# SEACAR Discrete Water Quality Analysis: Field Bottom pH

# Last compiled on 30 May, 2022

# Contents

Libraries	1
File Import	2
Data Filtering and Data Impacted by Specific Value Qualifiers	2
Managed Area Statistics	5
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	25
Appendix V: Managed Area Summary Box Plots	43

# 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)
opts_chunk$set(warning=FALSE, message=FALSE)
```

## 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 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"
}</pre>
```

```
data$Include <- as.logical(data$Include)</pre>
data$Include[grep("H", data$ValueQualifier[data$ProgramID==476])] <- TRUE
data <- data[!is.na(data$ResultValue),]</pre>
if(param_name!="Secchi_Depth"){
   data <- data[!is.na(data$RelativeDepth),]</pre>
   data <- data[data$RelativeDepth==depth,]</pre>
}
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$ResultValue>=-2,]
} else{
   data <- data[data$ResultValue>=0,]
data <- merge.data.frame(MA_All[,c("AreaID", "ManagedAreaName")],</pre>
                          data, by="ManagedAreaName", all=TRUE)
MA_Summ <- data %>%
   group_by(AreaID, ManagedAreaName) %>%
   summarize(ParameterName=parameter,
             RelativeDepth=depth,
             ActivityType=activity,
             N_Data=length(ResultValue[Include==TRUE & !is.na(ResultValue)]),
             N_Years=length(unique(Year[Include==TRUE & !is.na(Year)])),
             EarliestYear=min(Year[Include==TRUE]),
             LatestYear=max(Year[Include==TRUE]),
             SufficientData=ifelse(N_Data>0 & N_Years>=10, TRUE, FALSE))
data <- merge.data.frame(data, MA_Summ[,c("ManagedAreaName", "SufficientData")],</pre>
                          by="ManagedAreaName")
data$Use_In_Analysis <- ifelse(data$Include==TRUE & data$SufficientData==TRUE,
                                TRUE, FALSE)
MA_Summ <- MA_Summ %>%
   select(AreaID, ManagedAreaName, ParameterName, RelativeDepth, ActivityType,
          SufficientData, everything())
MA_Summ <- as.data.frame(MA_Summ[order(MA_Summ$ManagedAreaName), ])
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[na.omit(data$ProgramID!=476)] <- gsub("[^U]+", "",</pre>
                                                data$VQ_Plot[na.omit(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: 44974, Number Passed Filter: 44696
## Program 476 H Codes: 0 (0%)
## I Codes: 0 (0%)
## Q Codes: 0 (0%)
## U Codes: 0 (0%)
data summ <- data %>%
  group by (AreaID, ManagedAreaName) %>%
   summarize(ParameterName=parameter,
             RelativeDepth=depth,
             ActivityType=activity,
             N_Total=length(ResultValue),
             N_AnalysisUse=length(ResultValue[SufficientData==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$ValueQualifier))
data summ <- as.data.table(data summ[order(data summ$ManagedAreaName), ])</pre>
fwrite(data_summ, pasteO(out_dir,"/", param_name, "_", activity, "_", depth,
                          "_DataSummary.csv"), sep=",")
rm(data_summ)
MA_Include <- MA_Summ$ManagedAreaName[MA_Summ$SufficientData==TRUE &
                                          MA_Summ$N_Data<2000000]
n <- length(MA_Include)</pre>
MA_Exclude <- MA_Summ[MA_Summ$N_Years<10 & MA_Summ$N_Years>0,]
MA_Exclude <- MA_Exclude[,c("ManagedAreaName", "N_Years")]</pre>
z <- nrow(MA_Exclude)</pre>
```

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

# **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 SufficientData 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 ProgramName 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="|")
rm(Mon_Stats)
```

## 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 SufficientData 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

```
tauSeasonal <- function(data, independent, stats.median, stats.minYear,</pre>
                          stats.maxYear) {
   tau <- NULL
   tryCatch({ken <- kendallSeasonalTrendTest(</pre>
      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>
   rm(ken)
   }, warning = function(w) {
      print(w)
   }, error = function(e) {
      print(e)
   }, finally = {
      if (!exists("tau")) {
          tau <- NA
      if (!exists("p")) {
          p <- NA
```

```
if (!exists("slope")) {
         slope <- NA
      if (!exists("intercept")) {
         intercept <- NA
      if (!exists("trend")) {
         trend <- NA
   })
   KT <-c(unique(data$AreaID),</pre>
          unique(data$ManagedAreaName),
          stats.median,
          independent,
          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[9])) {
      KT <- tauSeasonal(data, FALSE, stats.median,</pre>
                          stats.minYear, stats.maxYear)
   if (is.null(KT.Stats) == TRUE) {
      KT.Stats <- KT
   } 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 {
         }
      }
```

```
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$ManagedAreaName
c_names <- c("AreaID", "ManagedAreaName", "Median", "Independent",</pre>
              "tau", "p", "SennSlope", "SennIntercept", "Trend")
if(n==0){
   KT.Stats <- data.frame(matrix(ncol=length(c_names),</pre>
                                   nrow=length(MA_Summ$ManagedAreaName)))
   colnames(KT.Stats) <- c_names</pre>
   KT.Stats[, c("AreaID", "ManagedAreaName")] <-</pre>
      MA_Summ[, c("AreaID", "ManagedAreaName")]
} else{
   for (i in 1:n) {
      x <- nrow(data[data$Use_In_Analysis == TRUE &
                          data$ManagedAreaName == MA Include[i], ])
      if (x>0) {
         KT.Stats <- runStats(data[data$Use_In_Analysis == TRUE &</pre>
                                         data$ManagedAreaName ==
                                         MA_Include[i], ])
      }
   }
   KT.Stats <- as.data.frame(KT.Stats)</pre>
   c_names <- c("AreaID", "ManagedAreaName", "Median", "Independent",</pre>
                 "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$tau <- round(as.numeric(KT.Stats$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>
}
KT.Stats <- merge.data.frame(MA_Summ, KT.Stats,</pre>
                                by=c("AreaID", "ManagedAreaName"), all=TRUE)
KT.Stats <- as.data.table(KT.Stats[order(KT.Stats$ManagedAreaName), ])</pre>
```

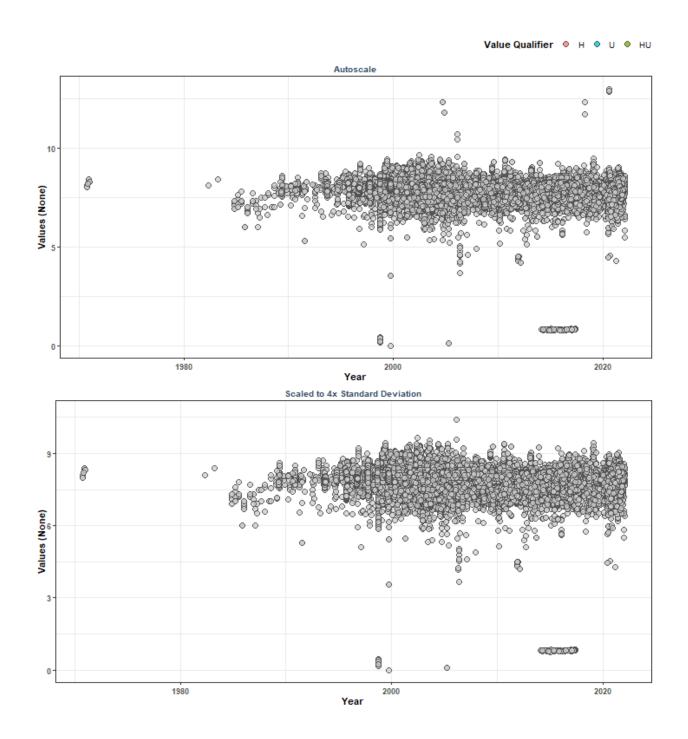
# Appendix I: Scatter Plot of Entire Dataset

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.

```
plot_theme <- theme_bw() + theme(text=element_text(family="Segoe UI"),</pre>
                                  title=element_text(face="bold"),
                                  plot.title=element_text(hjust=0.5, size=14, color="#314963"),
                                  plot.subtitle=element_text(hjust=0.5, size=10, color="#314963"),
                                  axis.text.x=element_text(face="bold"),
                                  axis.text.y=element text(face="bold"))
year_lower <- min(data$Year)</pre>
year_upper <- max(data$Year)</pre>
min_RV <- min(data$ResultValue)</pre>
mn RV <- mean(data$ResultValue[data$ResultValue <
                                   quantile(data$ResultValue, 0.98)])
sd_RV <- sd(data$ResultValue[data$ResultValue <</pre>
                                 quantile(data$ResultValue, 0.98)])
x_scale <- ifelse(year_upper - year_lower > 30, 10, 5)
y_scale <- mn_RV + 4 * sd_RV</pre>
p1 <- ggplot(data=data[data$Include==TRUE,],
             aes(x=SampleDate, y=ResultValue, fill=VQ_Plot)) +
   geom_point(shape=21, size=3, color="#333333", alpha=0.75) +
   labs(subtitle="Autoscale",
        x="Year", y=paste0("Values (", unit, ")"),
        fill="Value Qualifier") +
   plot theme +
   theme(legend.position="top", legend.box="horizontal",
         legend.justification="right") +
   scale_x_date(labels=date_format("%Y")) +
   {if(inc_H==TRUE){
      scale_fill_manual(values=c("H"= "#F8766D", "U"= "#00BFC4",
                                  "HU"="#7CAE00"), na.value="#cccccc")
   } else if(param_name=="Secchi_Depth"){
      scale_fill_manual(values=c("S"= "#F8766D", "U"= "#00BFC4",
                                  "SU"="#7CAE00"), na.value="#cccccc")
   } else {
      scale_fill_manual(values=c("U"= "#00BFC4"), na.value="#cccccc")
   }}
p2 <- ggplot(data=data[data$Include==TRUE,],
             aes(x=SampleDate, y=ResultValue, fill=VQ_Plot)) +
```

```
geom_point(shape=21, size=3, color="#333333", alpha=0.75) +
   ylim(min_RV, y_scale) +
   labs(subtitle="Scaled to 4x Standard Deviation",
        x="Year", y=paste0("Values (", unit, ")")) +
   plot_theme +
   theme(legend.position="none") +
   scale_x_date(labels=date_format("%Y")) +
   {if(inc H==TRUE){
      scale_fill_manual(values=c("H"= "#F8766D", "U"= "#00BFC4",
                                  "HU"="#7CAE00"), na.value="#cccccc")
   } else if(param_name=="Secchi_Depth"){
      scale_fill_manual(values=c("S"= "#F8766D", "U"= "#00BFC4",
                                  "SU"="#7CAE00"), na.value="#cccccc")
   } else {
      scale_fill_manual(values=c("U"= "#00BFC4"), na.value="#cccccc")
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>
   plot_theme + theme(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))
```

#### **Scatter Plot for Entire Dataset**



# 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 SufficientData 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$Include==TRUE, ],</pre>
             aes(x=Year, y=ResultValue, group=Year)) +
   geom_boxplot(color="#333333", fill="#cccccc", outlier.shape=21,
                outlier.size=3, outlier.color="#333333",
                outlier.fill="#cccccc", outlier.alpha=0.75) +
   labs(subtitle="Autoscale", x="Year",
        y=paste0("Values (", unit, ")")) +
   plot_theme
p2 <- ggplot(data=data[data$Include==TRUE, ],</pre>
             aes(x=Year, y=ResultValue, group=Year)) +
   geom_boxplot(color="#333333", fill="#cccccc", outlier.shape=21,
                outlier.size=3, outlier.color="#333333",
                outlier.fill="#cccccc", outlier.alpha=0.75) +
   labs(subtitle="Scaled to 4x Standard Deviation", x="Year",
        y=paste0("Values (", unit, ")")) +
   ylim(min_RV, y_scale) +
   plot_theme
p3 <- ggplot(data=data[data$Include==TRUE, ],
             aes(x=as.integer(Year), y=ResultValue, group=Year)) +
   geom_boxplot(color="#333333", fill="#cccccc", outlier.shape=21,
                outlier.size=3, outlier.color="#333333",
                outlier.fill="#cccccc", outlier.alpha=0.75) +
   labs(subtitle="Scaled to 4x Standard Deviation, Last 10 Years",
        x="Year", y=paste0("Values (", unit, ")")) +
   ylim(min_RV, y_scale) +
   scale_x_continuous(limits=c(max(data$Year) - 10.5, max(data$Year)+1),
                      breaks=seq(max(data$Year) - 10, max(data$Year), 2)) +
   plot_theme
set <- ggarrange(p1, p2, p3, ncol=1)</pre>
p0 <- ggplot() + labs(title="Summary Box Plots for Entire Data",
```

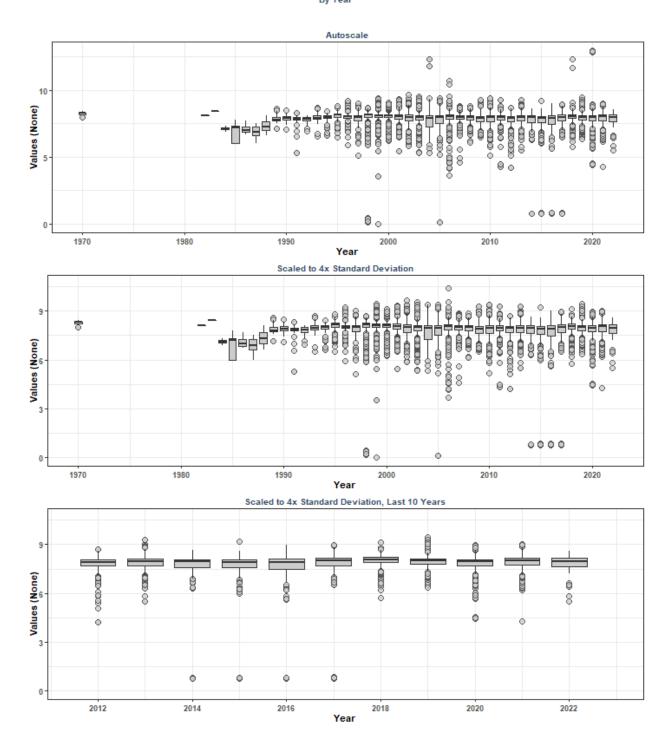
This set of box plots are grouped by year and month with the color being related to the month.

```
p1 <- ggplot(data=data[data$Include==TRUE, ],
             aes(x=YearMonthDec, y=ResultValue,
                 group=YearMonth, color=as.factor(Month))) +
   geom_boxplot(fill="#cccccc", outlier.size=1.5, outlier.alpha=0.75) +
   labs(subtitle="Autoscale", x="Year",
        y=paste0("Values (", unit, ")"), color="Month") +
   plot_theme +
   theme(legend.position="top", legend.box="horizontal") +
   guides(color=guide_legend(nrow=1))
p2 <- ggplot(data=data[data$Include==TRUE, ],</pre>
             aes(x=YearMonthDec, y=ResultValue,
                 group=YearMonth, color=as.factor(Month))) +
   geom_boxplot(fill="#cccccc", outlier.size=1.5, outlier.alpha=0.75) +
   labs(subtitle="Scaled to 4x Standard Deviation",
        x="Year", y=paste0("Values (", unit, ")")) +
  ylim(min_RV, y_scale) +
  plot theme +
   theme(legend.position="none", axis.text.x=element_text(face="bold"),
         axis.text.y=element_text(face="bold"))
p3 <- ggplot(data=data[data$Include==TRUE, ],
             aes(x=YearMonthDec, y=ResultValue,
                 group=YearMonth, color=as.factor(Month))) +
   geom_boxplot(fill="#cccccc", outlier.size=1.5, outlier.alpha=0.75) +
   labs(subtitle="Scaled to 4x Standard Deviation, Last 10 Years",
        x="Year", y=paste0("Values (", unit, ")")) +
   ylim(min_RV, y_scale) +
   scale_x_continuous(limits=c(max(data$Year) - 10.5, max(data$Year)+1),
                      breaks=seq(max(data$Year) - 10, max(data$Year), 2)) +
  plot theme +
  theme(legend.position="none")
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",</pre>
                      subtitle="By Year & Month") + plot_theme +
   theme(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))
```

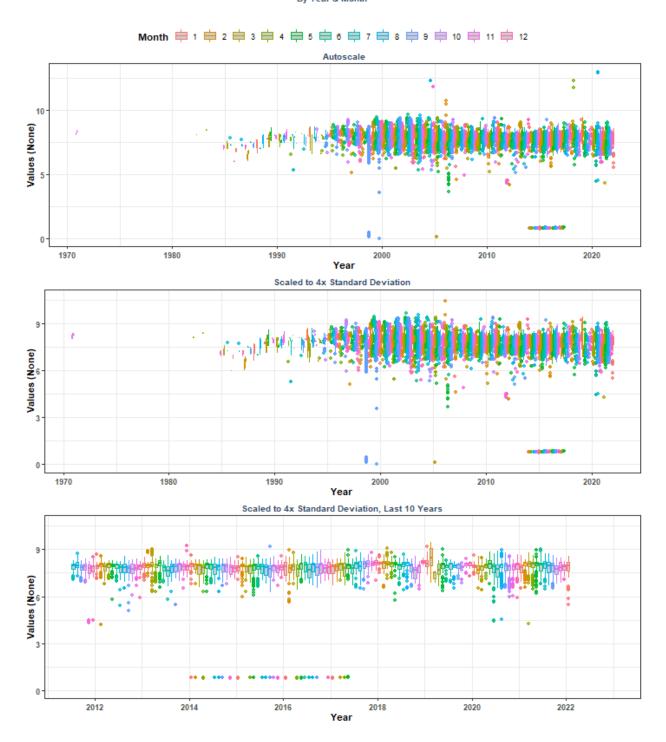
The following box plots are grouped by month with fill color being related to the month. This is designed to view potential seasonal trends.

```
p1 <- ggplot(data=data[data$Include==TRUE, ],</pre>
             aes(x=Month, y=ResultValue,
                 group=Month, fill=as.factor(Month))) +
   geom_boxplot(color="#333333", outlier.shape=21, outlier.size=3,
                outlier.color="#333333", outlier.alpha=0.75) +
   labs(subtitle="Autoscale", x="Month",
        y=paste0("Values (", unit, ")"), fill="Month") +
   scale x continuous(limits=c(0, 13), breaks=seq(3, 12, 3)) +
   plot theme +
   theme(legend.position="top", legend.box="horizontal") +
   guides(fill=guide_legend(nrow=1))
p2 <- ggplot(data=data[data$Include==TRUE, ],</pre>
             aes(x=Month, y=ResultValue,
                 group=Month, fill=as.factor(Month))) +
   geom_boxplot(color="#333333", outlier.shape=21, outlier.size=3,
                outlier.color="#333333", outlier.alpha=0.75) +
   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)) +
   plot_theme +
   theme(legend.position="none")
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(color="#333333", outlier.shape=21, outlier.size=3,
                outlier.color="#333333", outlier.alpha=0.75) +
   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)) +
  plot_theme +
  theme(legend.position="none")
leg <- get_legend(p1)</pre>
set <- ggarrange(leg, p1 + theme(legend.position="none"), p2, p3, ncol=1,
                 heights=c(0.1, 1, 1, 1))
p0 <- ggplot() + labs(title="Summary Box Plots for Entire Data",</pre>
                      subtitle="By Month") + plot_theme +
   theme(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>
```

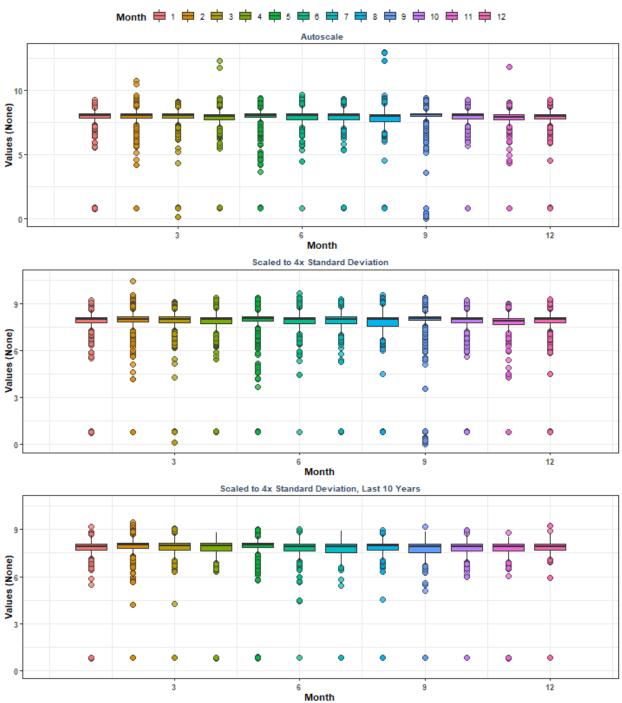
#### Summary Box Plots for Entire Data By Year



# Summary Box Plots for Entire Data By Year & Month



# Summary Box Plots for Entire Data By Month

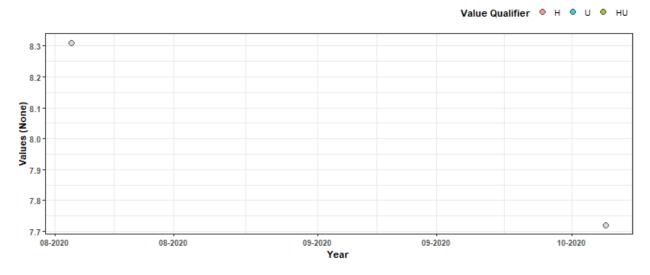


# Appendix III: Excluded Managed Areas

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.

