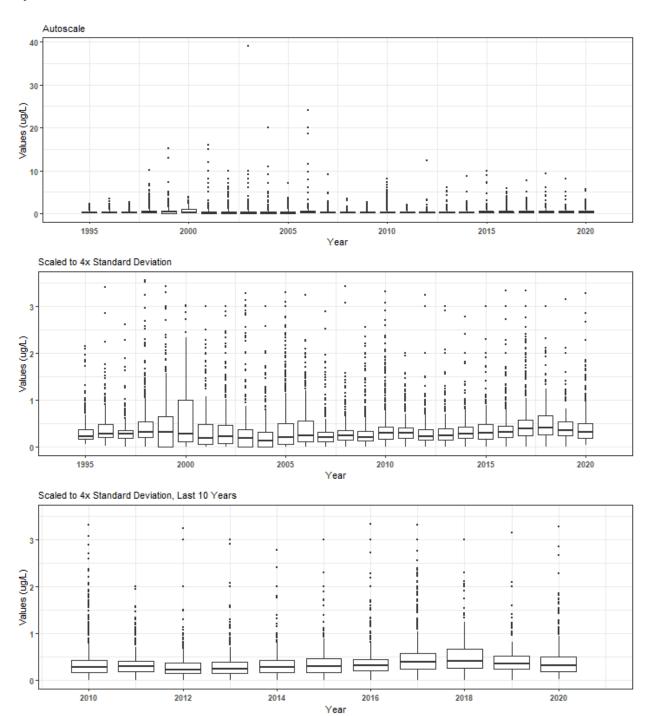
Summary Box Plots for Estero Bay Aquatic Preserve



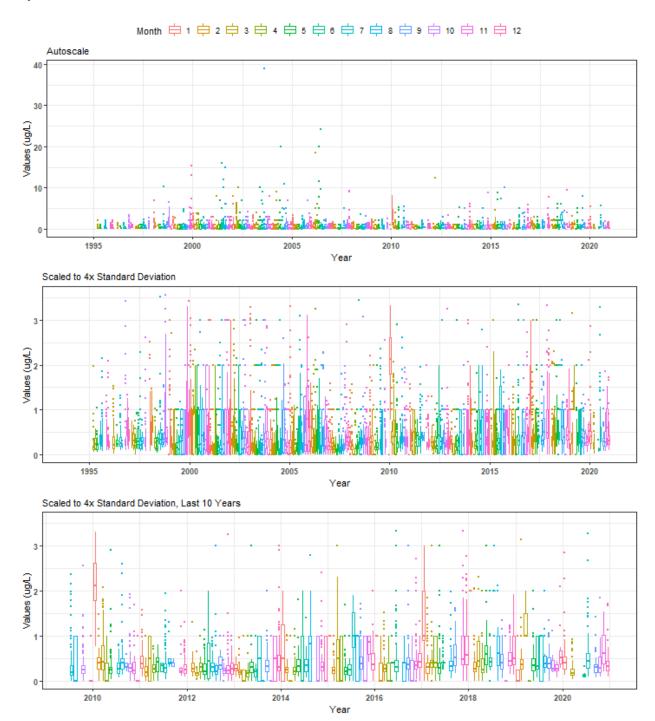
Summary Box Plots for Estero Bay Aquatic Preserve



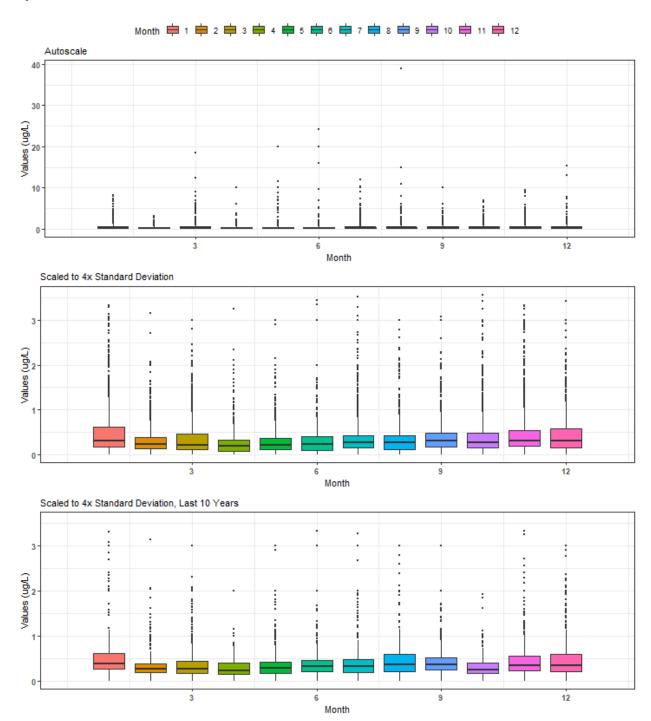
Summary Box Plots for Florida Keys National Marine Sanctuary



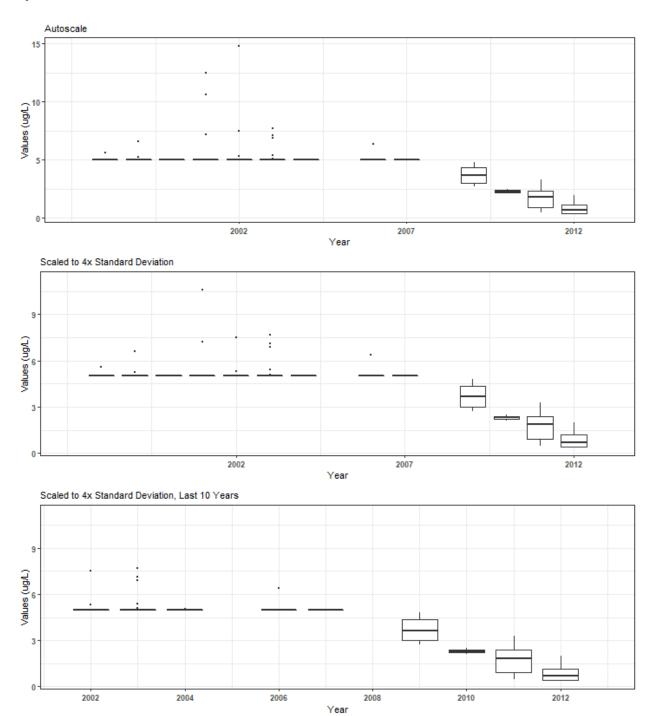
Summary Box Plots for Florida Keys National Marine Sanctuary



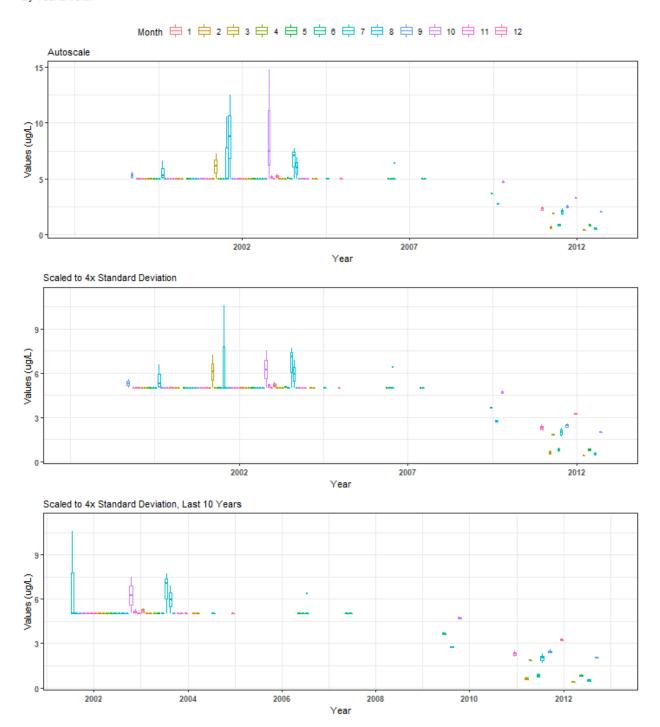
Summary Box Plots for Florida Keys National Marine Sanctuary



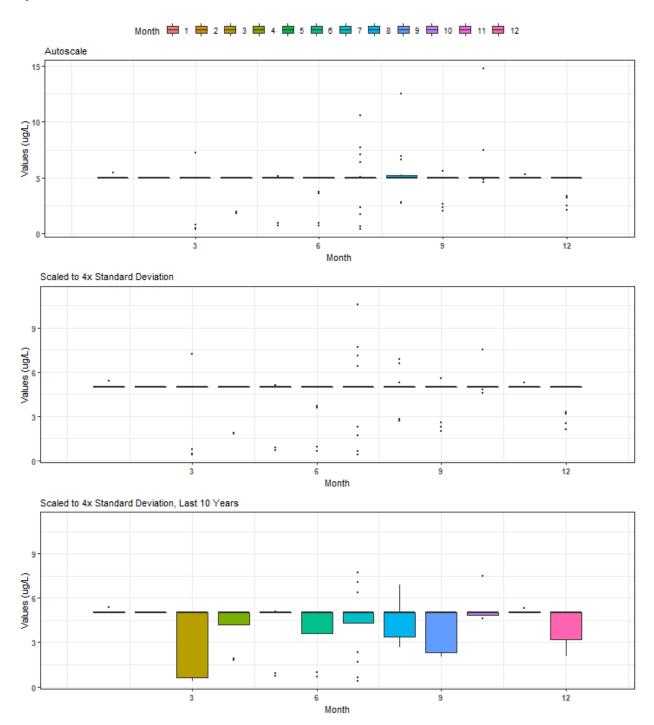
Summary Box Plots for Fort Pickens State Park Aquatic Preserve



Summary Box Plots for Fort Pickens State Park Aquatic Preserve

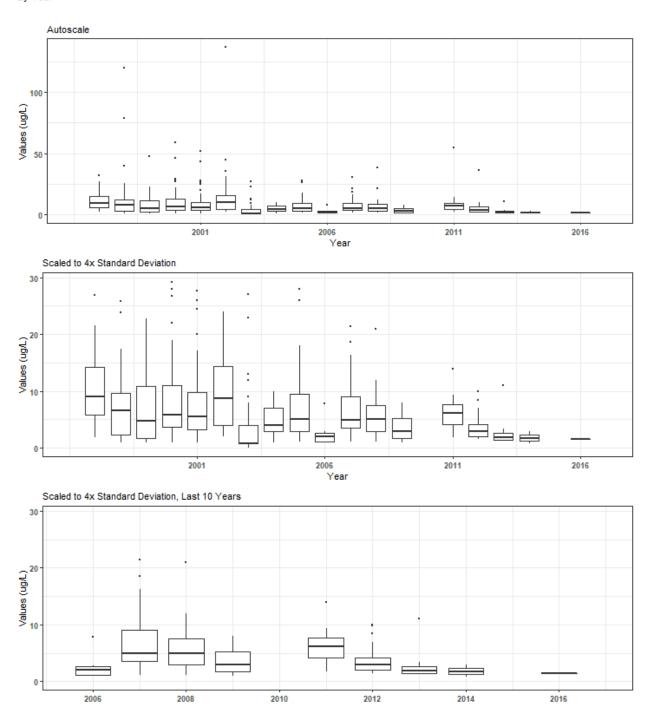


Summary Box Plots for Fort Pickens State Park Aquatic Preserve



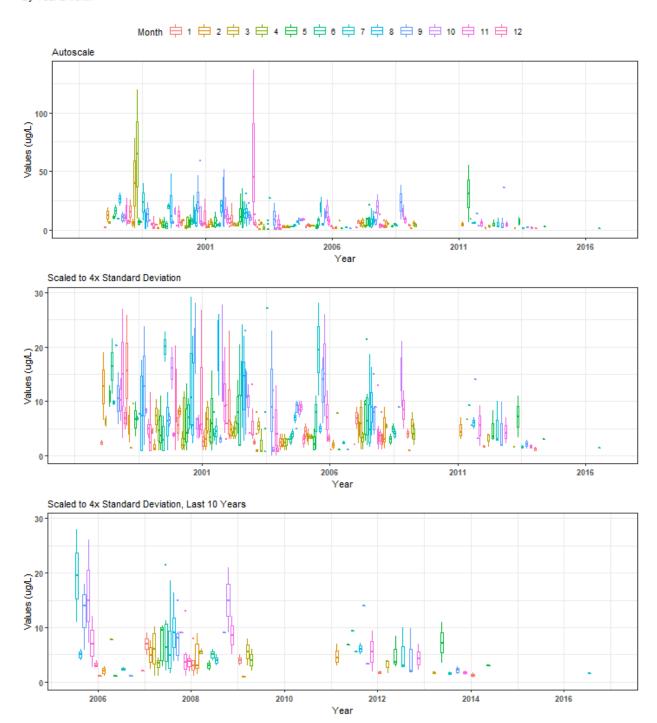
Summary Box Plots for Gasparilla Sound-Charlotte Harbor Aquatic Preserve

By Year

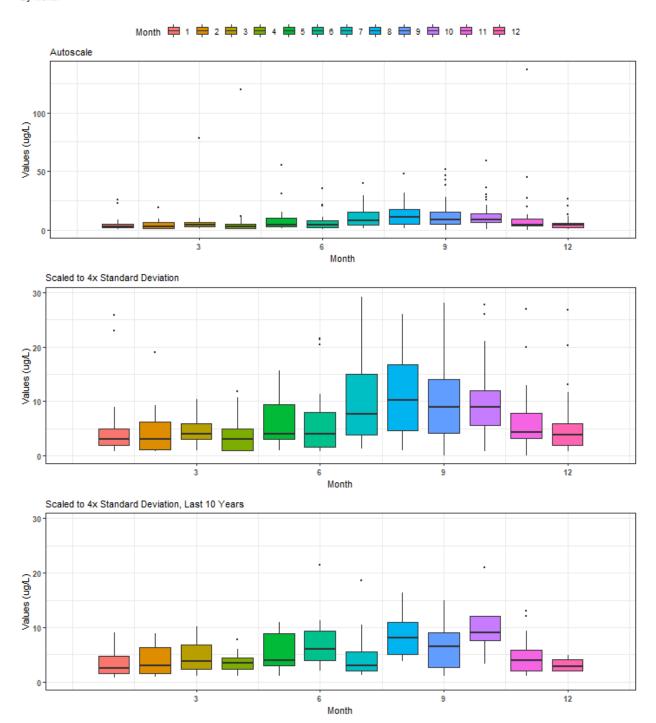


Year

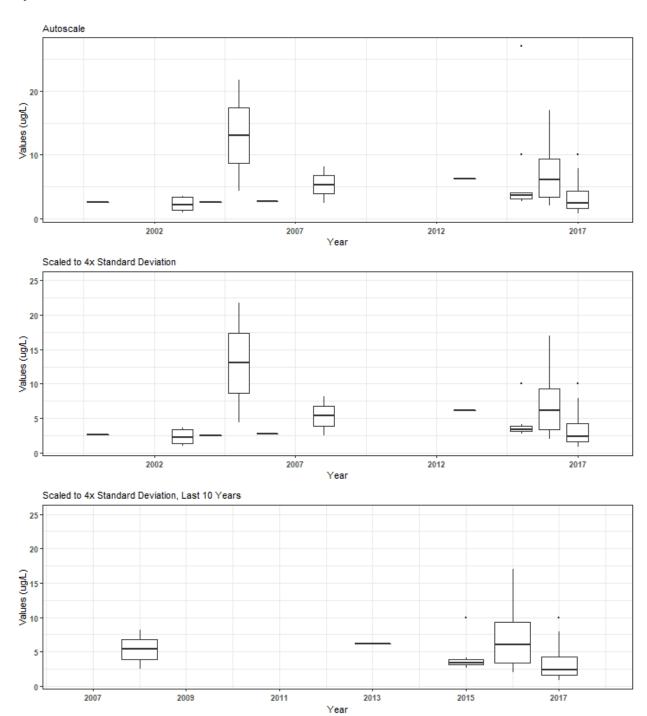
Summary Box Plots for Gasparilla Sound-Charlotte Harbor Aquatic Preserve



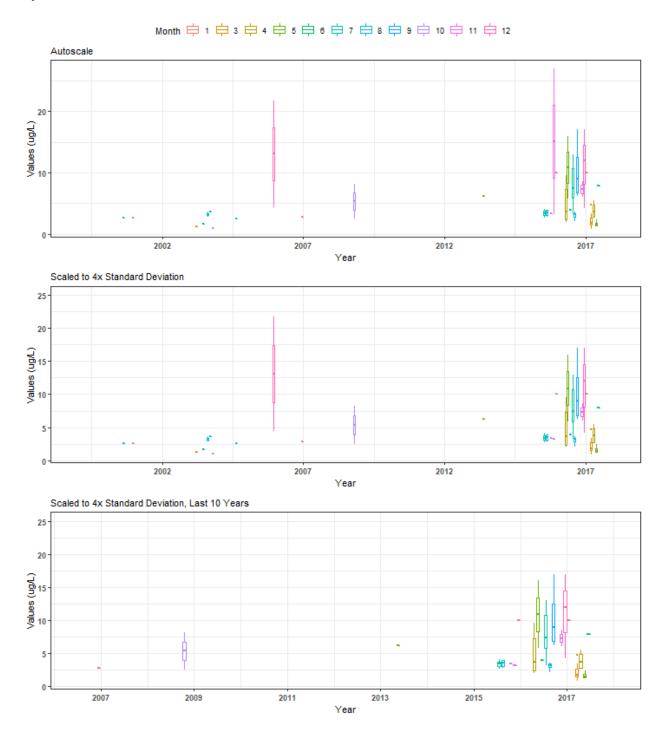
Summary Box Plots for Gasparilla Sound-Charlotte Harbor Aquatic Preserve



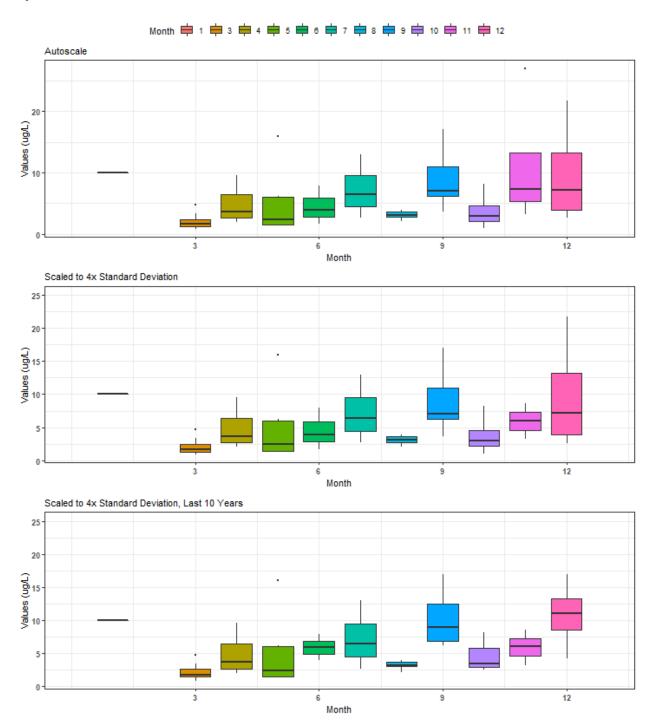
Summary Box Plots for Indian River-Malabar to Vero Beach Aquatic Preserve