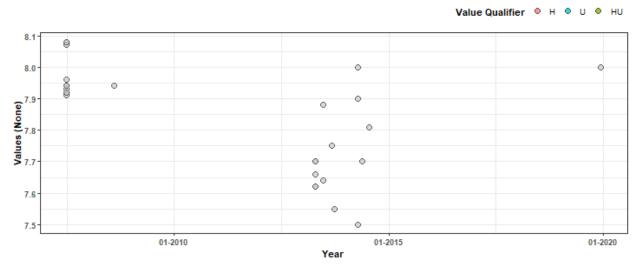
```
if(z==0){
  print("There are no managed areas that qualify.")
} else {
  for(i in 1:z){
     data$Include==TRUE, ],
                aes(x=SampleDate, y=ResultValue, fill=VQ_Plot)) +
        geom point(shape=21, size=3, color="#333333", alpha=0.75) +
        labs(title=paste0("Scatter Plot of Excluded Managed Area\n",
                         MA_Exclude$ManagedAreaName[i], " (",
                         MA_Exclude$N_Years[i], " Unique Years)"),
             subtitle="Autoscale", x="Year",
             y=pasteO("Values (", unit, ")"), fill="Value Qualifier") +
        plot_theme +
        theme(legend.position="top", legend.box="horizontal",
              legend.justification="right") +
        scale_x_date(labels=date_format("%m-%Y")) +
        {if(inc_H==TRUE){
           scale_fill_manual(values=c("H"= "#F8766D", "U"= "#00BFC4",
                                    "HU"="#7CAE00"), na.value="#cccccc")
        } else if(param_name=="Secchi_Depth"){
           scale_fill_manual(values=c("S"= "#F8766D", "U"= "#00BFC4",
                                    "SU"="#7CAE00"), na.value="#cccccc")
           scale_fill_manual(values=c("U"= "#00BFC4"), na.value="#ccccc")
        }}
     print(p1)
  }
}
```

#### Scatter Plot of Excluded Managed Area Coupon Bight Aquatic Preserve (1 Unique Years)

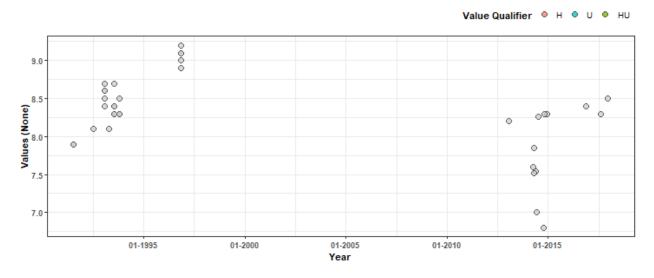


#### Scatter Plot of Excluded Managed Area Fort Clinch State Park Aquatic Preserve (5 Unique Years)

Autoscale

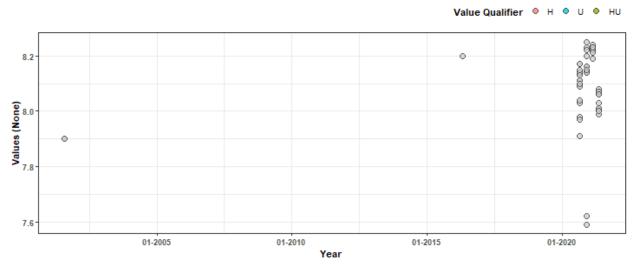


# Scatter Plot of Excluded Managed Area Fort Pickens State Park Aquatic Preserve (8 Unique Years)

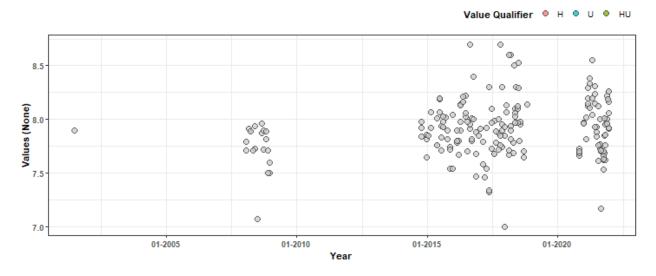


#### Scatter Plot of Excluded Managed Area Lignumvitae Key Aquatic Preserve (4 Unique Years)

Autoscale

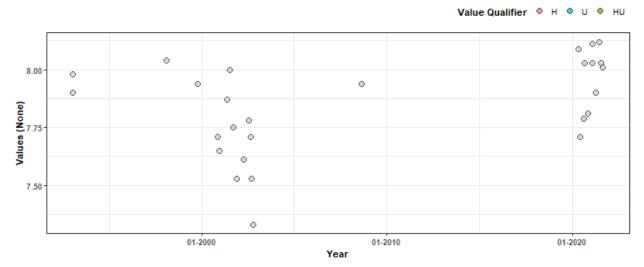


#### Scatter Plot of Excluded Managed Area Matlacha Pass Aquatic Preserve (9 Unique Years)

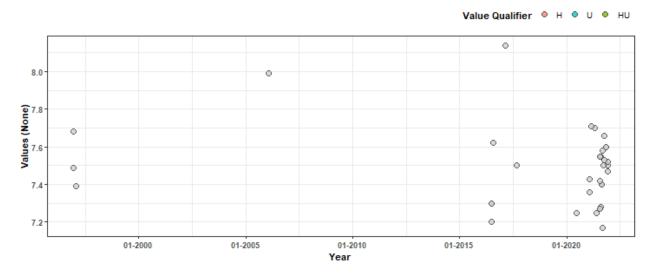


#### Scatter Plot of Excluded Managed Area Mosquito Lagoon Aquatic Preserve (9 Unique Years)

Autoscale

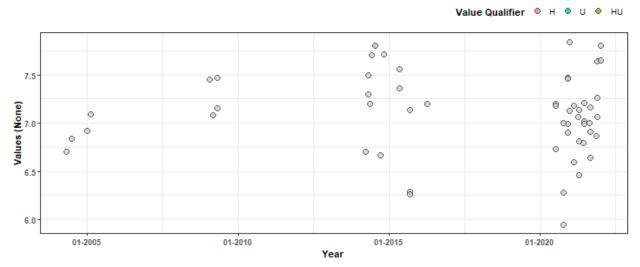


#### Scatter Plot of Excluded Managed Area North Fork St. Lucie Aquatic Preserve (7 Unique Years)

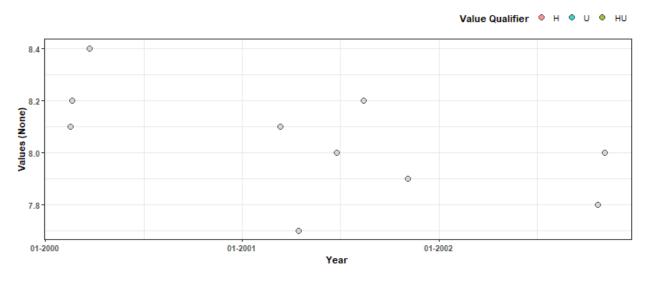


#### Scatter Plot of Excluded Managed Area Pellicer Creek Aquatic Preserve (9 Unique Years)

Autoscale

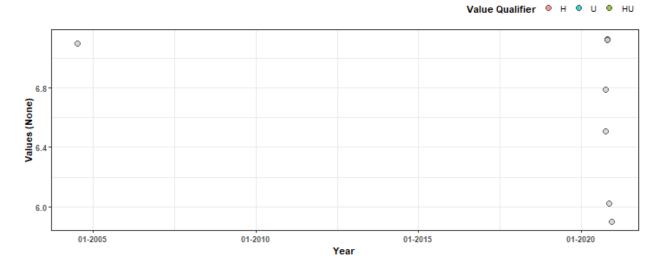


#### Scatter Plot of Excluded Managed Area Southeast Florida Coral Reef Ecosystem Conservation Area (3 Unique Years)



#### Scatter Plot of Excluded Managed Area Yellow River Marsh Aquatic Preserve (2 Unique Years)

Autoscale



## 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

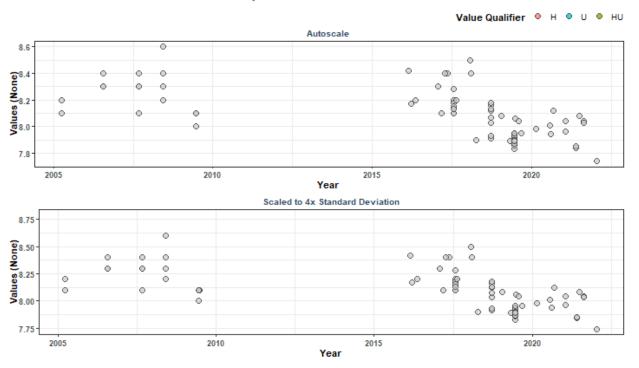
- 1. Use the data set that only has SufficientData of TRUE for the desired managed area
- 2. Determine the earliest and latest year of the data to create x-axis scale and intervals
- 3. Determine the minimum, mean, and standard deviation for the data to be used for y-axis scales
  - Excludes the top 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 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

```
sd_RV <- sd(plot_data$ResultValue[plot_data$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$ManagedAreaName==MA_Include[i]]
s_slope <- KT.Stats$SennSlope[KT.Stats$ManagedAreaName==MA_Include[i]]</pre>
s int <- KT.Stats$SennIntercept[KT.Stats$ManagedAreaName==MA Include[i]]
trend <- KT.Stats$Trend[KT.Stats$ManagedAreaName==MA Include[i]]</pre>
p <- KT.Stats$p[KT.Stats$ManagedAreaName==MA Include[i]]</pre>
model <- lm(ResultValue ~ DecDate,</pre>
            data=plot_data)
m_int <- coef(model)[[1]]</pre>
m_slope <- coef(model)[[2]]</pre>
rm(model)
p1 <- ggplot(data=plot_data,
             aes(x=DecDate, y=ResultValue, fill=VQ_Plot)) +
   geom_point(shape=21, size=3, color="#333333", alpha=0.75) +
   geom_abline(aes(slope=s_slope, intercept=s_int),
               color="blue", size=1.2) +
   labs(subtitle="Autoscale",
        x="Year", y=paste0("Values (", unit, ")"),
        fill="Value Qualifier") +
   plot theme +
   theme(legend.position="top", legend.box="horizontal",
         legend.justification="right") +
   {if(inc H==TRUE){
      scale_fill_manual(values=c("H"= "#F8766D", "U"= "#00BFC4",
                                  "HU"="#7CAE00"), na.value="#cccccc")
   } else if(param_name=="Secchi_Depth"){
      scale_fill_manual(values=c("S"= "#F8766D", "U"= "#00BFC4",
                                  "SU"="#7CAE00"), na.value="#cccccc")
   } else {
      scale_fill_manual(values=c("U"= "#00BFC4"), na.value="#cccccc")
   }}
p2 <- ggplot(data=plot_data,</pre>
             aes(x=DecDate, y=ResultValue, fill=VQ_Plot)) +
   geom_point(shape=21, size=3, color="#333333", alpha=0.75) +
   geom_abline(aes(slope=s_slope, intercept=s_int),
               color="blue", size=1.2) +
   ylim(min_RV, y_scale) +
   labs(subtitle="Scaled to 4x Standard Deviation",
        x="Year", y=paste0("Values (", unit, ")")) +
   plot theme +
   theme(legend.position="none") +
   {if(inc_H==TRUE){
      scale_fill_manual(values=c("H"= "#F8766D", "U"= "#00BFC4",
                                  "HU"="#7CAE00"), na.value="#cccccc")
   } else if(param_name=="Secchi_Depth"){
      scale_fill_manual(values=c("S"= "#F8766D", "U"= "#00BFC4",
                                  "SU"="#7CAE00"), na.value="#cccccc")
```

```
} else {
            scale_fill_manual(values=c("U"= "#00BFC4"), na.value="#cccccc")
         }}
      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 ",</pre>
                                           MA_Include[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)) +
         plot_theme + theme(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)
      rm(KTset, leg)
   }
}
```

#### Data Points with Trendlines for Alligator Harbor Aquatic Preserve

Senn Slope=-0.0227272727272726. Senn Intercept=88.8400000000001 Trend=-1, tau=-0.5452, p=0 Linear Trendline: y=-0.0195604752630904x + 47.5410954301561

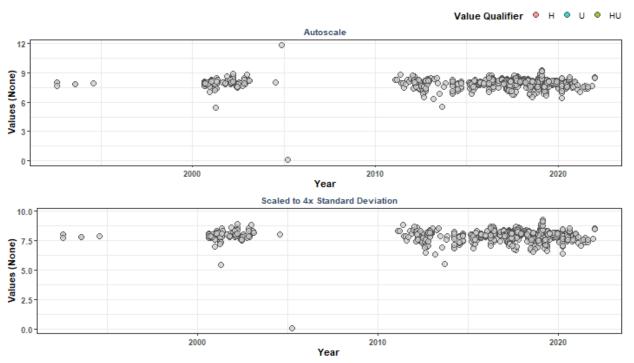


Data Points with Trendlines for Apalachicola Bay Aquatic Preserve

Senn Slope=0, Senn Intercept=12.6706842105263

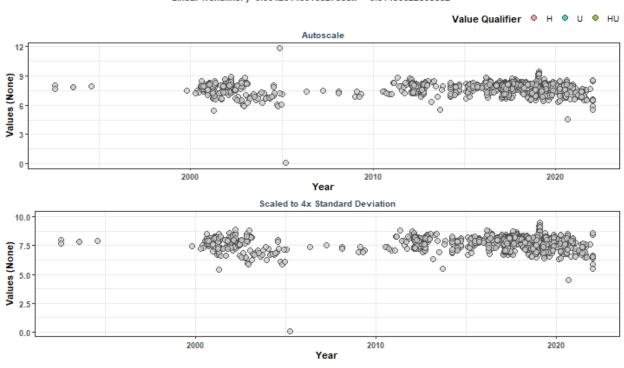
Trend=0, tau=-0.012, p=0.6304

Linear Trendline: y=0.000558656236993909x + 6.79436688918005



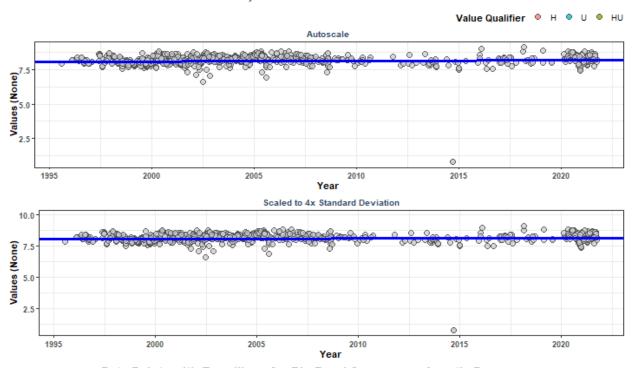
#### Data Points with Trendlines for Apalachicola National Estuarine Research Reserve

Senn Slope=0, Senn Intercept=17.1057900359874 Trend=0, tau=-0.006, p=0.8969 Linear Trendline: y=0.00426446016327935x + -0.81430622805552



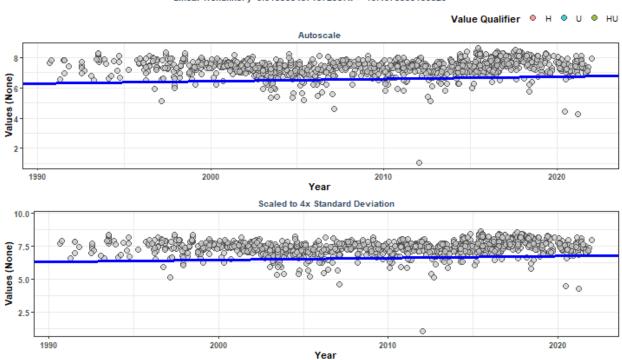
#### Data Points with Trendlines for Banana River Aquatic Preserve

Senn Slope=0.00428571428571428, Senn Intercept=-0.470000000000133 Trend=1, tau=0.0637, p=0.0042 Linear Trendline: y=0.0000557401713461378x + 8.0859285641281



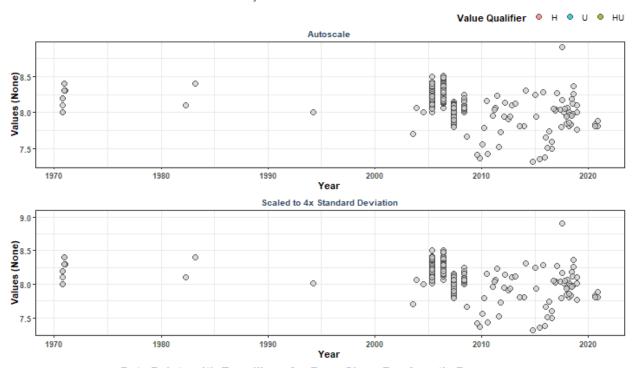
#### Data Points with Trendlines for Big Bend Seagrasses Aquatic Preserve