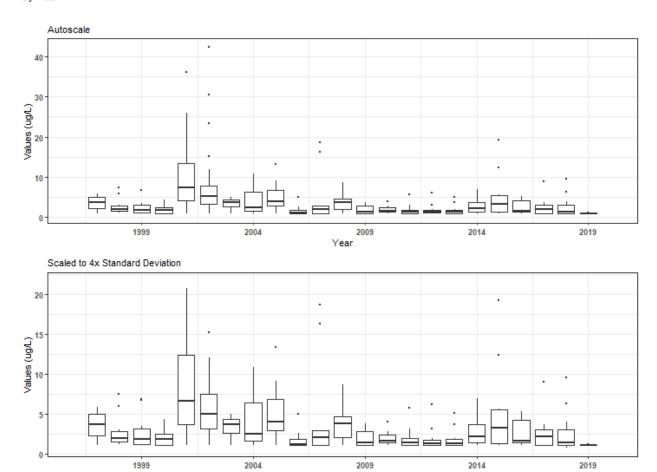
Summary Box Plots for Indian River-Malabar to Vero Beach Aquatic Preserve

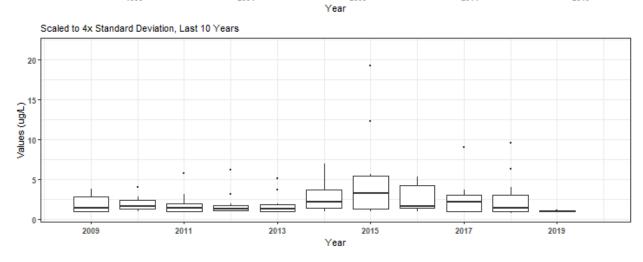


Summary Box Plots for Indian River-Malabar to Vero Beach Aquatic Preserve

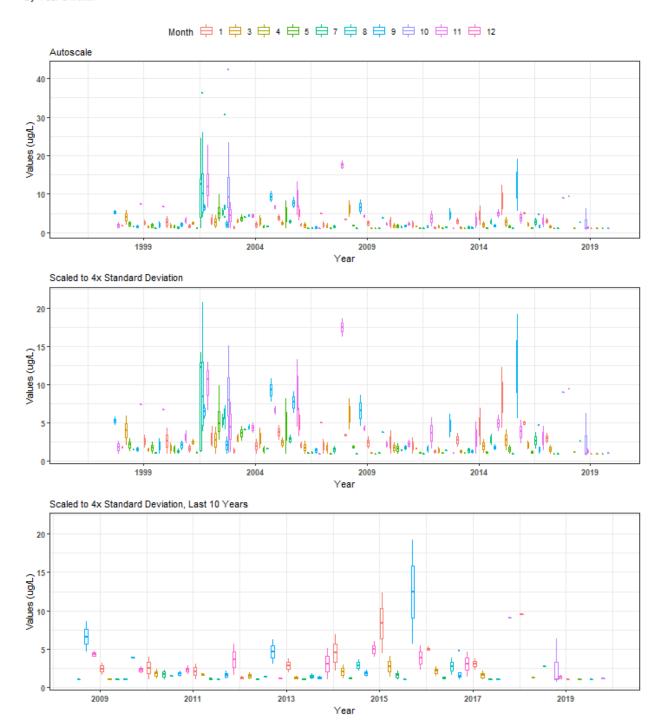


Summary Box Plots for Jensen Beach to Jupiter Inlet Aquatic Preserve

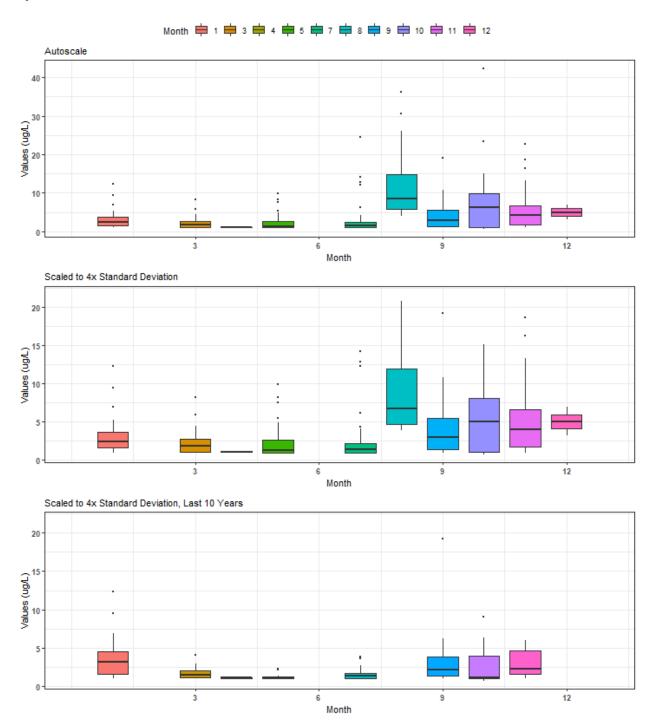




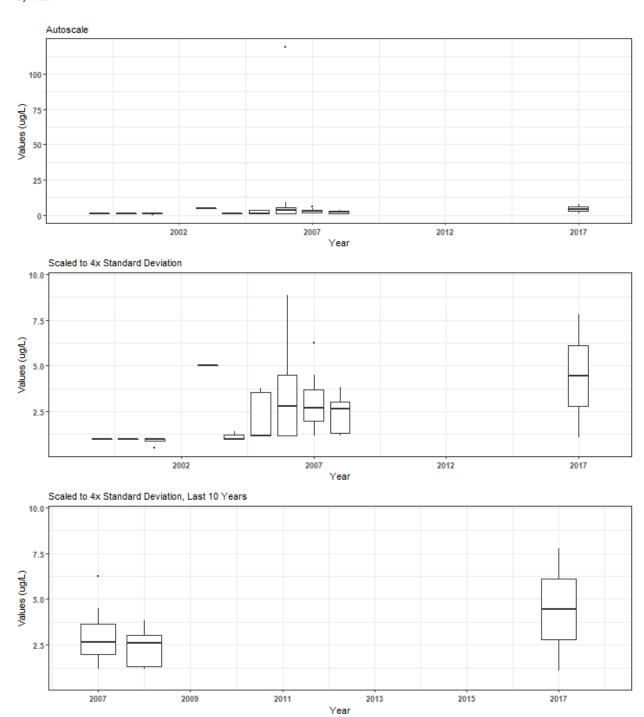
Summary Box Plots for Jensen Beach to Jupiter Inlet Aquatic Preserve



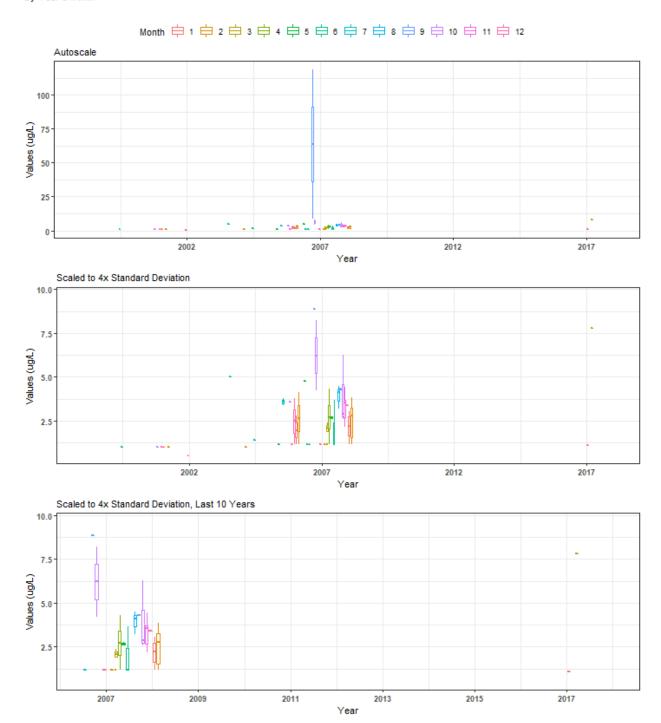
Summary Box Plots for Jensen Beach to Jupiter Inlet Aquatic Preserve



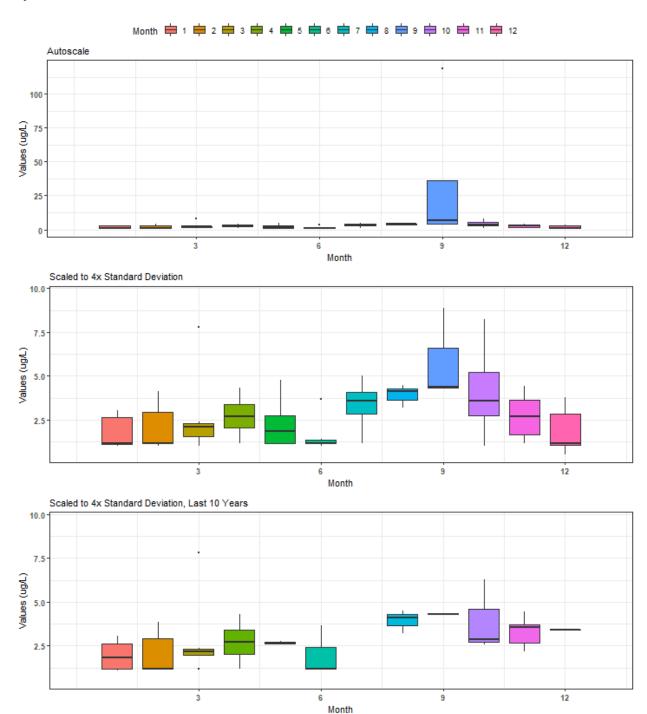
Summary Box Plots for Lemon Bay Aquatic Preserve