Senn Slope=0.01458333333333333, Senn Intercept=-22.7221969696969 Trend=1, tau=0.1547, p=0 Linear Trendline: y=0.0133334371572987x + -19.4075353133529

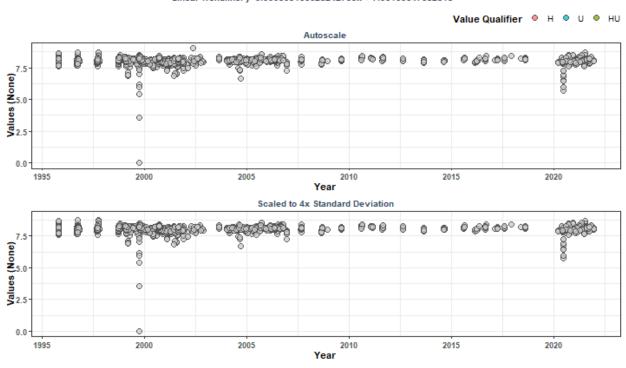


#### Data Points with Trendlines for Biscayne Bay Aquatic Preserve

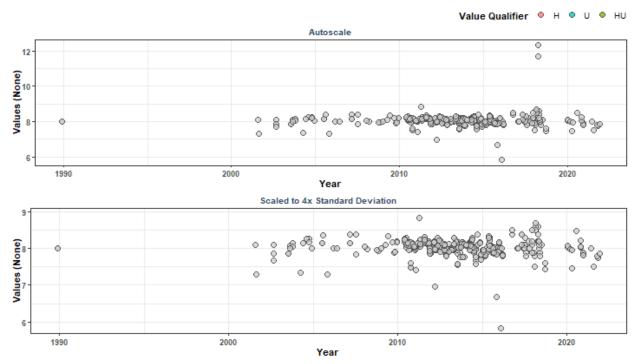
Senn Slope=-0.09333333333333337, Senn Intercept=11.4517850970348 Trend=-1, tau=-0.3373, p=0 Linear Trendline: y=-0.00767816451785355x + 23.523406907408



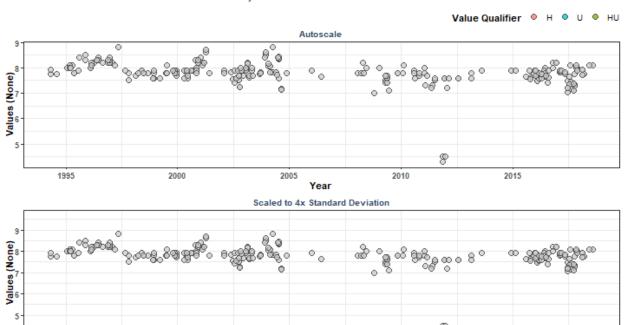
#### Data Points with Trendlines for Boca Ciega Bay Aquatic Preserve



#### Data Points with Trendlines for Cape Haze Aquatic Preserve

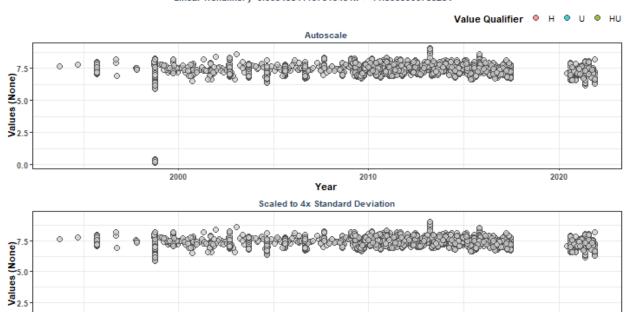


#### Data Points with Trendlines for Cape Romano-Ten Thousand Islands Aquatic Preserve



Year

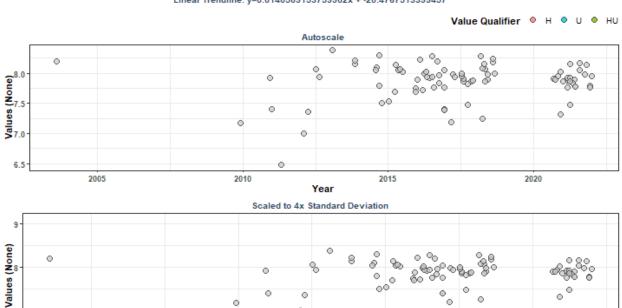
#### Data Points with Trendlines for Cockroach Bay Aquatic Preserve



#### Data Points with Trendlines for Estero Bay Aquatic Preserve

Year

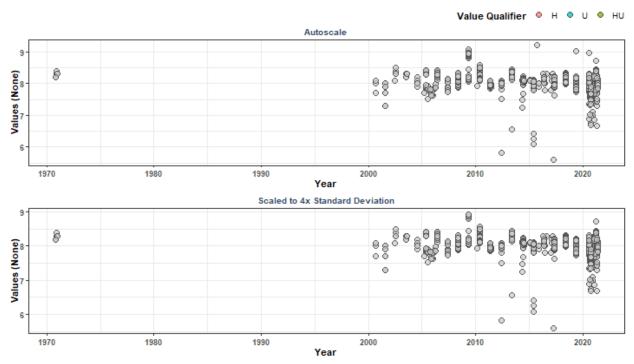
0.0



Year

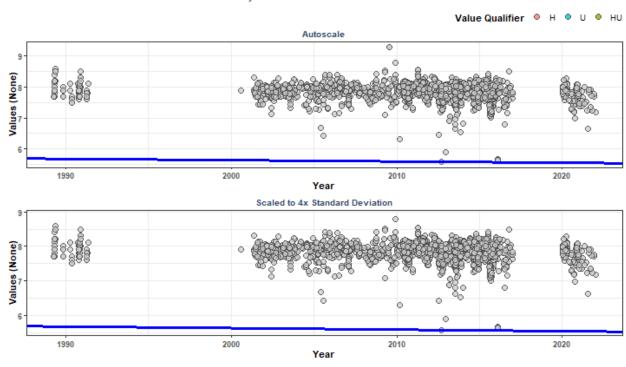
#### Data Points with Trendlines for Florida Keys National Marine Sanctuary

Senn Slope=-0.0109999999999999, Senn Intercept=14.778 Trend=-1, tau=-0.0935, p=0 Linear Trendline: y=-0.00956569686028797x + 27.3488669156959

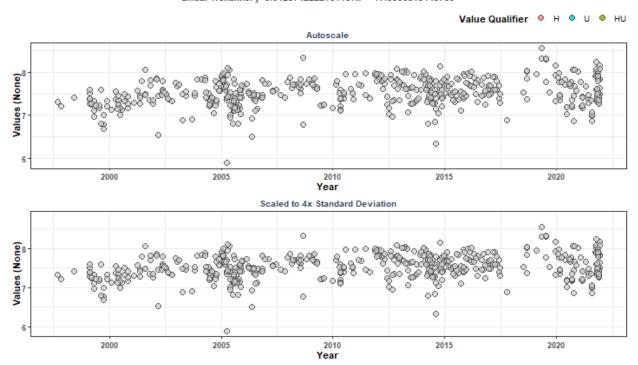


#### Data Points with Trendlines for Gasparilla Sound-Charlotte Harbor Aquatic Preserve

Senn Slope=-0.00428571428571432, Senn Intercept=14.2043750000001 Trend=-1, tau=-0.0645, p=0 Linear Trendline: y=-0.00372966469572219x + 15.3731815640876

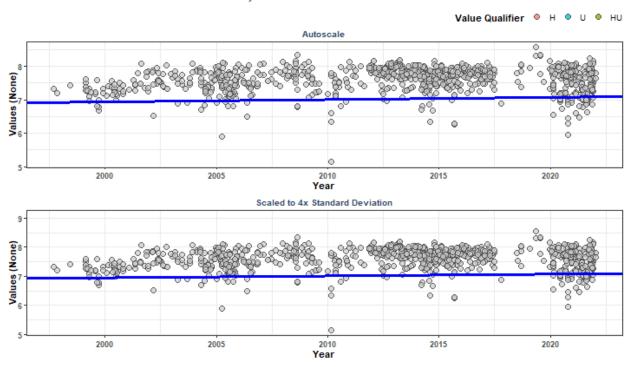


#### Data Points with Trendlines for Guana River Marsh Aquatic Preserve



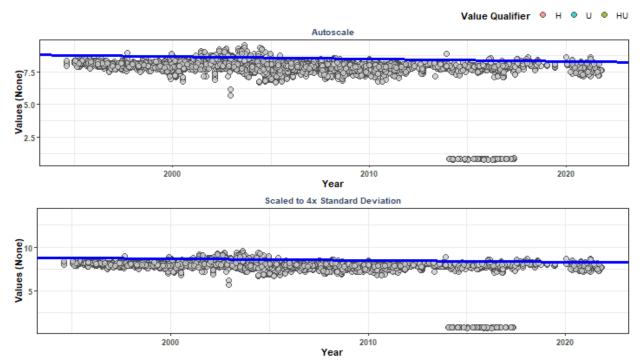
#### Data Points with Trendlines for Guana Tolomato Matanzas National Estuarine Research Reserve

Senn Slope=0.0066666666666666667, Senn Intercept=-6.3880000000001 Trend=1, tau=0.0939, p=0.0002 Linear Trendline: y=0.00838116947823506x + -9.26274954189448



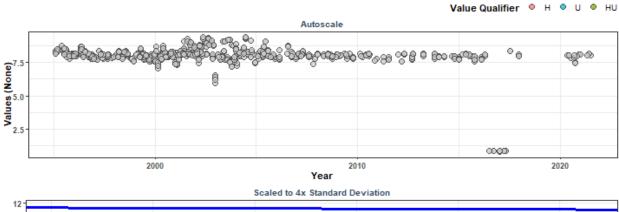
#### Data Points with Trendlines for Indian River-Malabar to Vero Beach Aquatic Preserve

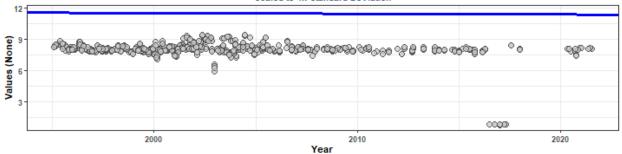
Senn Slope=-0.01875, Senn Intercept=46.1844298245614 Trend=-1, tau=-0.2558, p=0 Linear Trendline: y=-0.0928784715894398x + 193.95407122974



#### Data Points with Trendlines for Indian River-Vero Beach to Ft. Pierce Aquatic Preserve

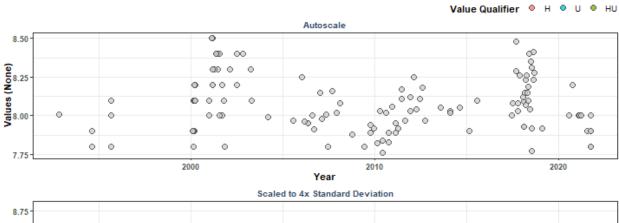
Senn Slope=-0.00705882352941175, Senn Intercept=25.5924337121212 Trend=-1, tau=-0.0901, p=0 Linear Trendline: y=-0.0478344640667845x + 103.823160662827

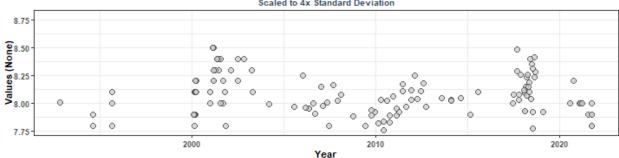




#### Data Points with Trendlines for Jensen Beach to Jupiter Inlet Aquatic Preserve

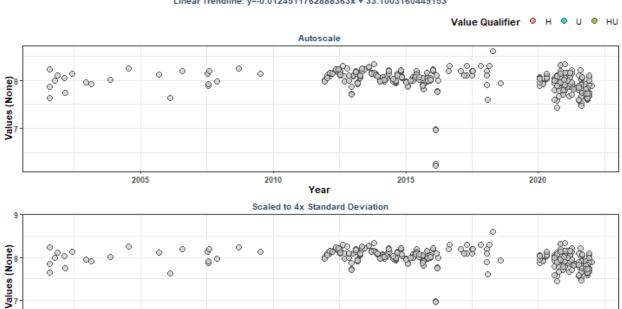
Senn Slope=0, Senn Intercept=9.85498581011348 Trend=0, tau=0.0015, p=1 Linear Trendline: y=-0.00180051590960288x + 11.6899892239701





#### Data Points with Trendlines for Lemon Bay Aquatic Preserve

Senn Slope=-0.01833333333333334, Senn Intercept=39.076339285714 Trend=-1, tau=-0.3016, p=0 Linear Trendline: y=-0.0124511762888363x + 33.1003160449153



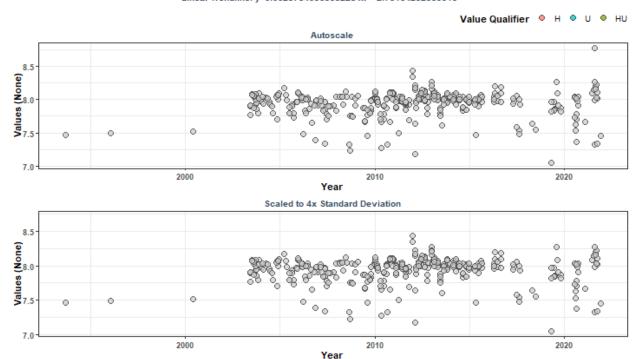
Year

2015

2020

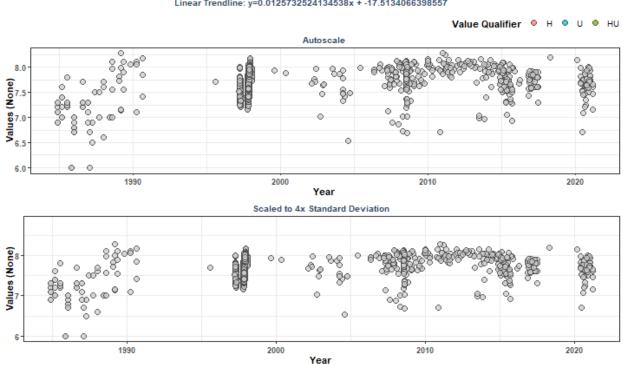
2010

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



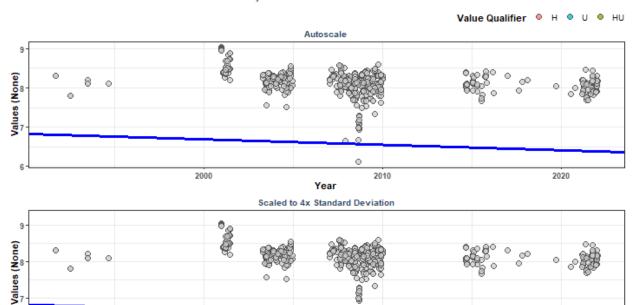
#### Data Points with Trendlines for Nassau River-St. Johns River Marshes Aquatic Preserve

Senn Slope=0.0164285714285714, Senn Intercept=-12.5160000000001 Trend=1, tau=0.1053, p=0 Linear Trendline: y=0.0125732524134538x + -17.5134066398557



#### Data Points with Trendlines for Nature Coast Aquatic Preserve

Senn Slope=-0.01400000000000001, Senn Intercept=34.686 Trend=-1, tau=-0.2326, p=0 Linear Trendline: y=-0.011016543349694x + 30.2517948269761



#### Data Points with Trendlines for Pine Island Sound Aquatic Preserve

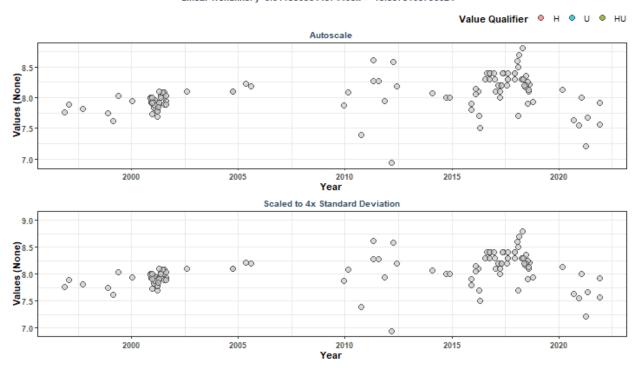
Year

2010

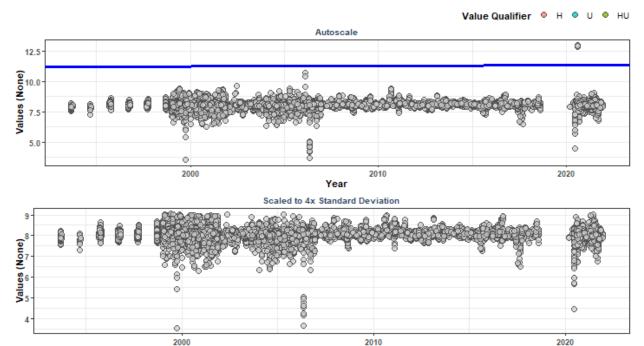
2020

2000

Senn Slope=0.011764705882353, Senn Intercept=-13.245625 Trend=1, tau=0.2418, p=0.0015 Linear Trendline: y=0.0118803644374403x + -15.8378105796024



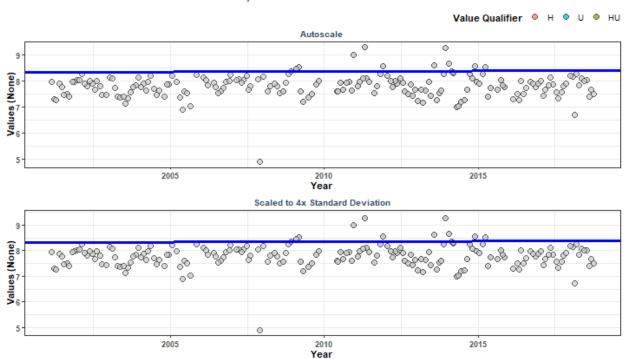
#### Data Points with Trendlines for Pinellas County Aquatic Preserve



#### Data Points with Trendlines for Rocky Bayou State Park Aquatic Preserve

Year

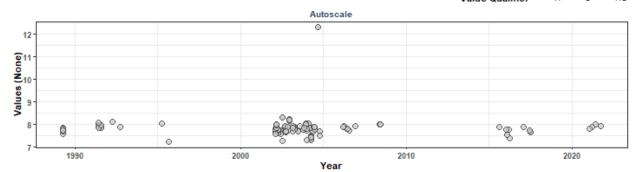
Senn Slope=0.0033333333333333326, Senn Intercept=1.65708333333312 Trend=0, tau=0.0399, p=0.4283 Linear Trendline: y=0.00660546256584331x + -5.45484640568164



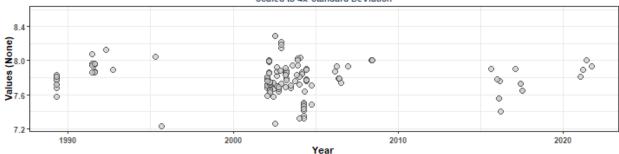
#### Data Points with Trendlines for Rookery Bay Aquatic Preserve

Senn Slope=-0.00692307692307691, Senn Intercept=13.5839438502674 Trend=-1, tau=-0.105, p=0.019 Linear Trendline: y=0.000654026174736869x + 6.5857595492979



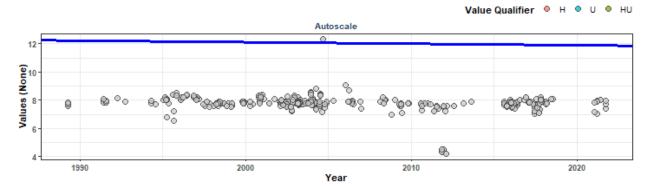


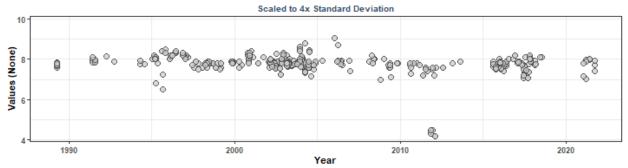
#### Scaled to 4x Standard Deviation



#### Data Points with Trendlines for Rookery Bay National Estuarine Research Reserve

Senn Slope=-0.0107692307692308, Senn Intercept=33.6287873931625 Trend=-1, tau=-0.2119, p=0 Linear Trendline: y=-0.0165147342935684x + 40.9191759835324



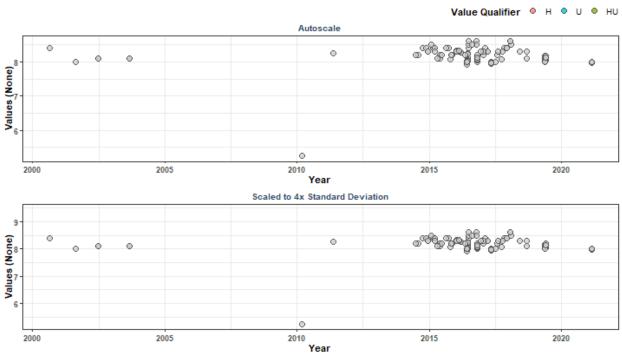


Data Points with Trendlines for St. Andrews State Park Aquatic Preserve

Senn Slope=0.01500000000000001, Senn Intercept=-27.110000000001

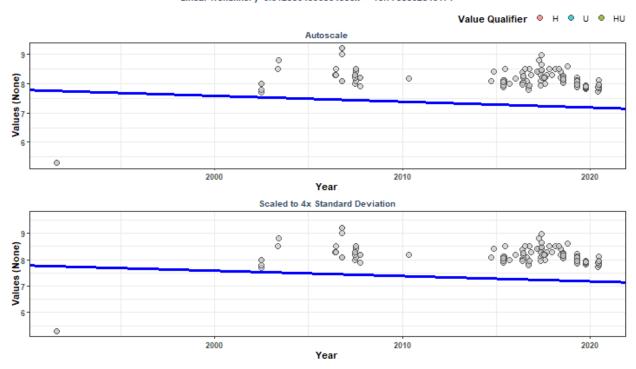
Trend=0, tau=0.0687, p=0.1355

Linear Trendline: y=0.00769431970447494x + -7.34923685505622



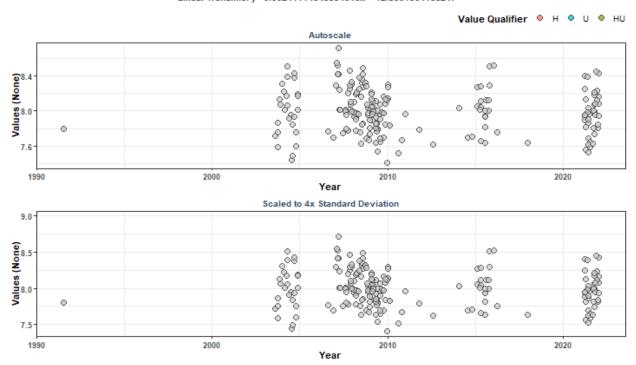
#### Data Points with Trendlines for St. Joseph Bay Aquatic Preserve

Senn Slope=-0.02, Senn Intercept=47.58357142857 Trend=-1, tau=-0.1791, p=0.0017 Linear Trendline: y=0.0123504390864556x + -16.7758802840171



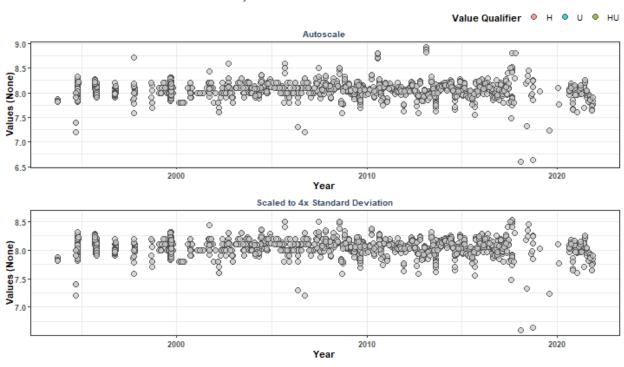
#### Data Points with Trendlines for St. Martins Marsh Aquatic Preserve

Senn Slope=-0.00571428571428572, Senn Intercept=17.7268137254903 Trend=0, tau=-0.0892, p=0.0717 Linear Trendline: y=-0.00241114848834646x + 12.8501364168247

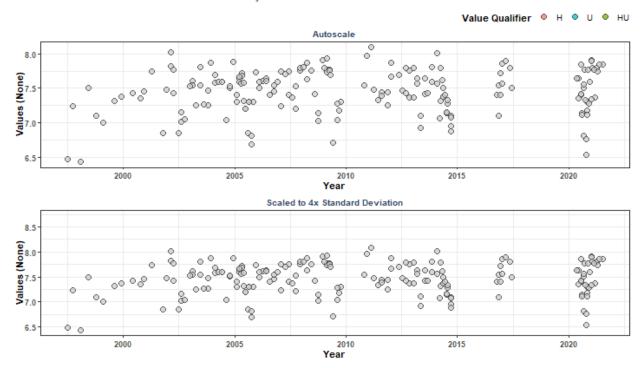


#### Data Points with Trendlines for Terra Ceia Aquatic Preserve

Senn Slope=0.0033333333333333326, Senn Intercept=5.48075630252081 Trend=1, tau=0.0497, p=0 Linear Trendline: y=-0.000179651632942128x + 8.41182910522596



#### Data Points with Trendlines for Tomoka Marsh Aquatic Preserve



#### 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 SufficientData 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:

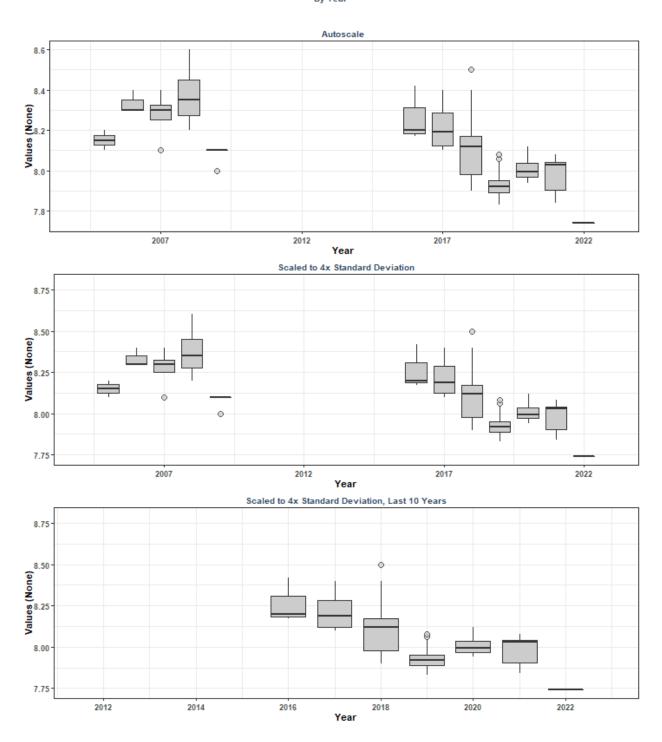
- 1. Y-axis autoscaled
- 2. Y-axis set to be mean + 4 times the standard deviation
- 3. Y-axis set to be mean + 4 times the standard deviation for most recent 10 years of data

```
if(n==0){
   print("There are no managed areas that qualify.")
} else {
   for (i in 1:n) {
      plot_data <- data[data$SufficientData==TRUE &</pre>
                            data$ManagedAreaName==MA Include[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[plot_data$ResultValue <</pre>
                                              quantile(data$ResultValue, 0.98)])
      sd_RV <- sd(plot_data$ResultValue[plot_data$ResultValue <</pre>
                                             quantile(data$ResultValue, 0.98)])
      x_scale <- ifelse(year_upper - year_lower > 30, 10, 5)
      y_scale \leftarrow mn_RV + 4 * sd_RV
      ##Year plots
      p1 <- ggplot(data=plot_data,</pre>
                   aes(x=Year, y=ResultValue, group=Year)) +
         geom_boxplot(color="#333333", fill="#cccccc", outlier.shape=21,
                       outlier.size=3, outlier.color="#333333",
                      outlier.fill="#cccccc", outlier.alpha=0.75) +
         labs(subtitle="Autoscale",
              x="Year", y=paste0("Values (", unit, ")")) +
         scale_x_continuous(limits=c(year_lower - 1, year_upper + 1),
                             breaks=rev(seq(year_upper,
                                            year lower, -x scale))) +
         plot_theme
      p2 <- ggplot(data=plot_data,</pre>
                   aes(x=Year, y=ResultValue, group=Year)) +
         geom_boxplot(color="#333333", fill="#cccccc", outlier.shape=21,
                      outlier.size=3, outlier.color="#333333",
                      outlier.fill="#cccccc", outlier.alpha=0.75) +
         labs(subtitle="Scaled to 4x Standard Deviation",
              x="Year", y=paste0("Values (", unit, ")")) +
         ylim(min_RV, y_scale) +
         scale_x_continuous(limits=c(year_lower - 1, year_upper + 1),
                             breaks=rev(seq(year_upper,
                                            year_lower, -x_scale))) +
         plot_theme
      p3 <- ggplot(data=plot_data[plot_data$Year >= year_upper - 10, ],
                   aes(x=Year, y=ResultValue, group=Year)) +
         geom_boxplot(color="#333333", fill="#cccccc", outlier.shape=21,
                      outlier.size=3, outlier.color="#333333",
                       outlier.fill="#cccccc", outlier.alpha=0.75) +
         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))) +
         plot_theme
```

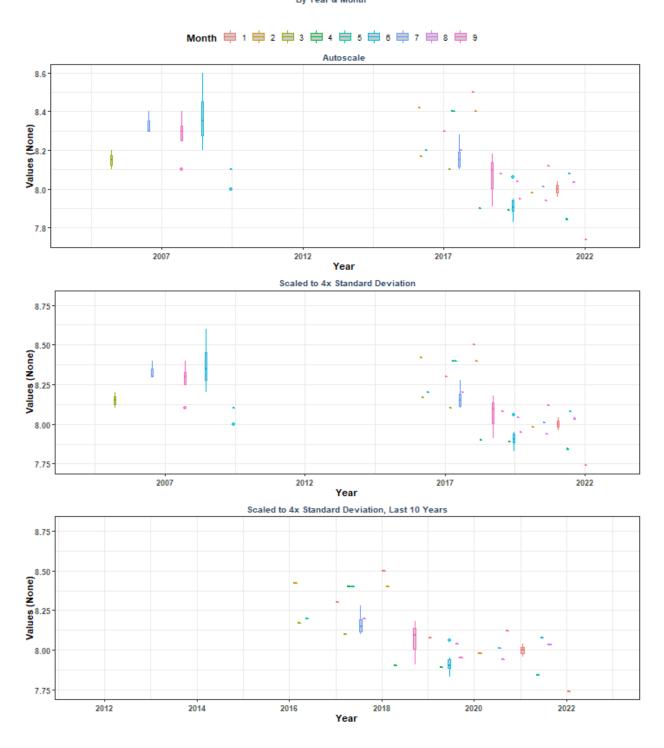
```
Yset <- ggarrange(p1, p2, p3, ncol=1)</pre>
p0 <- ggplot() + labs(title=paste0("Summary Box Plots for ",
                                   MA_Include[i]),
                      subtitle="By Year") +
   plot_theme + theme(panel.border=element_blank(),
                      panel.grid.major=element_blank(),
                      panel.grid.minor=element_blank(),
                      axis.line=element_blank())
## Year & Month Plots
p4 <- ggplot(data=plot_data,
             aes(x=YearMonthDec, y=ResultValue,
                 group=YearMonth, color=as.factor(Month))) +
   geom_boxplot(fill="#cccccc", outlier.size=1.5, outlier.alpha=0.75) +
   labs(subtitle="Autoscale",
        x="Year", y=paste0("Values (", unit, ")"), color="Month") +
   scale_x_continuous(limits=c(year_lower - 1, year_upper + 1),
                      breaks=rev(seq(year_upper,
                                     year_lower, -x_scale))) +
   plot_theme +
   theme(legend.position="none")
p5 <- ggplot(data=plot_data,
             aes(x=YearMonthDec, y=ResultValue,
                 group=YearMonth, color=as.factor(Month))) +
   geom_boxplot(fill="#cccccc", outlier.size=1.5, outlier.alpha=0.75) +
   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))) +
   plot_theme +
   theme(legend.position="top", legend.box="horizontal") +
   guides(color=guide_legend(nrow=1))
p6 <- ggplot(data=plot_data[plot_data$Year >= year_upper - 10, ],
             aes(x=YearMonthDec, y=ResultValue,
                 group=YearMonth, color=as.factor(Month))) +
   geom_boxplot(fill="#cccccc", outlier.size=1.5, outlier.alpha=0.75) +
   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))) +
   plot_theme +
   theme(legend.position="none")
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_Include[i]),
                       subtitle="By Year & Month") + plot_theme +
   theme(panel.border=element blank(),
         panel.grid.major=element_blank(),
         panel.grid.minor=element_blank(), axis.line=element_blank())
## Month Plots
p7 <- ggplot(data=plot_data,
             aes(x=Month, y=ResultValue,
                 group=Month, fill=as.factor(Month))) +
   geom_boxplot(color="#333333", outlier.shape=21, outlier.size=3,
                outlier.color="#333333", outlier.alpha=0.75) +
   labs(subtitle="Autoscale",
        x="Month", y=paste0("Values (", unit, ")"), fill="Month") +
   scale_x_continuous(limits=c(0, 13), breaks=seq(3, 12, 3)) +
   plot_theme +
   theme(legend.position="none")
p8 <- ggplot(data=plot_data,
             aes(x=Month, y=ResultValue,
                 group=Month, fill=as.factor(Month))) +
   geom_boxplot(color="#333333", outlier.shape=21, outlier.size=3,
                outlier.color="#333333", outlier.alpha=0.75) +
   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)) +
   plot_theme +
   theme(legend.position="top", legend.box="horizontal") +
   guides(fill=guide_legend(nrow=1))
p9 <- ggplot(data=plot_data[plot_data$Year >= year_upper - 10, ],
             aes(x=Month, y=ResultValue,
                 group=Month, fill=as.factor(Month))) +
   geom_boxplot(color="#333333", outlier.shape=21, outlier.size=3,
                outlier.color="#333333", outlier.alpha=0.75) +
   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)) +
   plot theme +
   theme(legend.position="none")
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_Include[i]),
                        subtitle="By Month") + plot_theme +
   theme(panel.border=element_blank(),
         panel.grid.major=element_blank(),
```

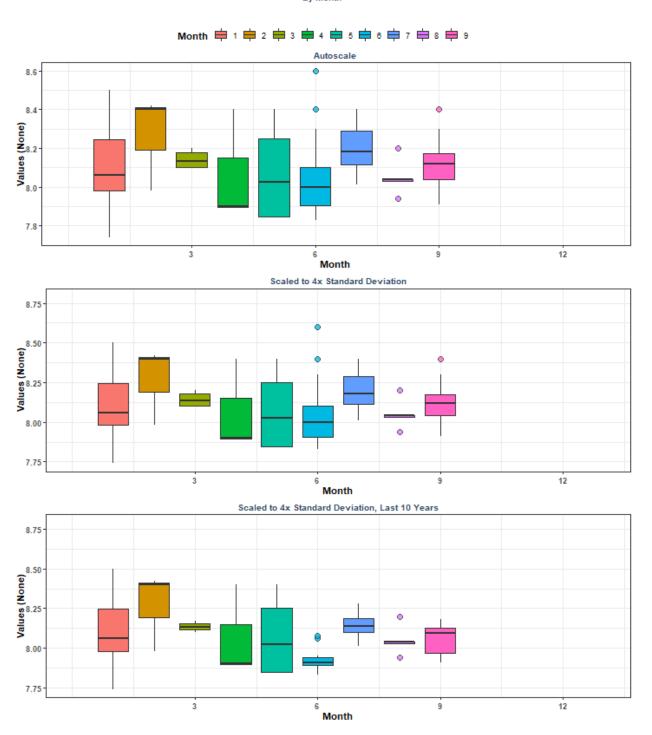
### Summary Box Plots for Alligator Harbor Aquatic Preserve



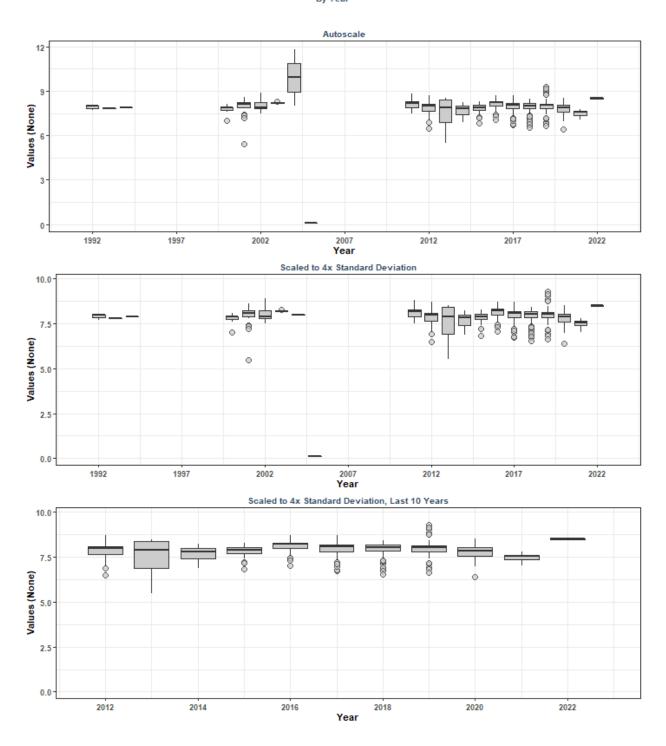
# Summary Box Plots for Alligator Harbor Aquatic Preserve By Year & Month