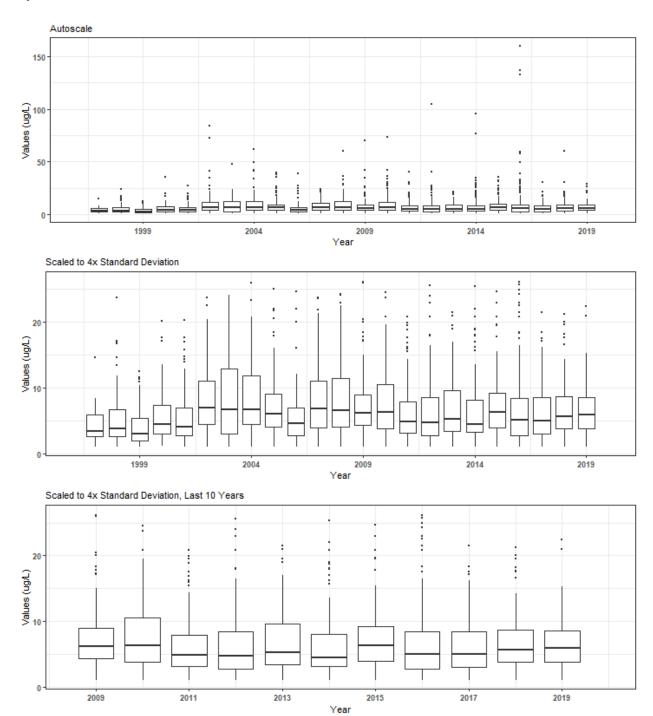
Summary Box Plots for Lemon Bay Aquatic Preserve



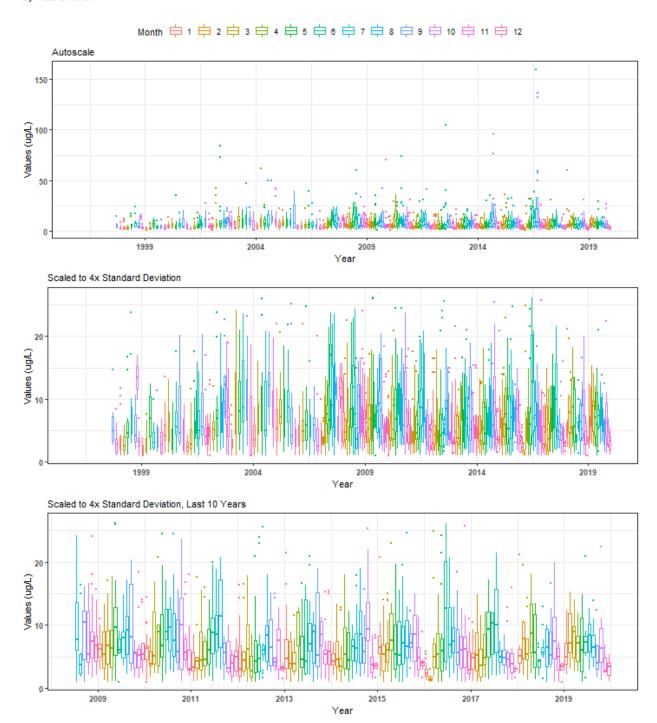
Summary Box Plots for Lemon Bay Aquatic Preserve



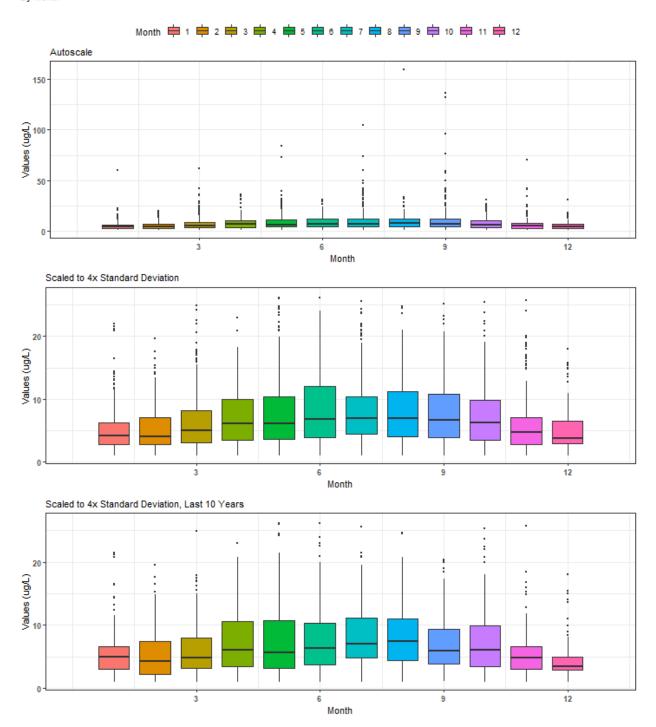
Summary Box Plots for Loxahatchee River-Lake Worth Creek Aquatic Preserve



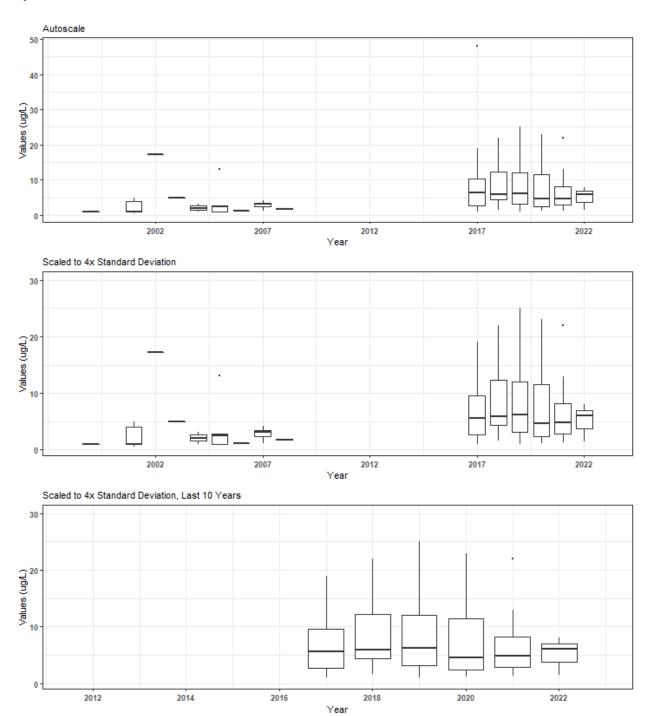
Summary Box Plots for Loxahatchee River-Lake Worth Creek Aquatic Preserve



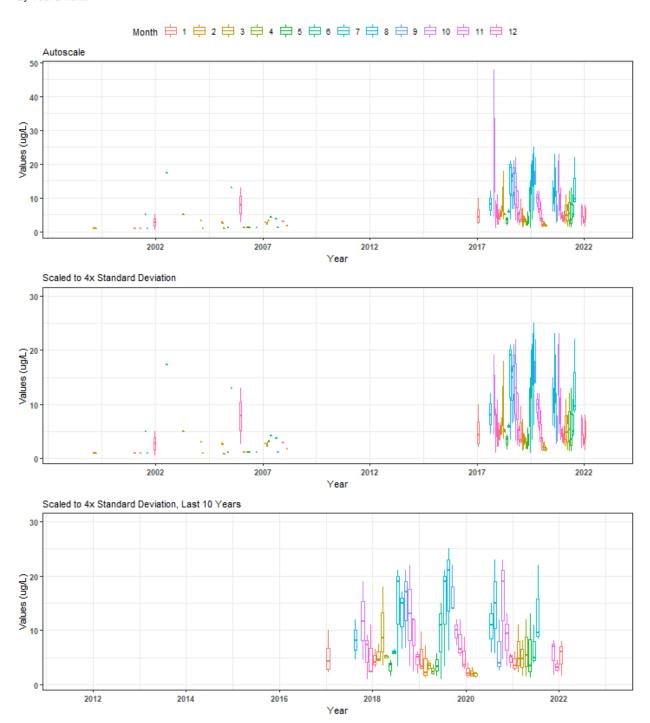
Summary Box Plots for Loxahatchee River-Lake Worth Creek Aquatic Preserve



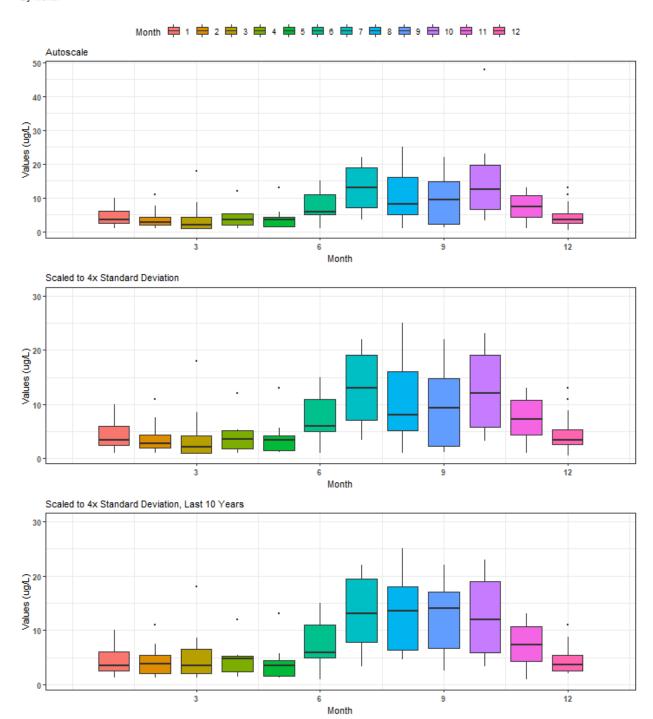
Summary Box Plots for Matlacha Pass Aquatic Preserve