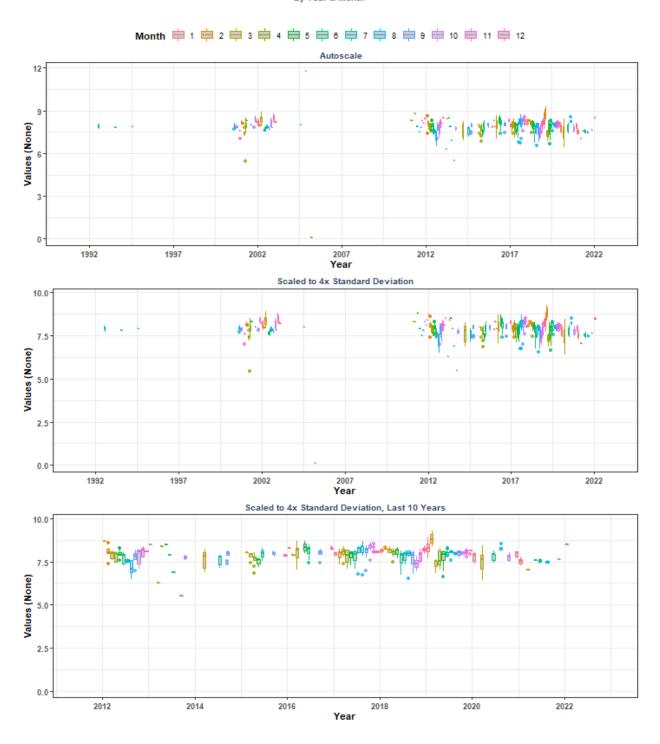
### Summary Box Plots for Alligator Harbor Aquatic Preserve By Month



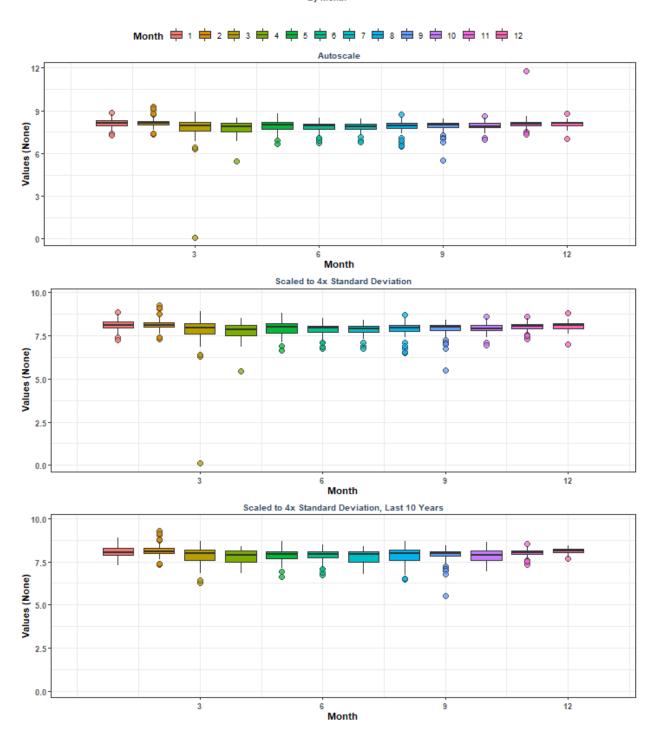
### Summary Box Plots for Apalachicola Bay Aquatic Preserve By Year



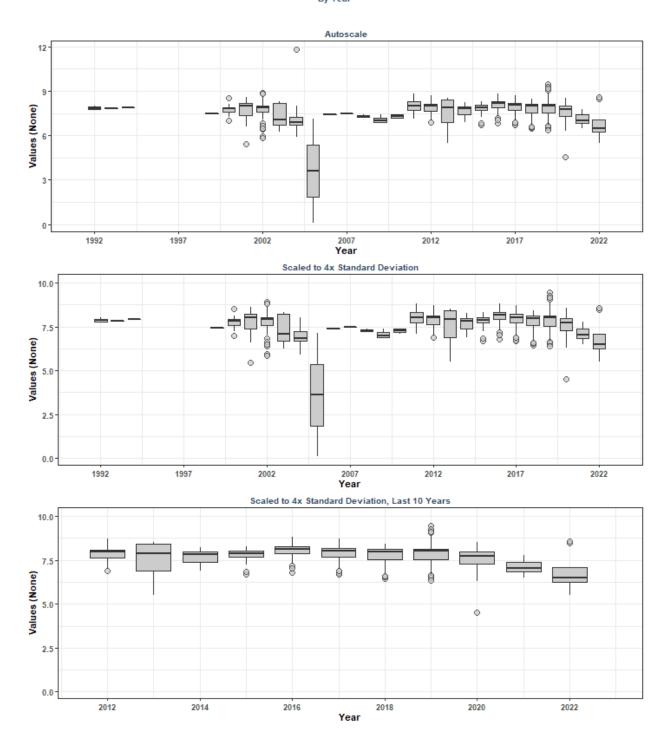
### Summary Box Plots for Apalachicola Bay Aquatic Preserve By Year & Month



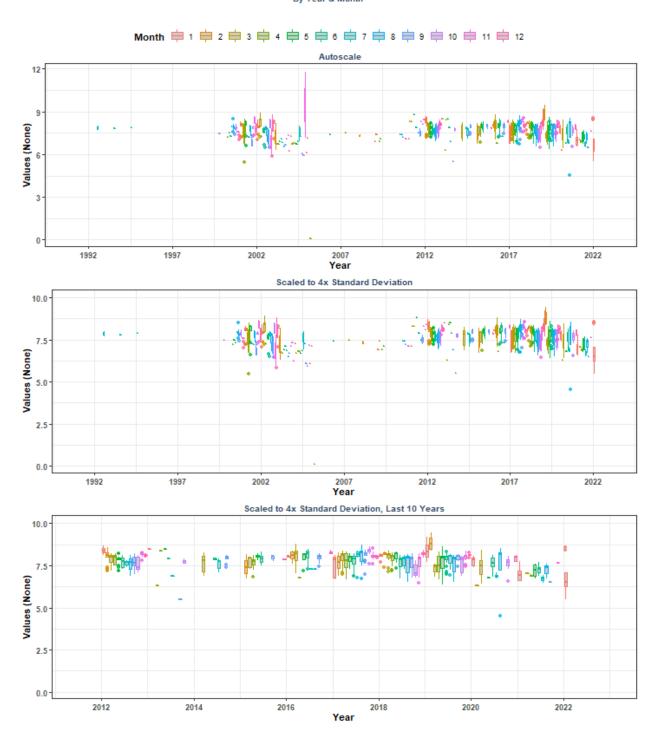
# Summary Box Plots for Apalachicola Bay Aquatic Preserve By Month



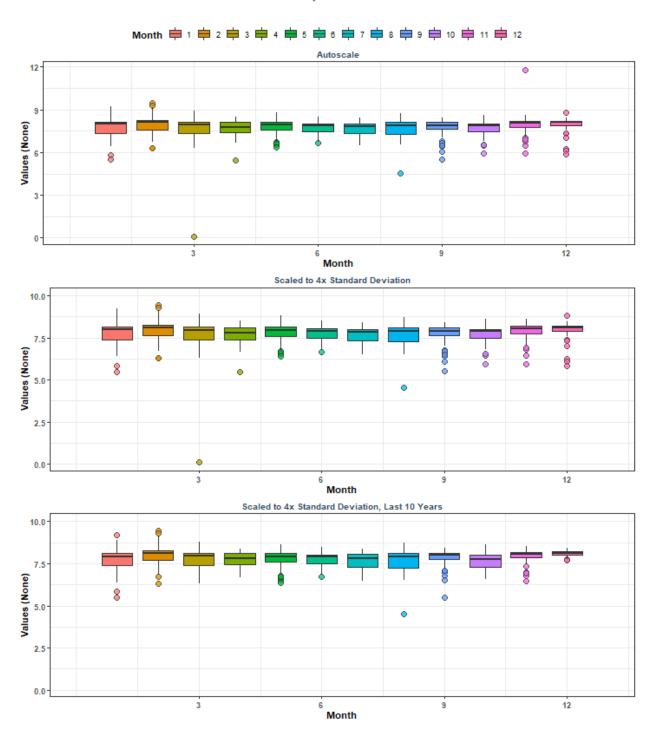
### Summary Box Plots for Apalachicola National Estuarine Research Reserve By Year



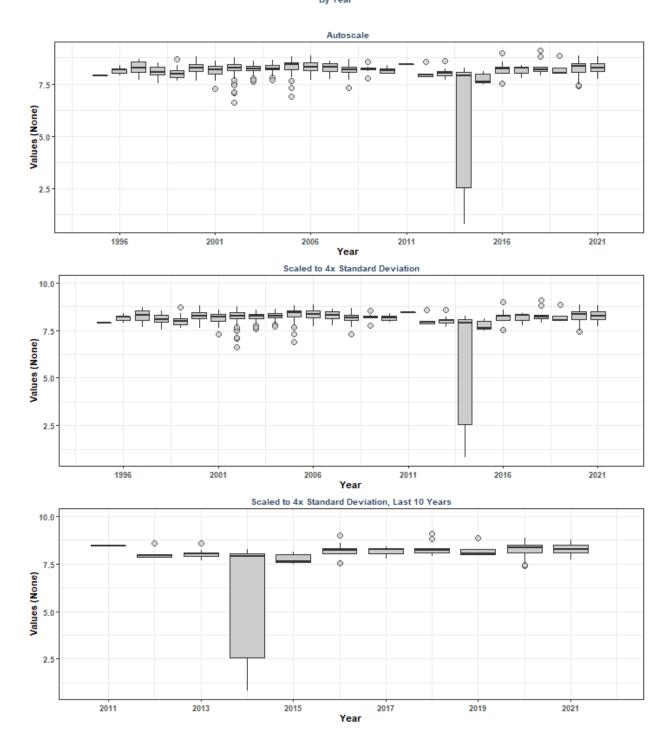
#### Summary Box Plots for Apalachicola National Estuarine Research Reserve By Year & Month



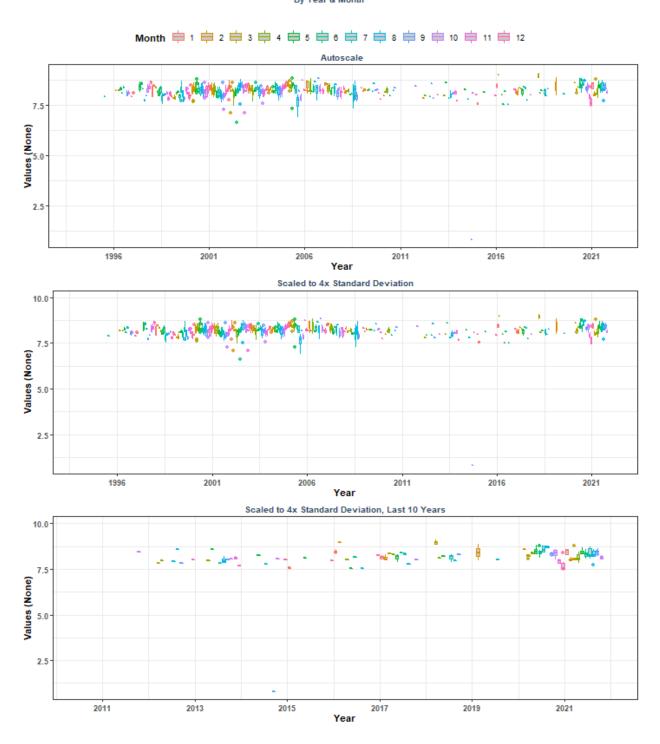
# Summary Box Plots for Apalachicola National Estuarine Research Reserve By Month



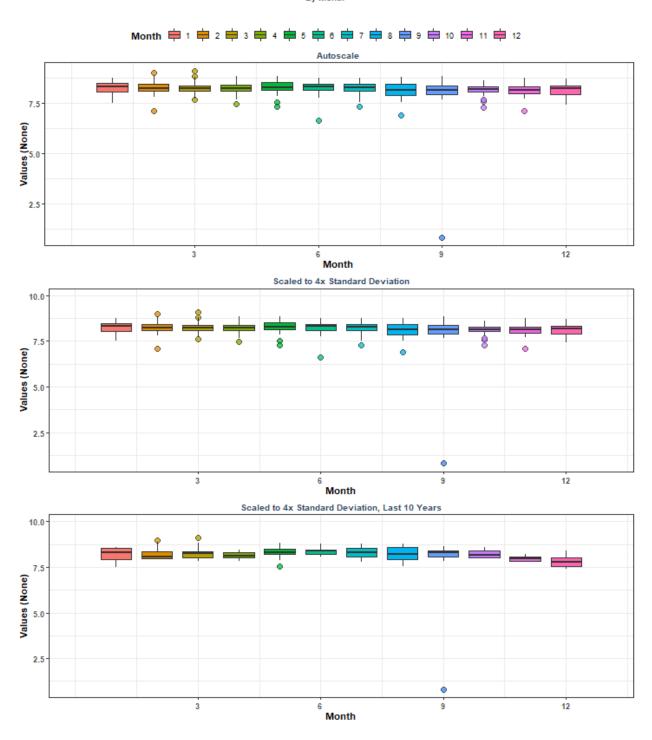
### Summary Box Plots for Banana River Aquatic Preserve By Year



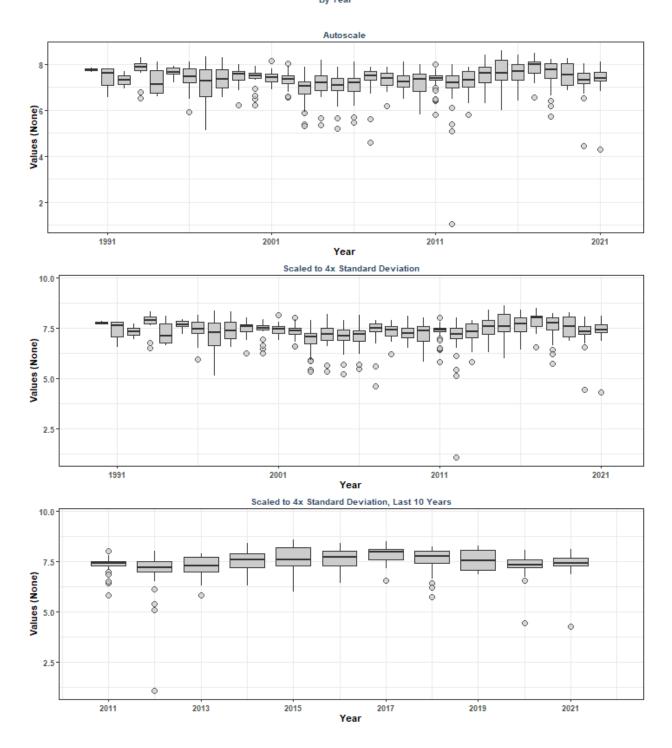
### Summary Box Plots for Banana River Aquatic Preserve By Year & Month



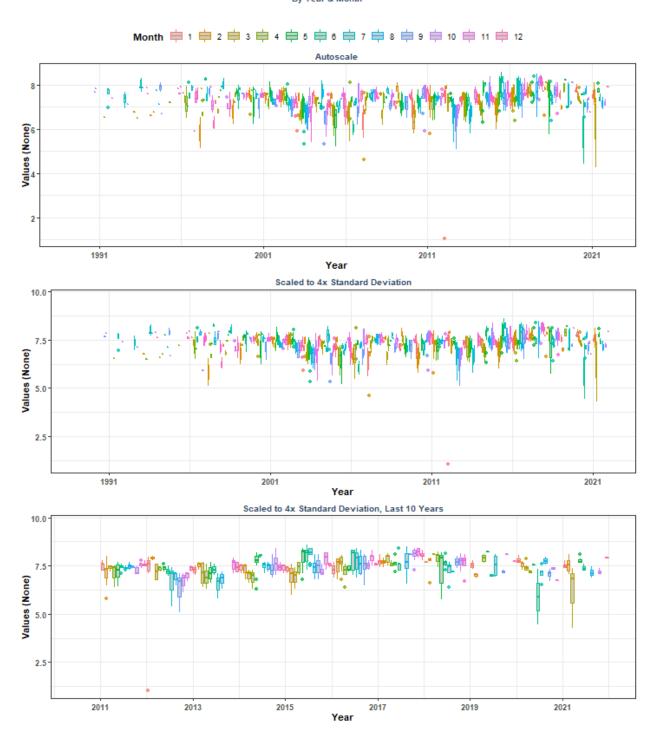
### Summary Box Plots for Banana River Aquatic Preserve By Month



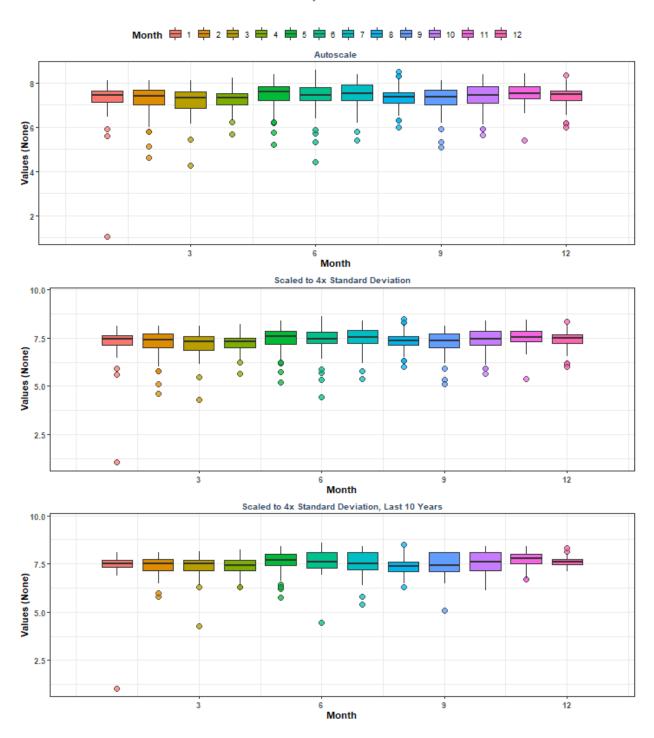
### Summary Box Plots for Big Bend Seagrasses Aquatic Preserve By Year



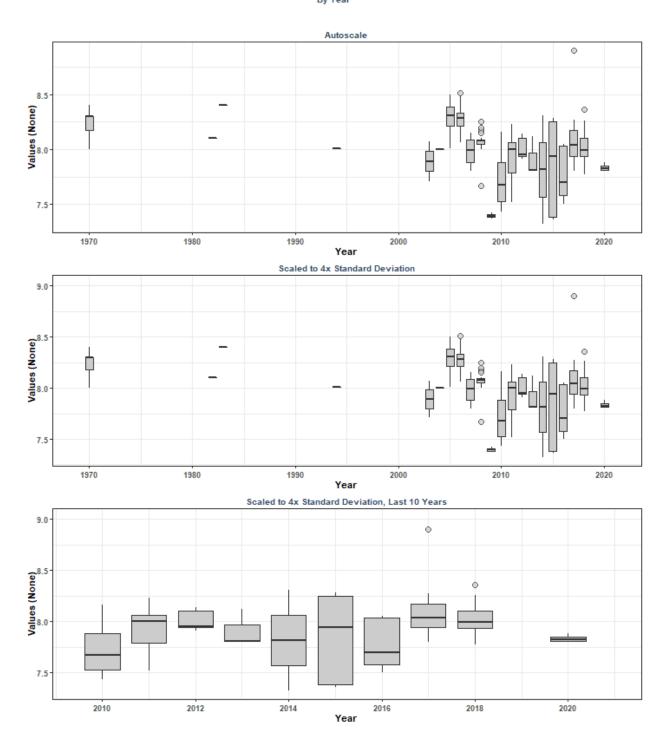
### Summary Box Plots for Big Bend Seagrasses Aquatic Preserve By Year & Month



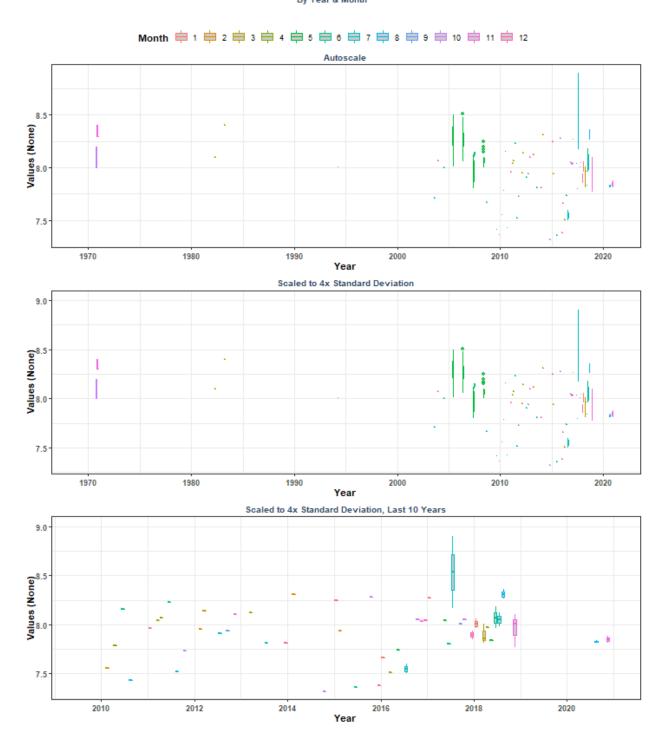
#### Summary Box Plots for Big Bend Seagrasses Aquatic Preserve By Month



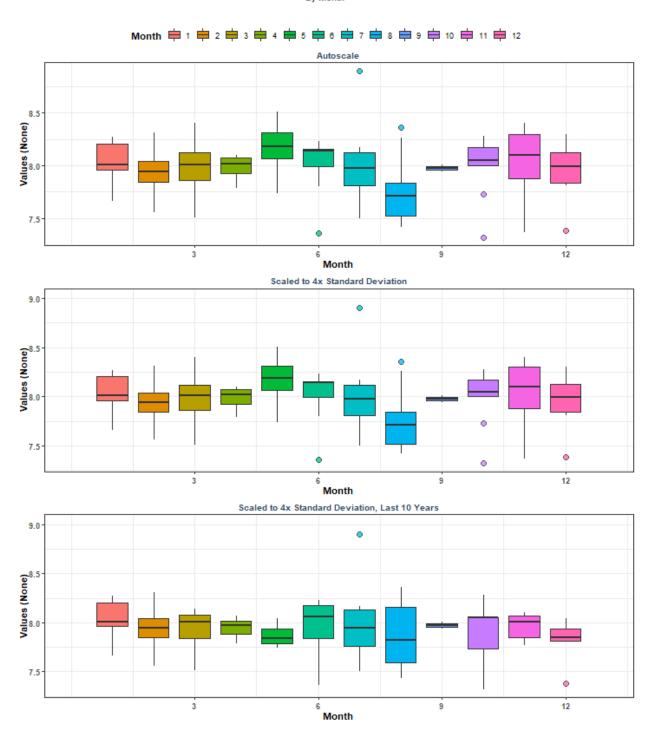
### Summary Box Plots for Biscayne Bay Aquatic Preserve By Year



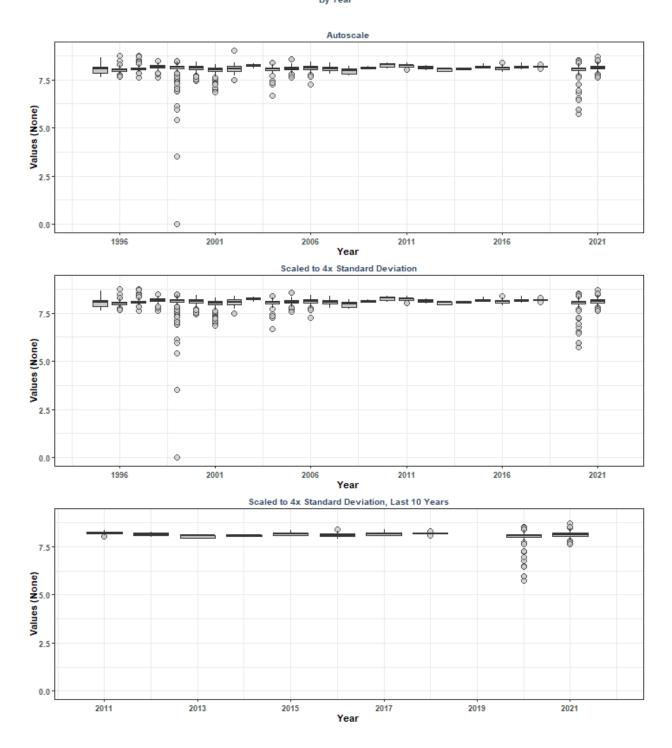
### Summary Box Plots for Biscayne Bay Aquatic Preserve By Year & Month



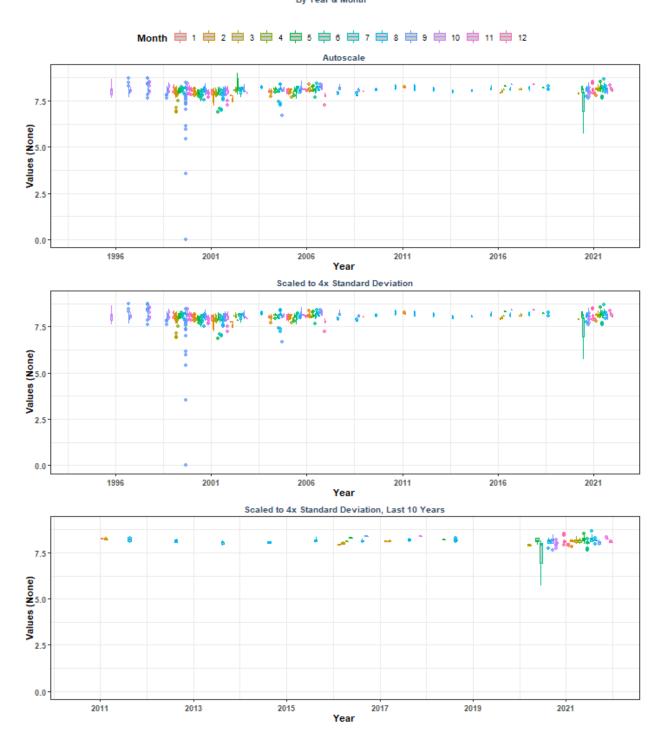
### Summary Box Plots for Biscayne Bay Aquatic Preserve By Month



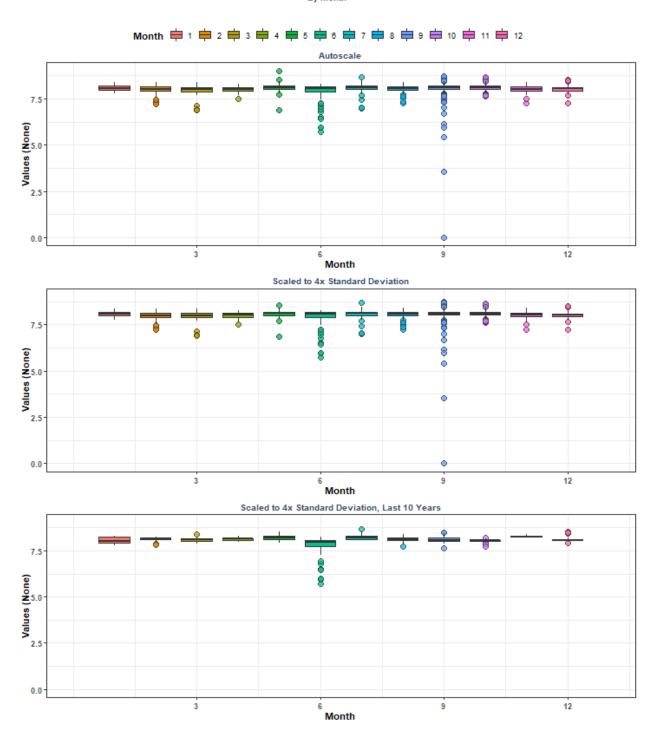
### Summary Box Plots for Boca Ciega Bay Aquatic Preserve By Year



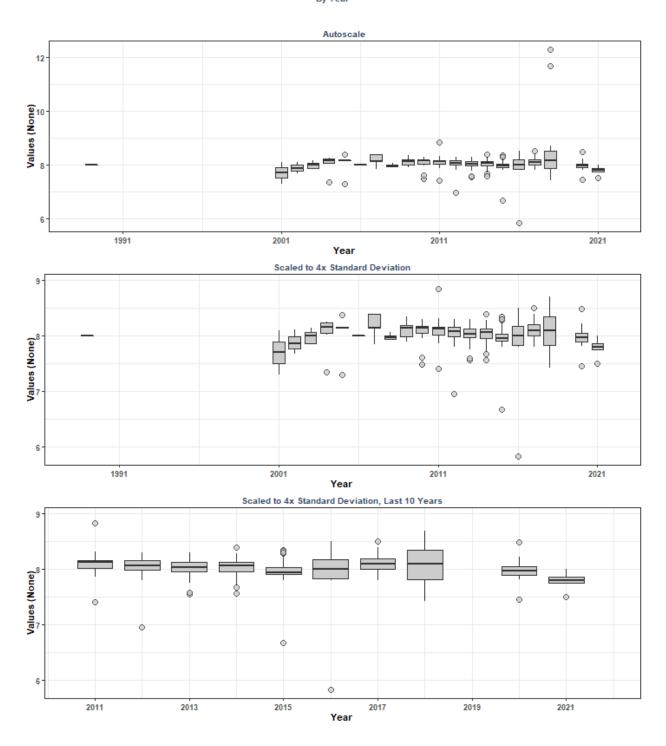
#### Summary Box Plots for Boca Ciega Bay Aquatic Preserve By Year & Month



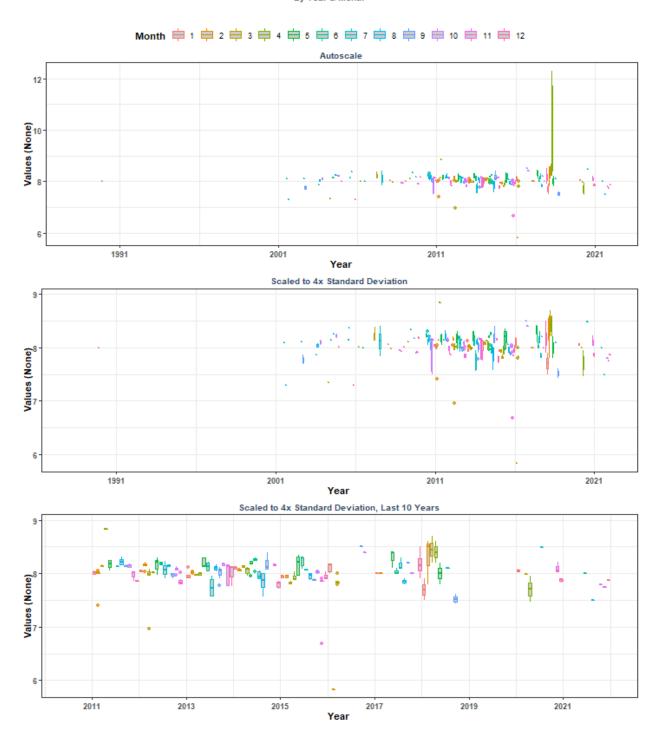
### Summary Box Plots for Boca Ciega Bay Aquatic Preserve By Month



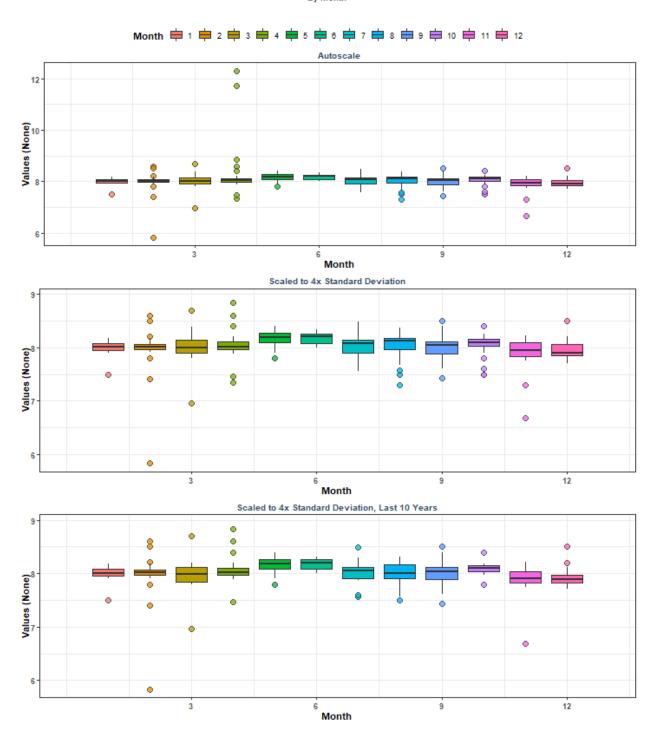
### Summary Box Plots for Cape Haze Aquatic Preserve $$_{\rm By\,Year}$$



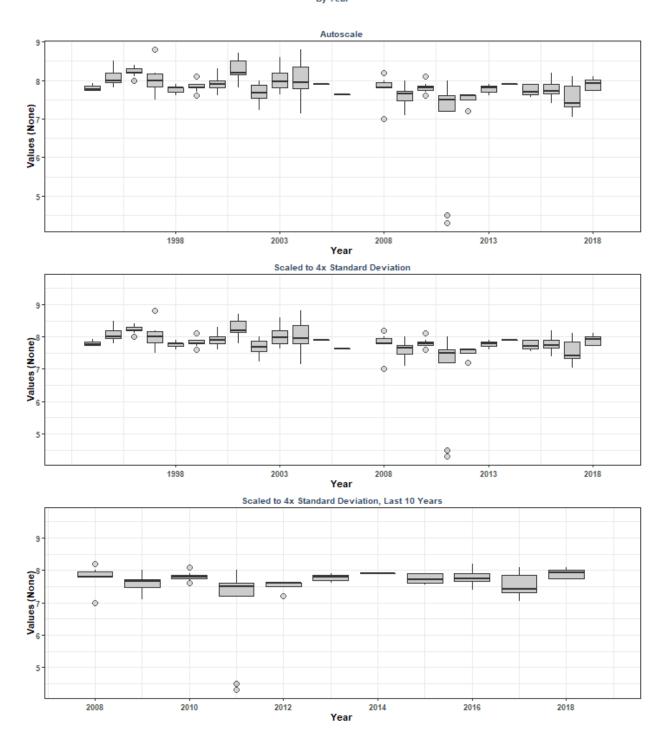
### Summary Box Plots for Cape Haze Aquatic Preserve By Year & Month



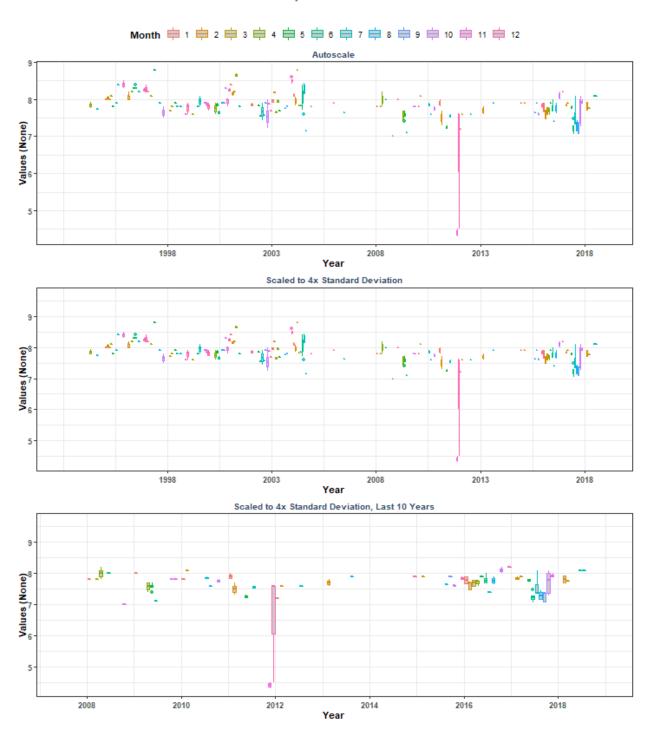
### Summary Box Plots for Cape Haze Aquatic Preserve By Month



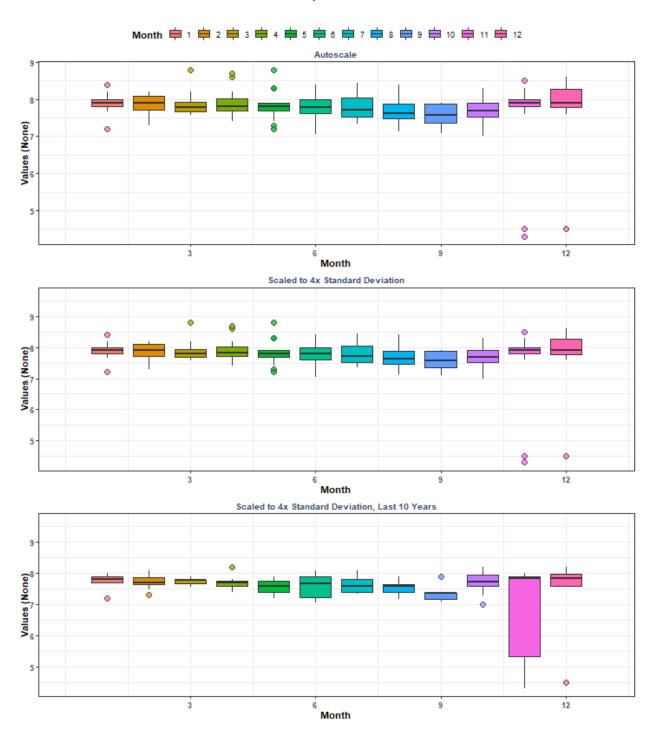
### Summary Box Plots for Cape Romano-Ten Thousand Islands Aquatic Preserve $$_{\rm By\ Year}$$