Summary Box Plots for Matlacha Pass Aquatic Preserve

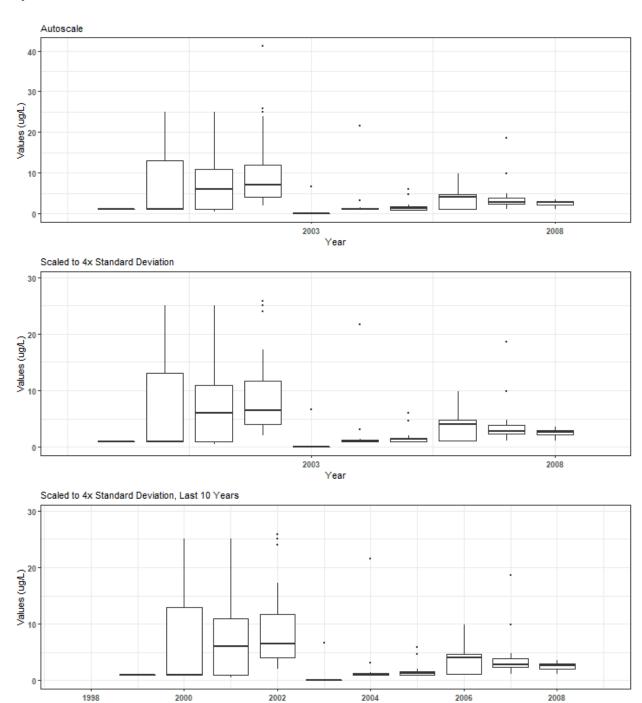


Summary Box Plots for Matlacha Pass Aquatic Preserve



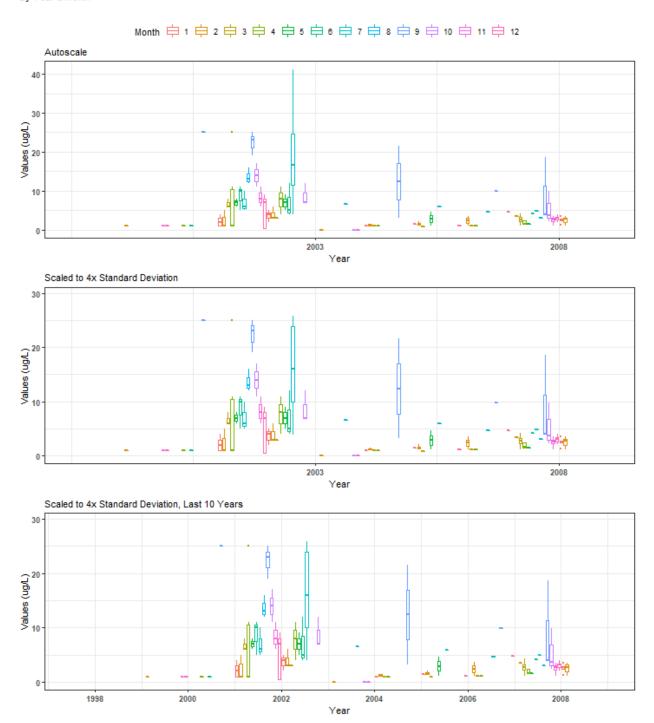
Summary Box Plots for Pine Island Sound Aquatic Preserve

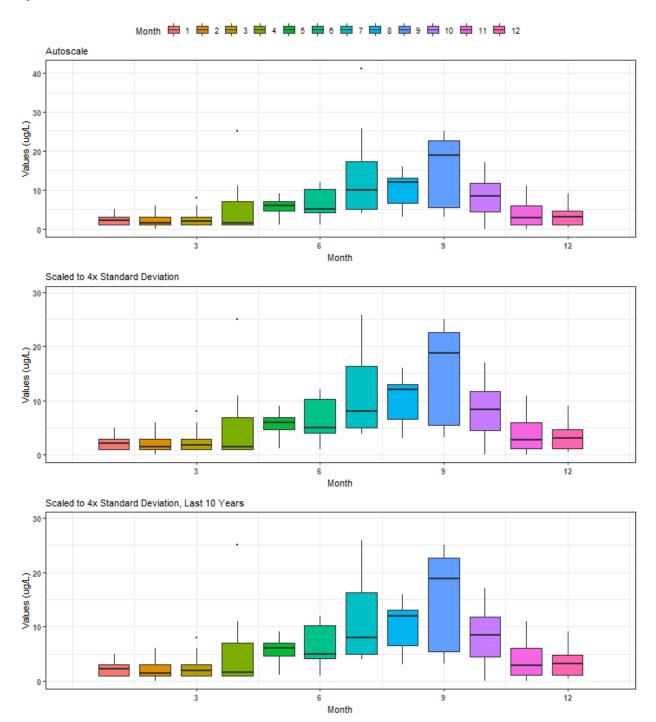
By Year



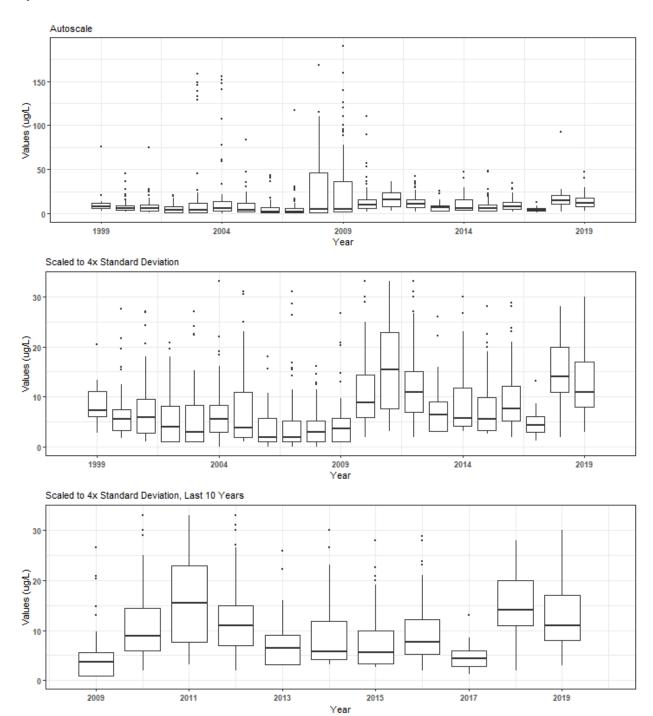
Year

Summary Box Plots for Pine Island Sound Aquatic Preserve

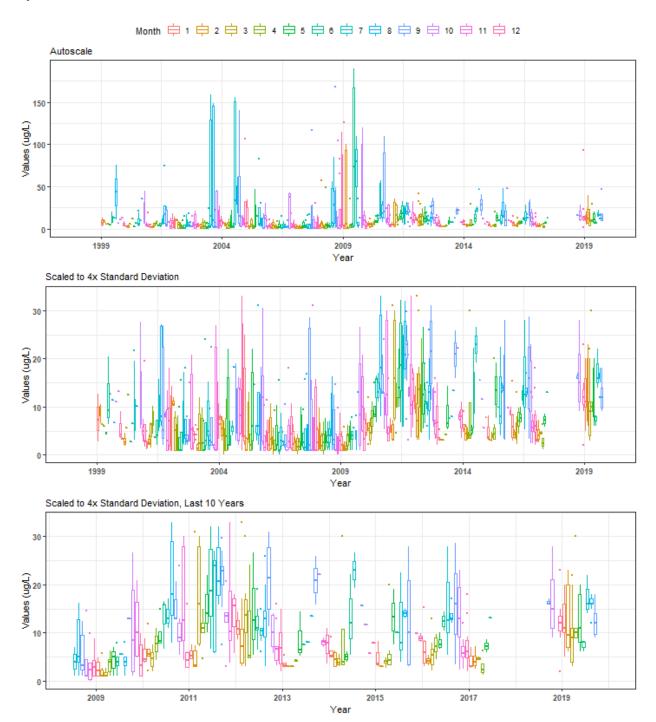


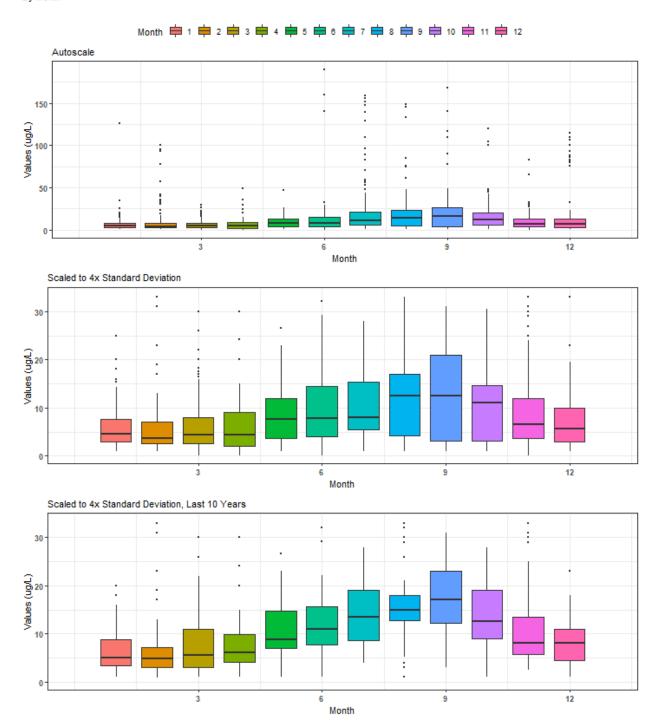


Summary Box Plots for Pinellas County Aquatic Preserve



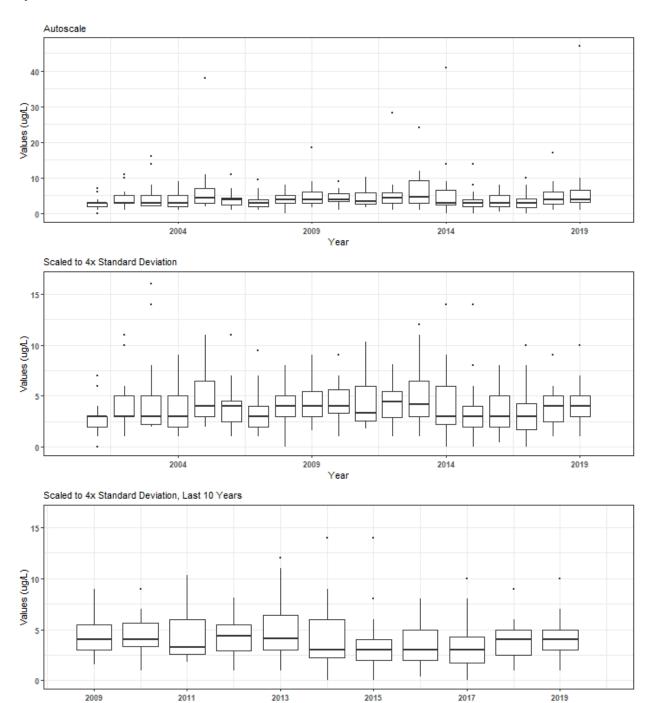
Summary Box Plots for Pinellas County Aquatic Preserve