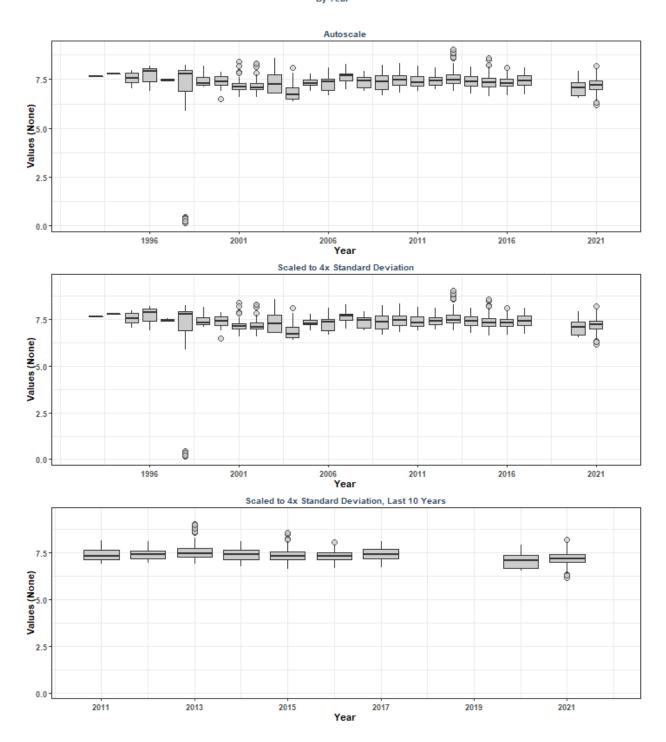
#### Summary Box Plots for Cape Romano-Ten Thousand Islands Aquatic Preserve By Year & Month



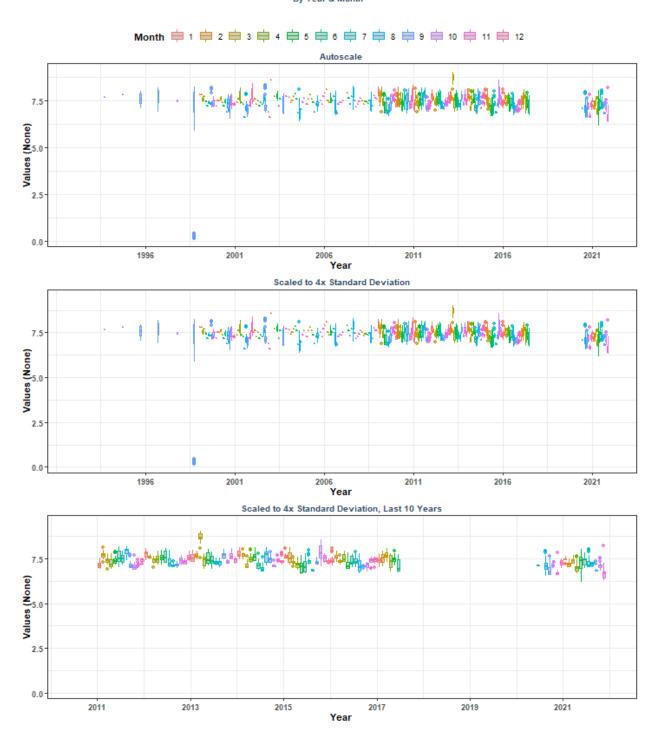
## Summary Box Plots for Cape Romano-Ten Thousand Islands Aquatic Preserve By Month



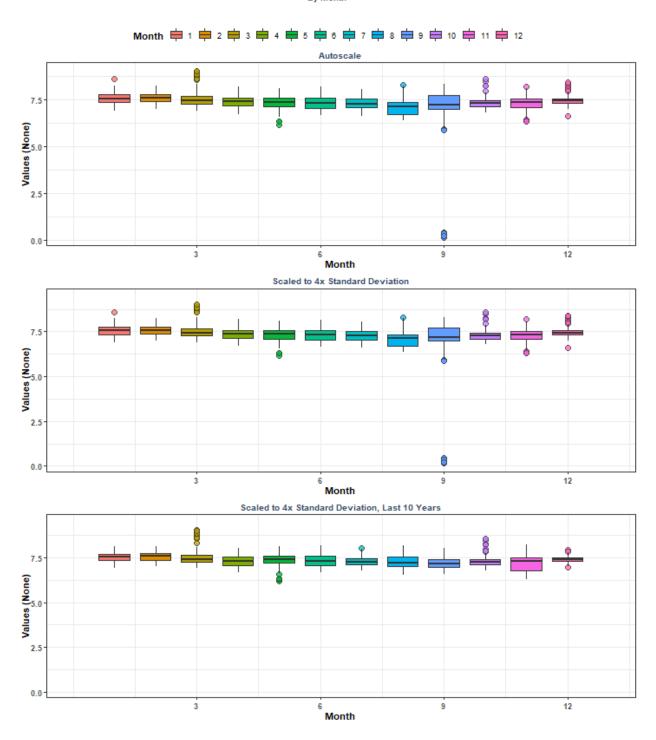
## Summary Box Plots for Cockroach Bay Aquatic Preserve By Year



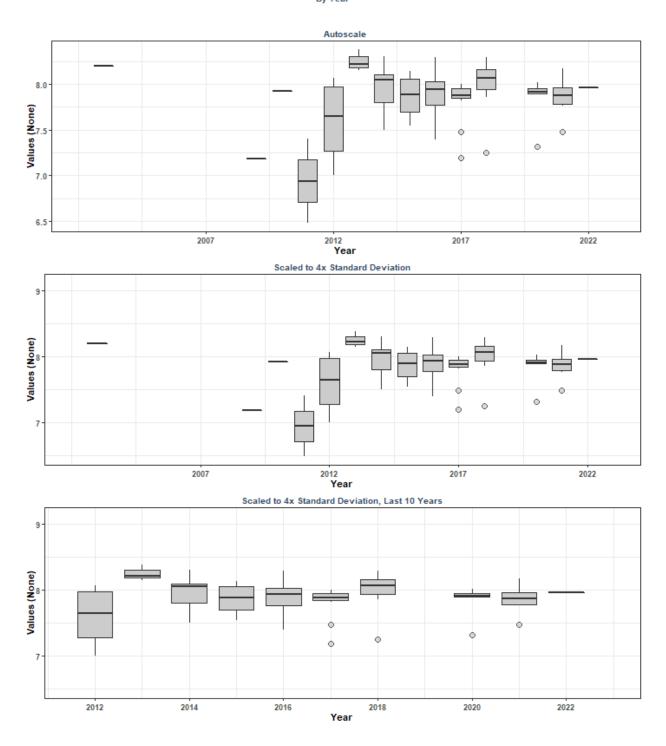
#### Summary Box Plots for Cockroach Bay Aquatic Preserve By Year & Month



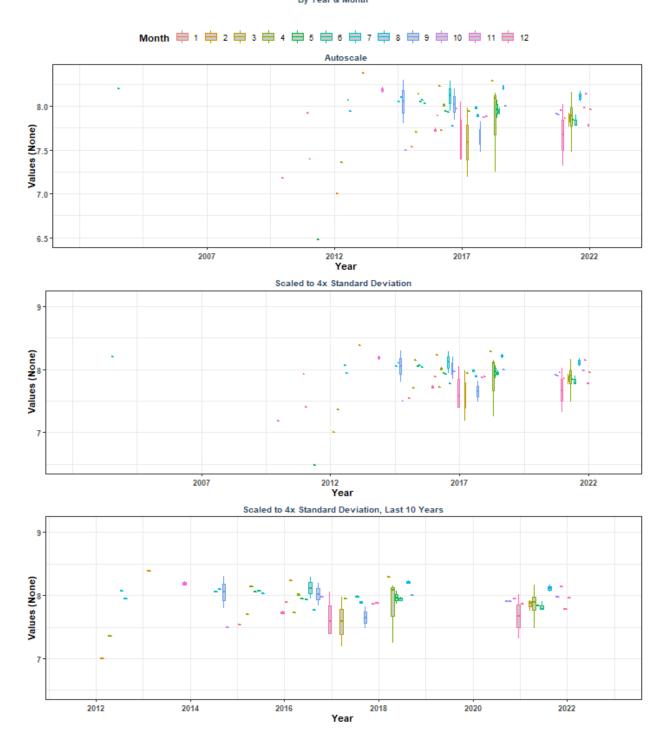
## Summary Box Plots for Cockroach Bay Aquatic Preserve By Month



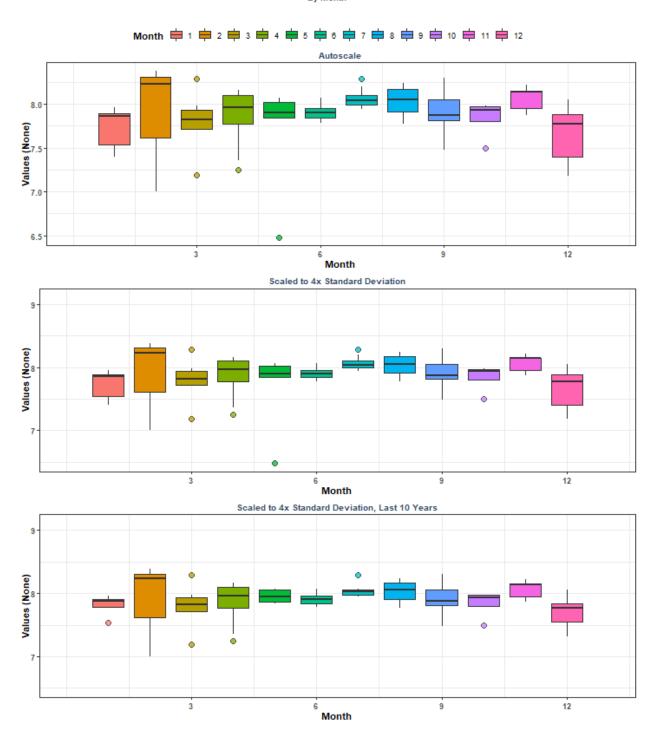
## Summary Box Plots for Estero Bay Aquatic Preserve By Year



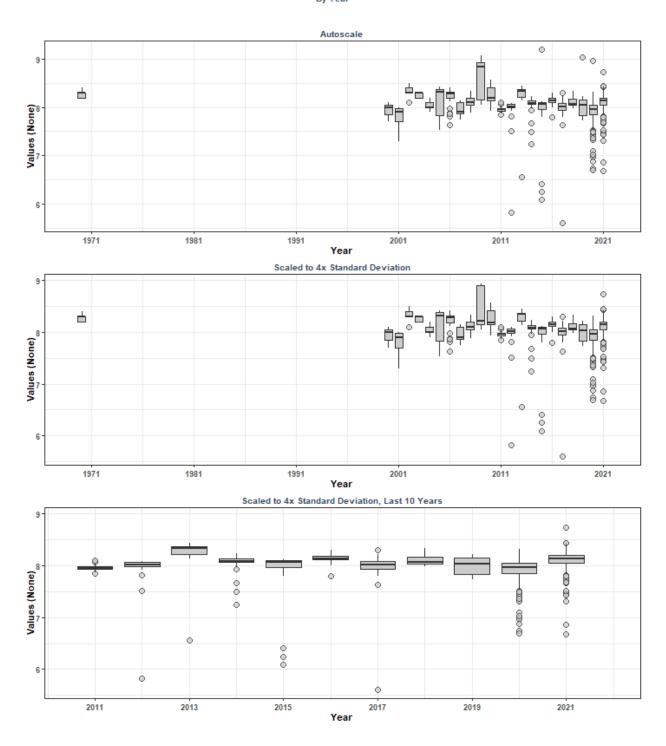
#### Summary Box Plots for Estero Bay Aquatic Preserve By Year & Month



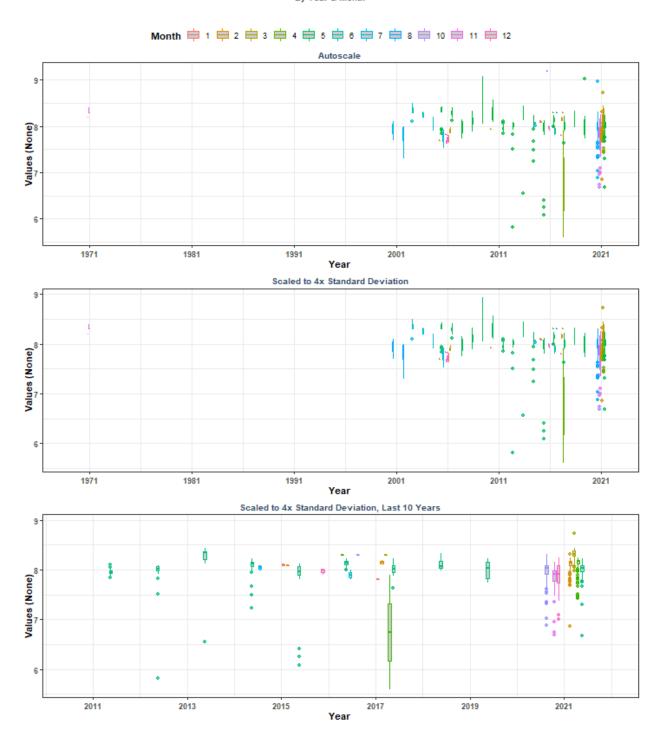
## Summary Box Plots for Estero Bay Aquatic Preserve By Month



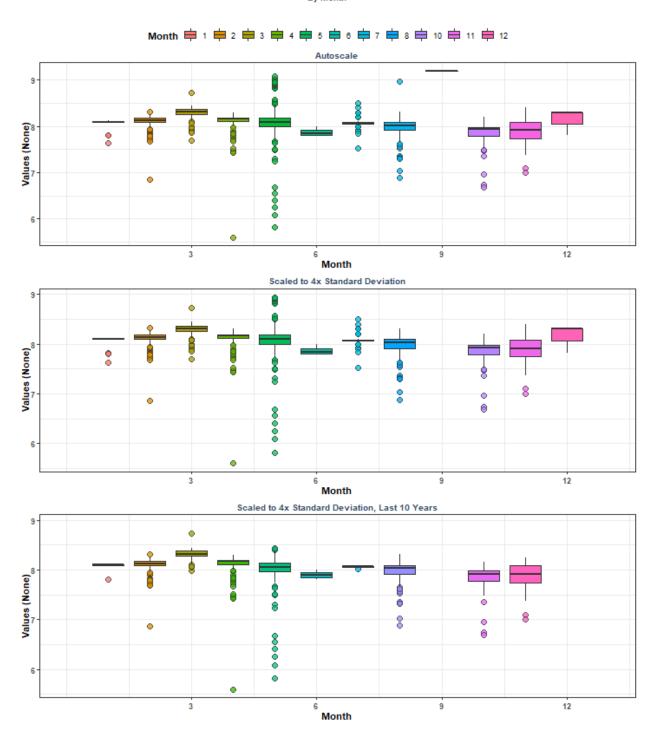
# Summary Box Plots for Florida Keys National Marine Sanctuary By Year



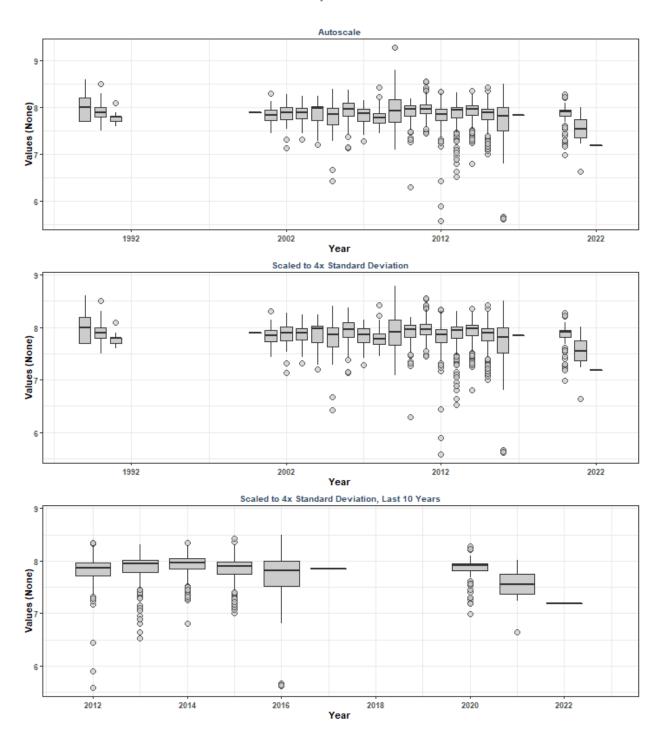
#### Summary Box Plots for Florida Keys National Marine Sanctuary By Year & Month



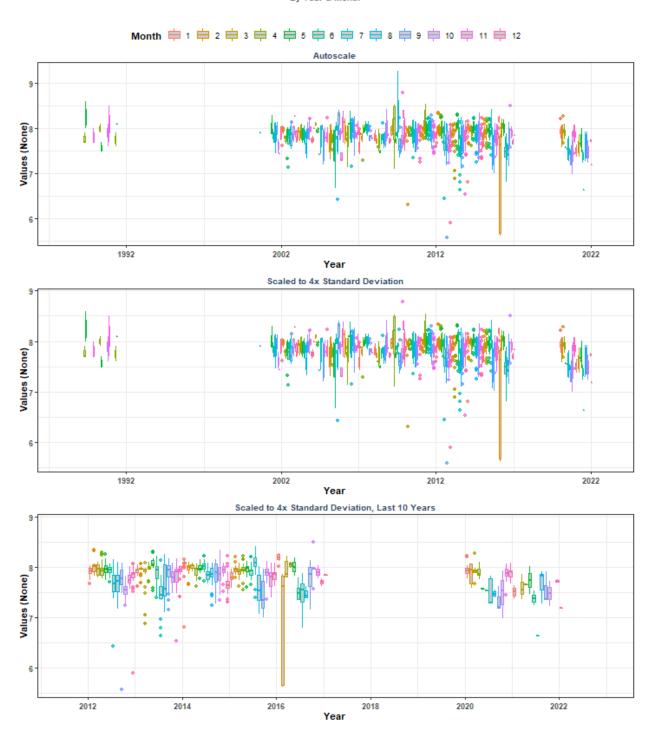
## Summary Box Plots for Florida Keys National Marine Sanctuary By Month



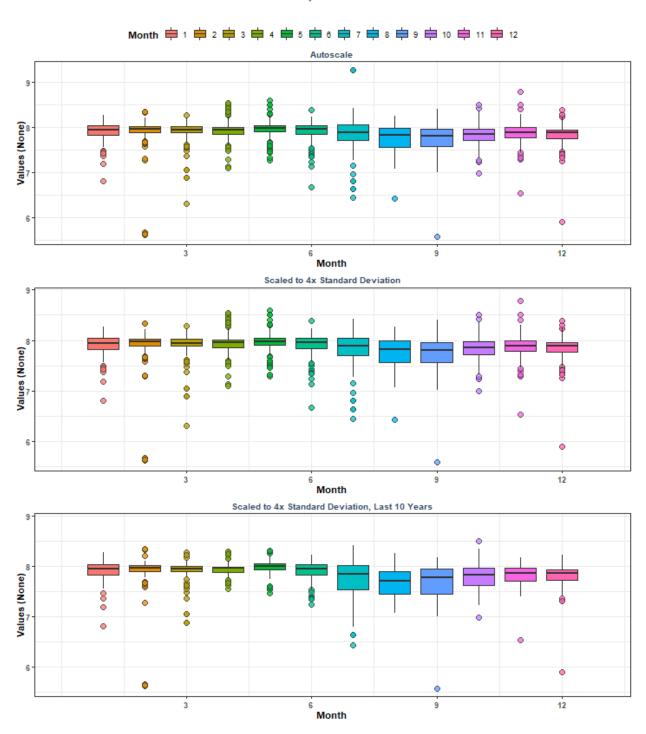
## Summary Box Plots for Gasparilla Sound-Charlotte Harbor Aquatic Preserve By Year