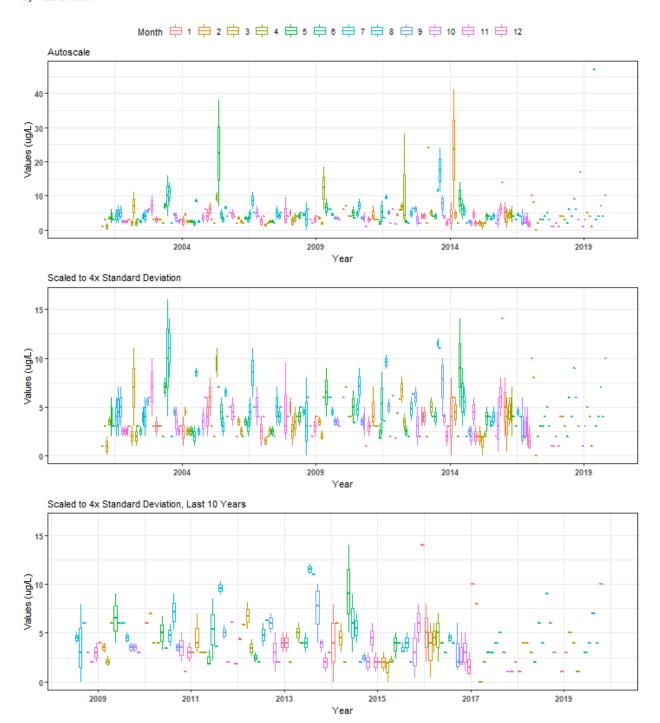
Summary Box Plots for Rocky Bayou State Park Aquatic Preserve

By Year

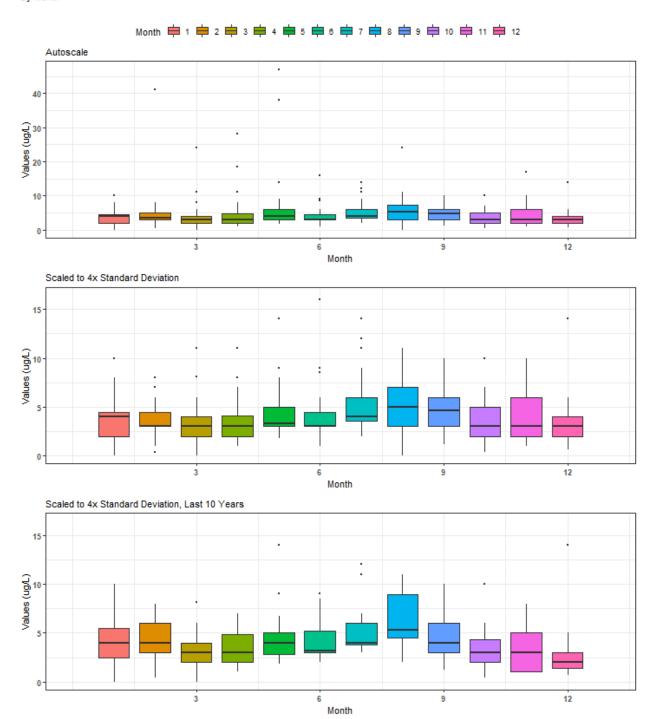


Year

Summary Box Plots for Rocky Bayou State Park Aquatic Preserve

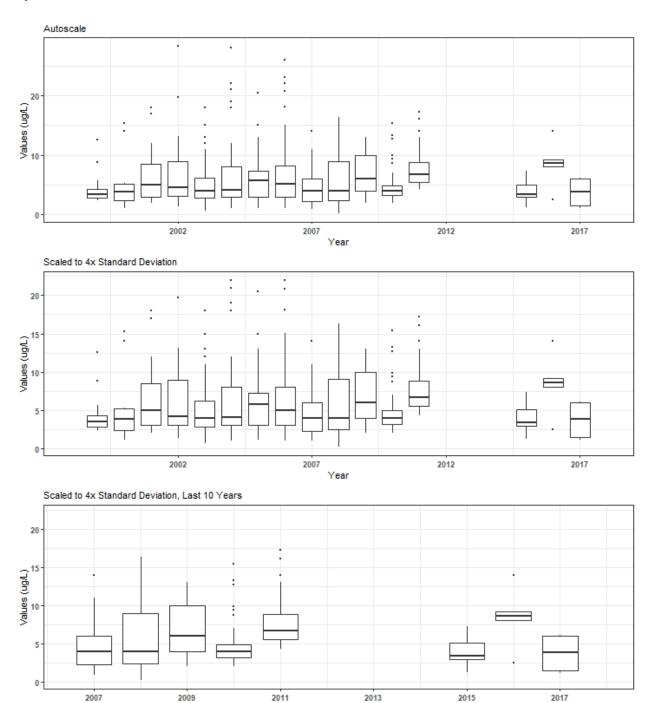


Summary Box Plots for Rocky Bayou State Park Aquatic Preserve

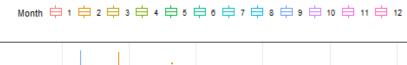


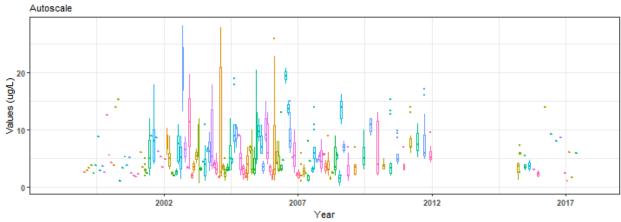
Summary Box Plots for Rookery Bay Aquatic Preserve

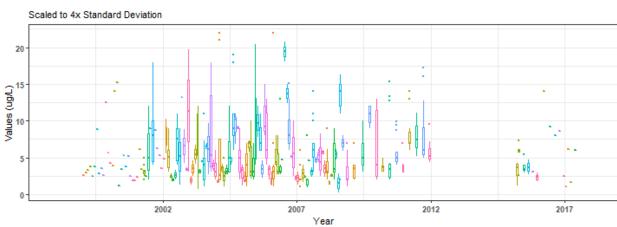
By Year

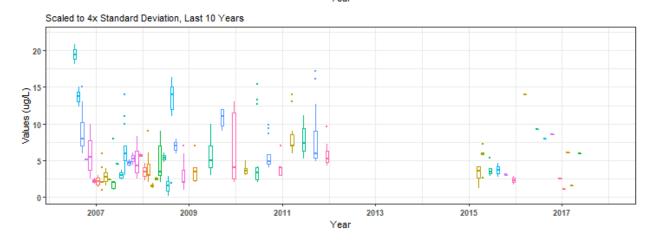


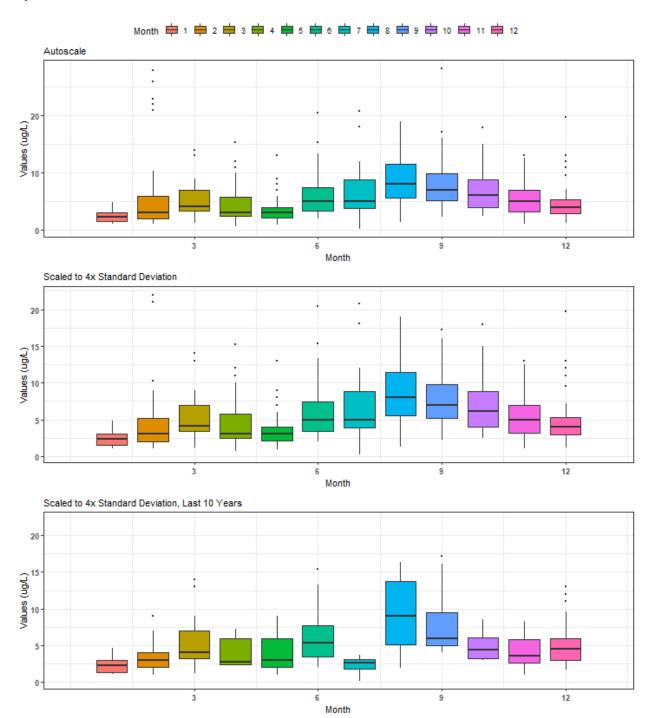
Year





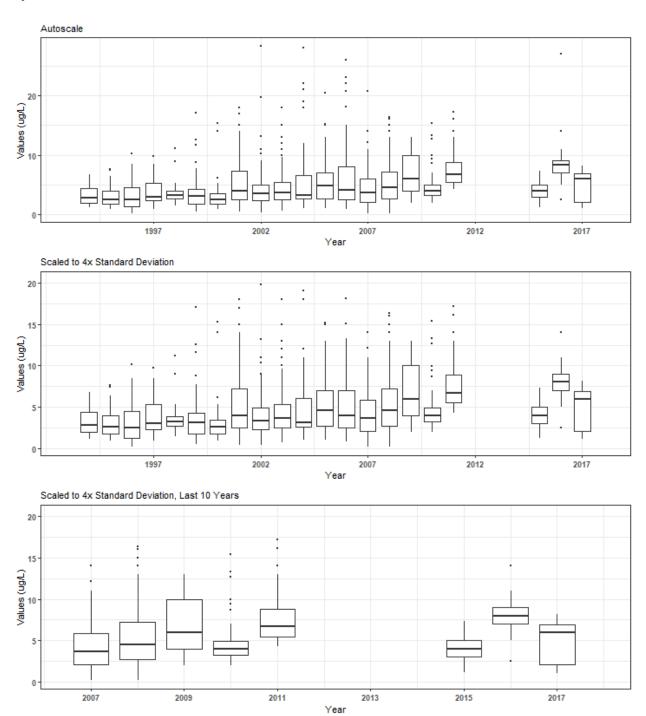




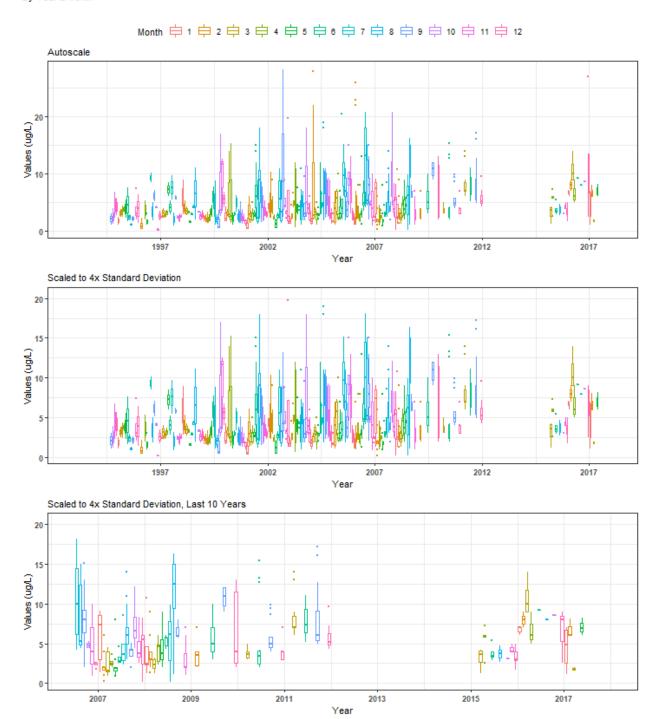


Summary Box Plots for Rookery Bay National Estuarine Research Reserve

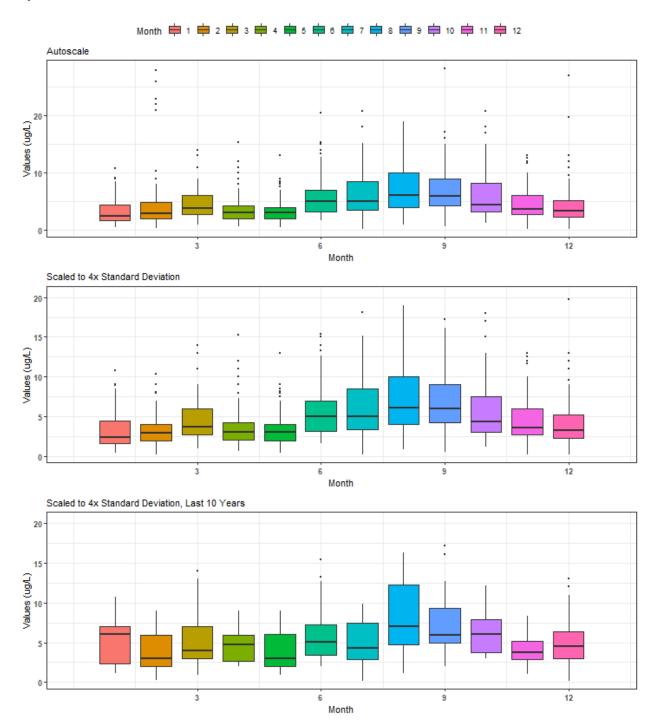
By Year



Summary Box Plots for Rookery Bay National Estuarine Research Reserve



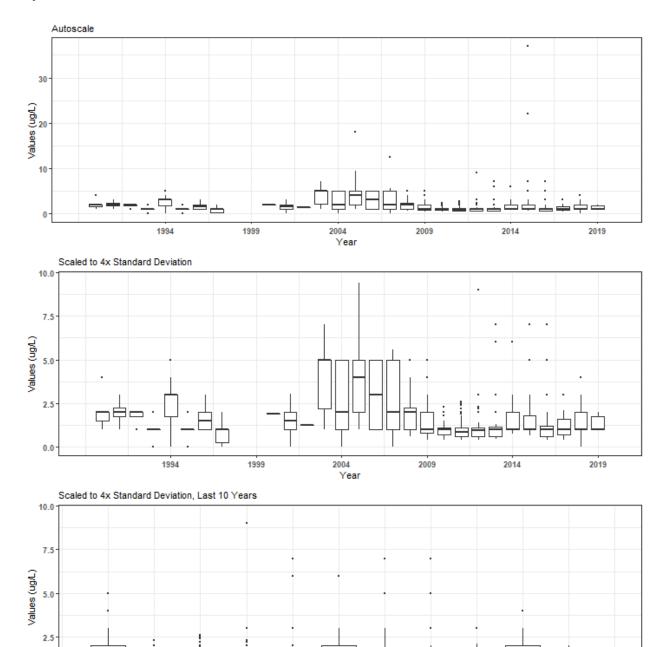
Summary Box Plots for Rookery Bay National Estuarine Research Reserve



Summary Box Plots for St. Andrews State Park Aquatic Preserve

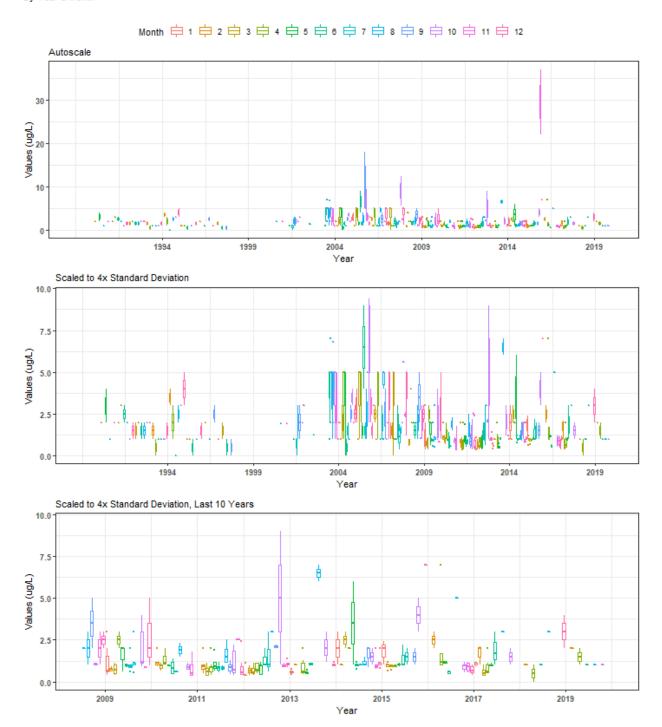
By Year

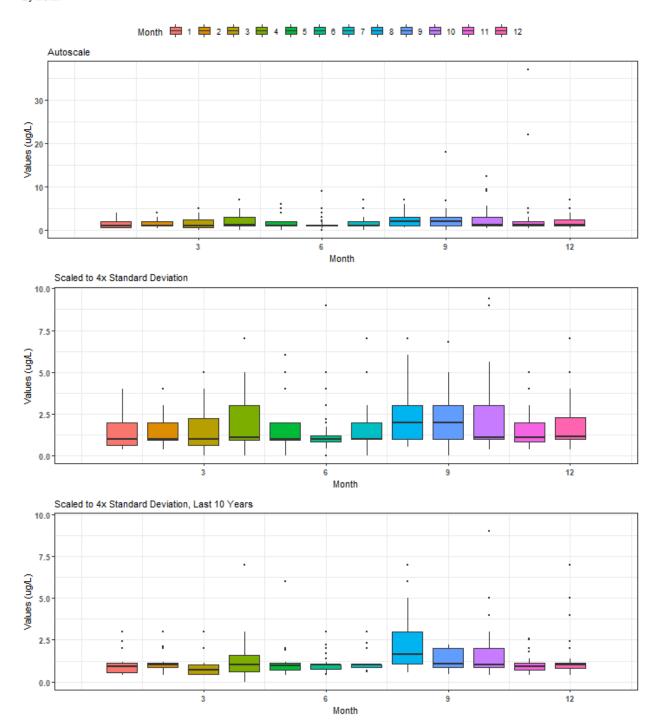
0.0



Year

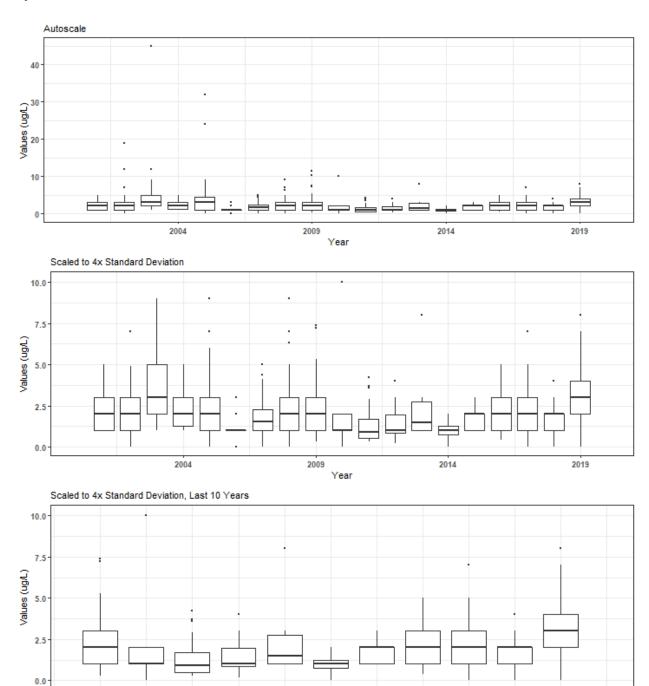
Summary Box Plots for St. Andrews State Park Aquatic Preserve





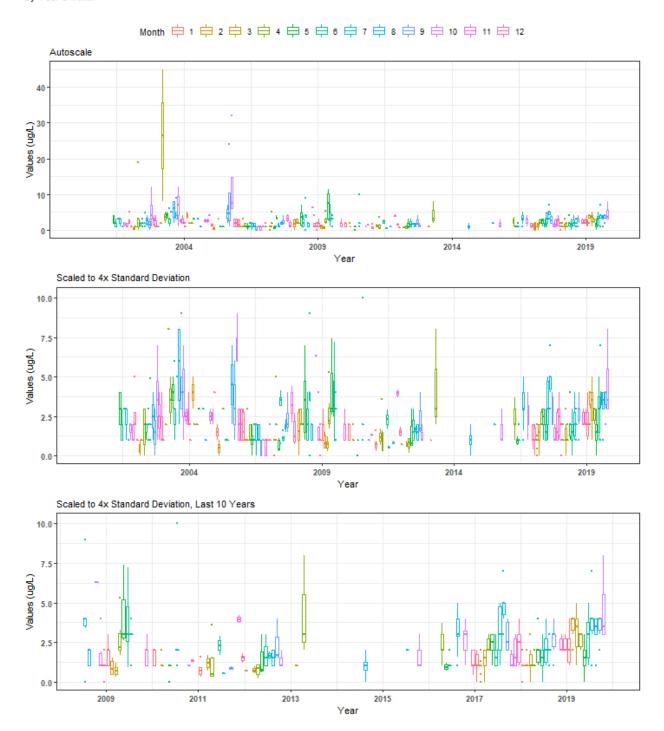
Summary Box Plots for St. Joseph Bay Aquatic Preserve

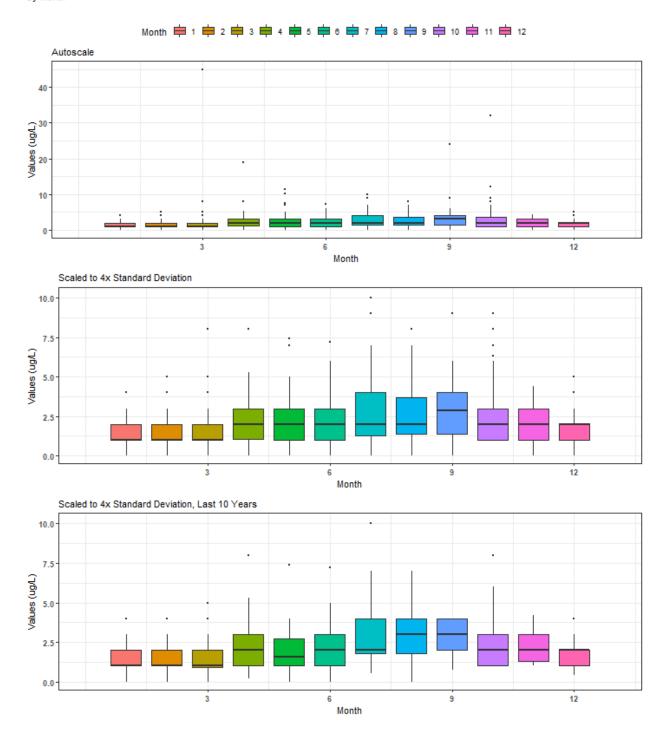
By Year



Year

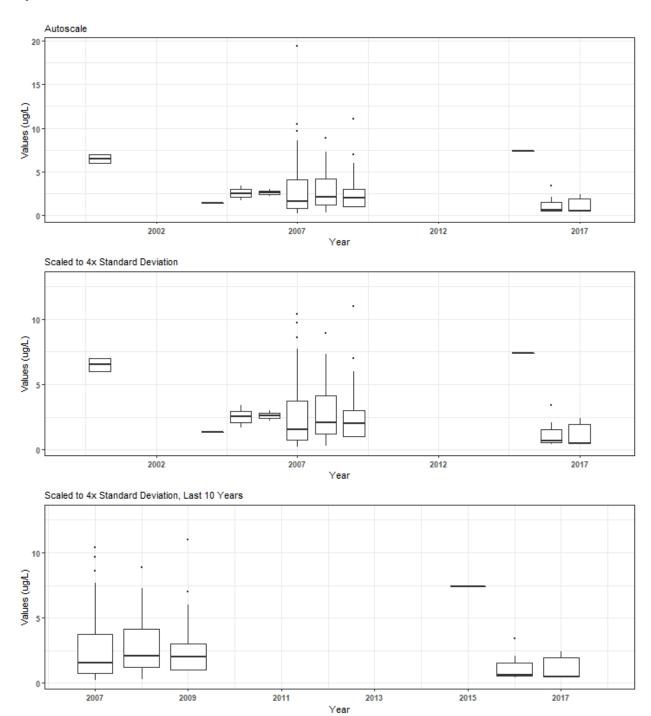
Summary Box Plots for St. Joseph Bay Aquatic Preserve



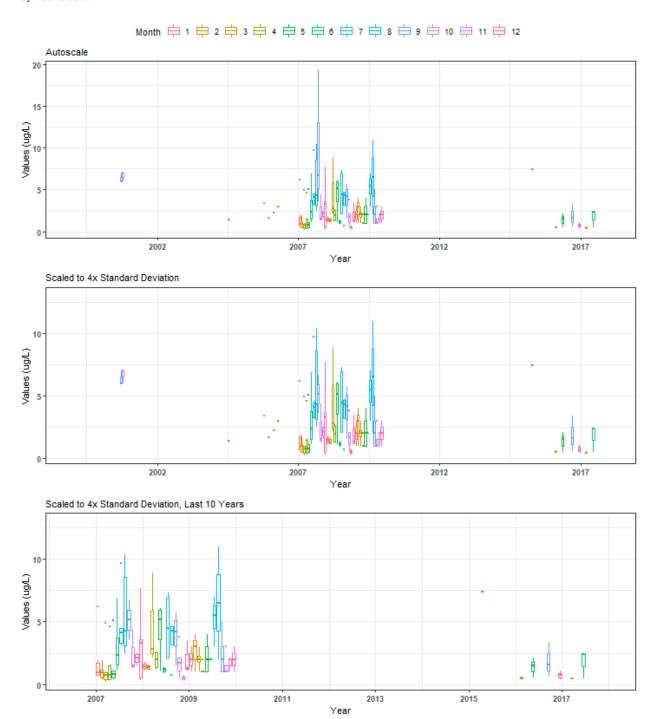


Summary Box Plots for St. Martins Marsh Aquatic Preserve

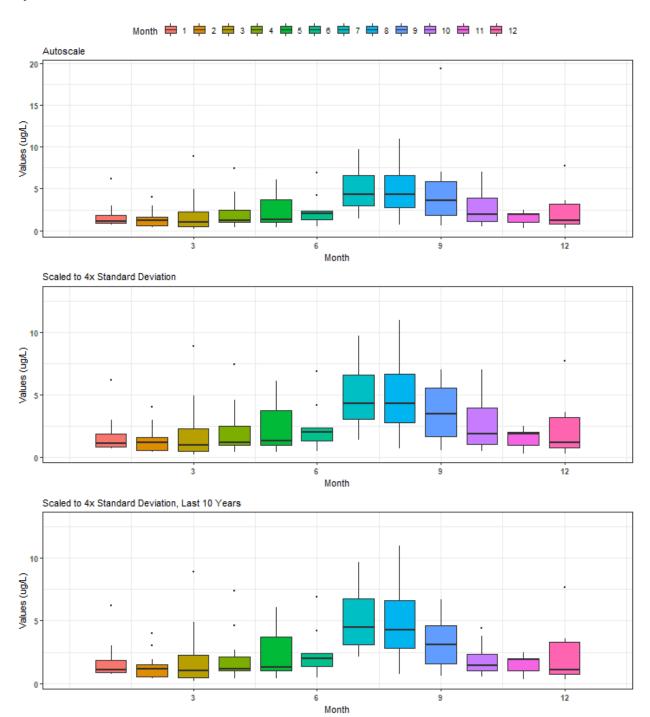
By Year



Summary Box Plots for St. Martins Marsh Aquatic Preserve

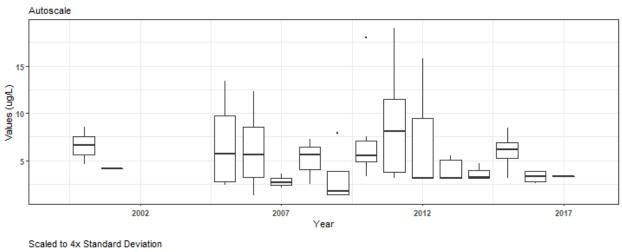


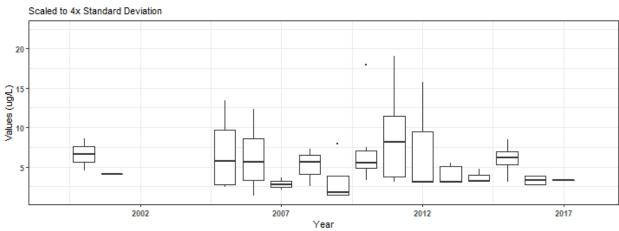
Summary Box Plots for St. Martins Marsh Aquatic Preserve

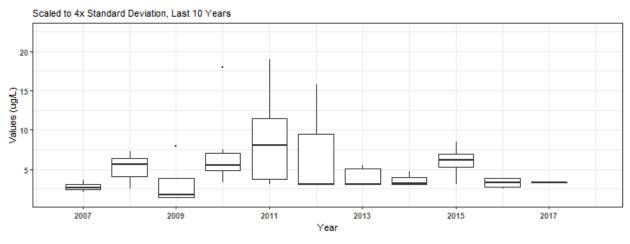


Summary Box Plots for Terra Ceia Aquatic Preserve

By Year







Summary Box Plots for Terra Ceia Aquatic Preserve