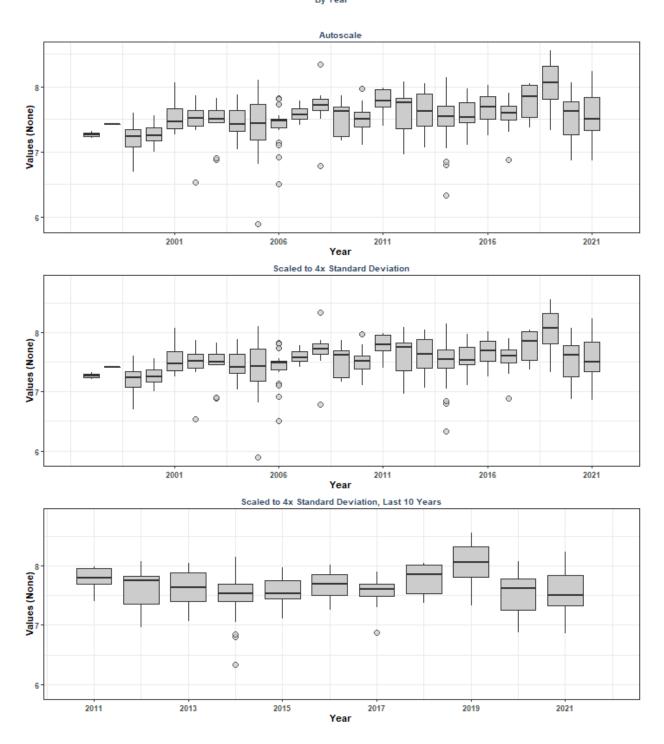
#### Summary Box Plots for Gasparilla Sound-Charlotte Harbor Aquatic Preserve By Year & Month



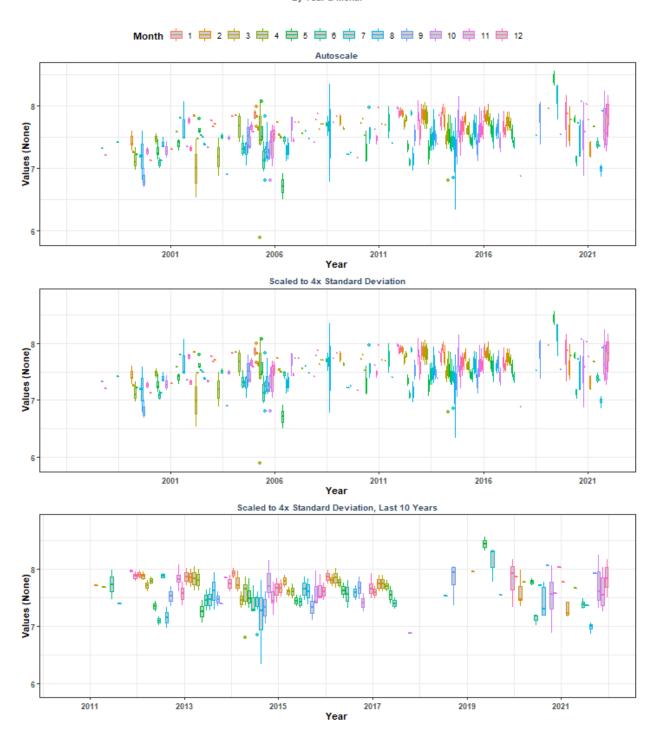
## Summary Box Plots for Gasparilla Sound-Charlotte Harbor Aquatic Preserve By Month



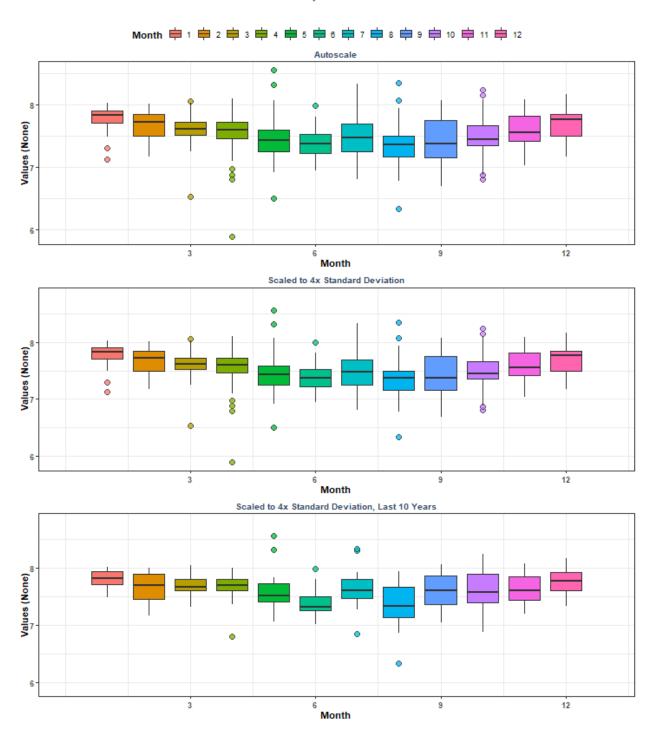
## Summary Box Plots for Guana River Marsh Aquatic Preserve By Year



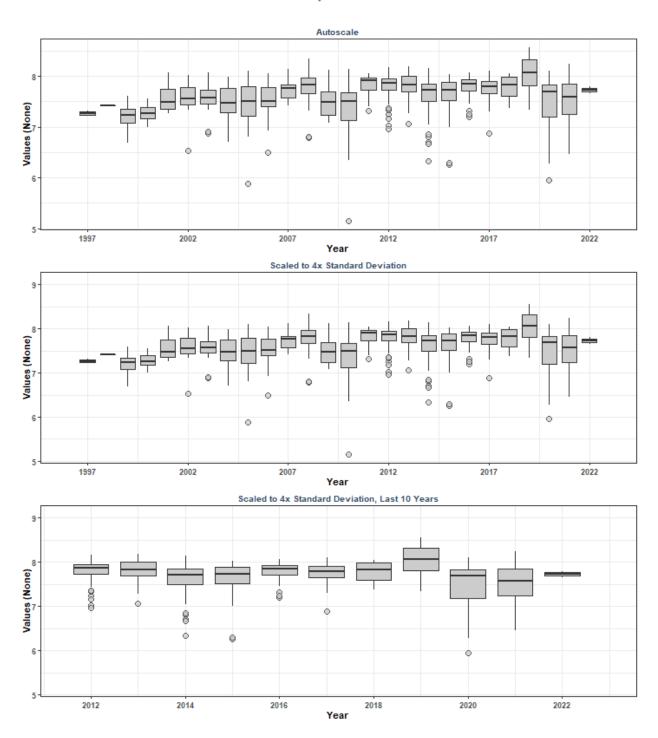
## Summary Box Plots for Guana River Marsh Aquatic Preserve By Year & Month



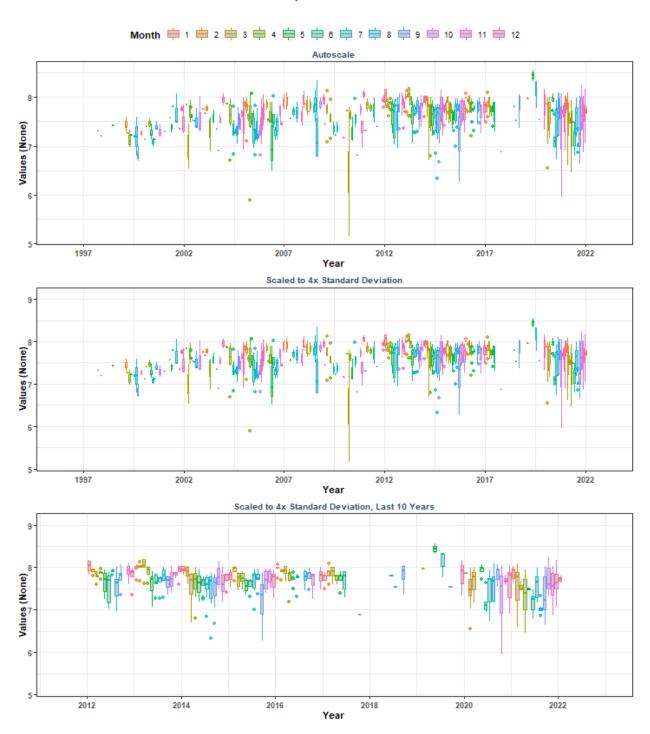
## Summary Box Plots for Guana River Marsh Aquatic Preserve By Month



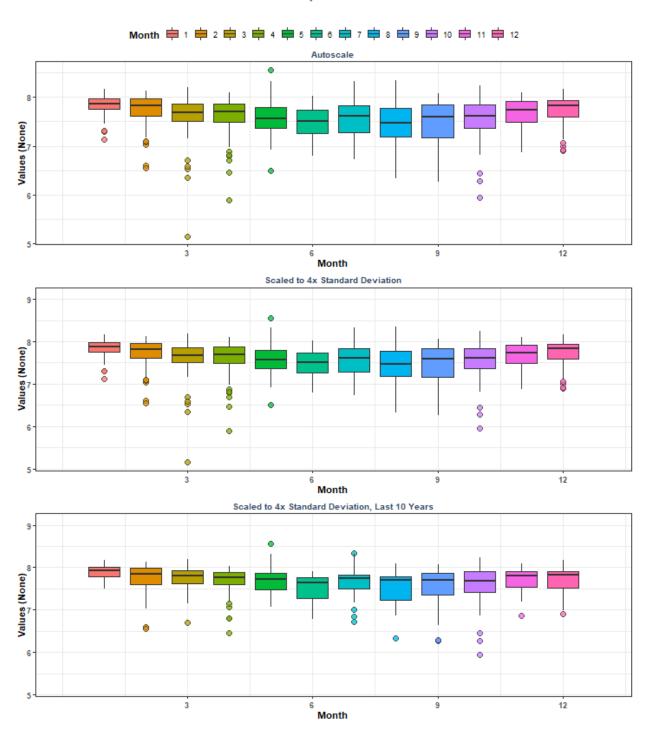
## Summary Box Plots for Guana Tolomato Matanzas National Estuarine Research Reserve $$_{\rm By\ Year}$$



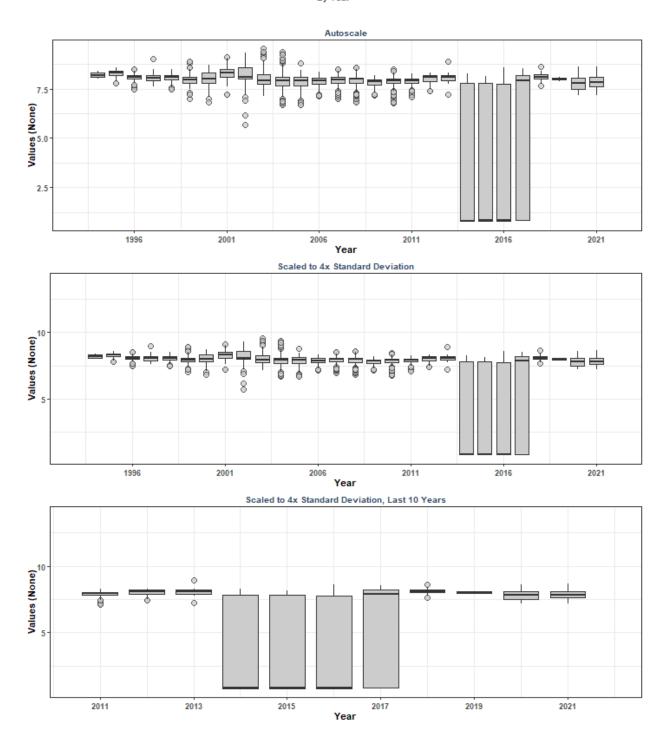
#### Summary Box Plots for Guana Tolomato Matanzas National Estuarine Research Reserve By Year & Month



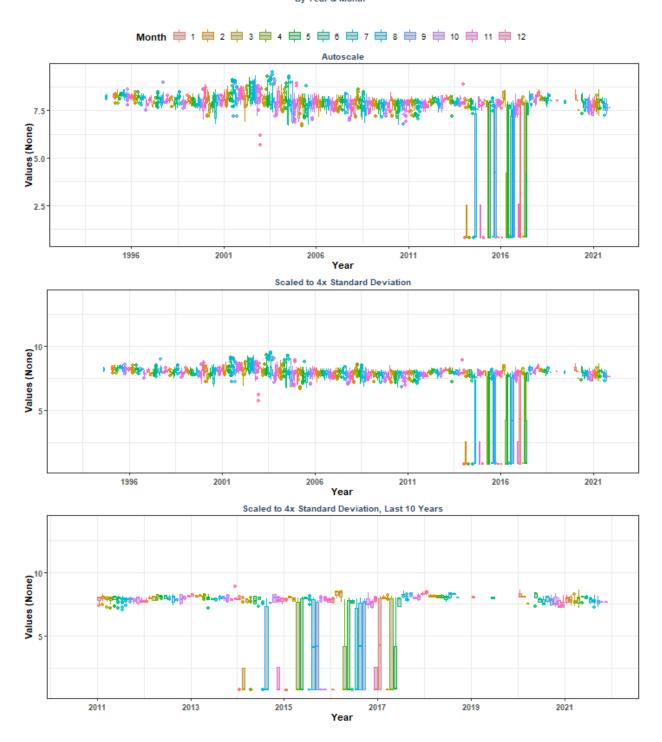
#### Summary Box Plots for Guana Tolomato Matanzas National Estuarine Research Reserve By Month



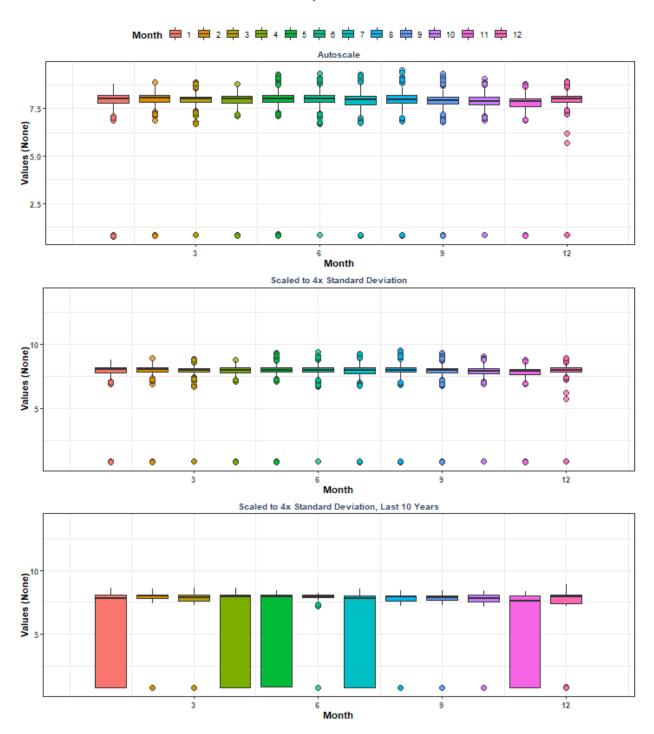
## Summary Box Plots for Indian River-Malabar to Vero Beach Aquatic Preserve $$_{\rm By\ Year}$$



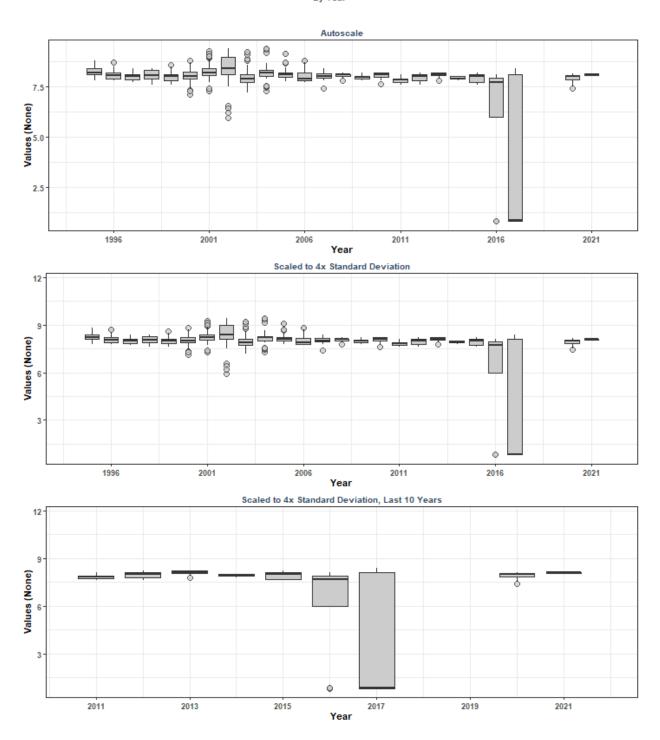
#### Summary Box Plots for Indian River-Malabar to Vero Beach Aquatic Preserve By Year & Month



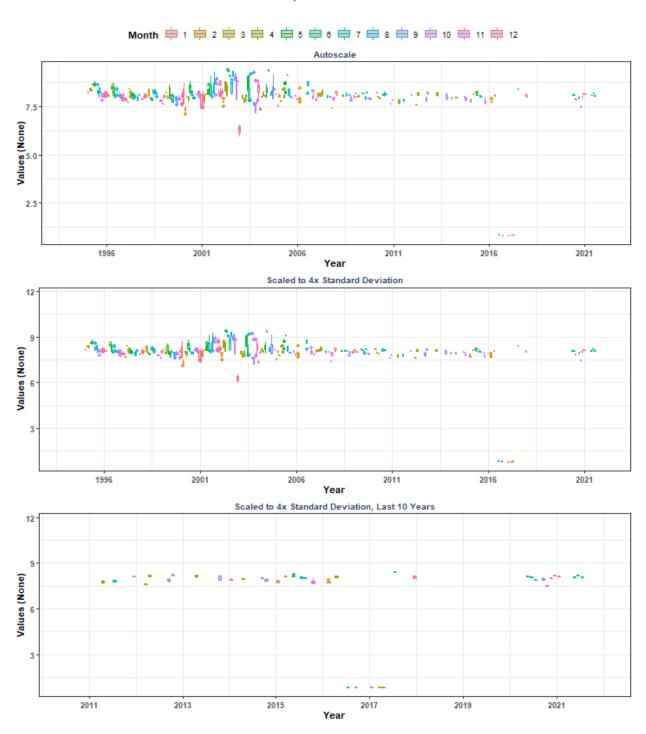
## Summary Box Plots for Indian River-Malabar to Vero Beach Aquatic Preserve By Month



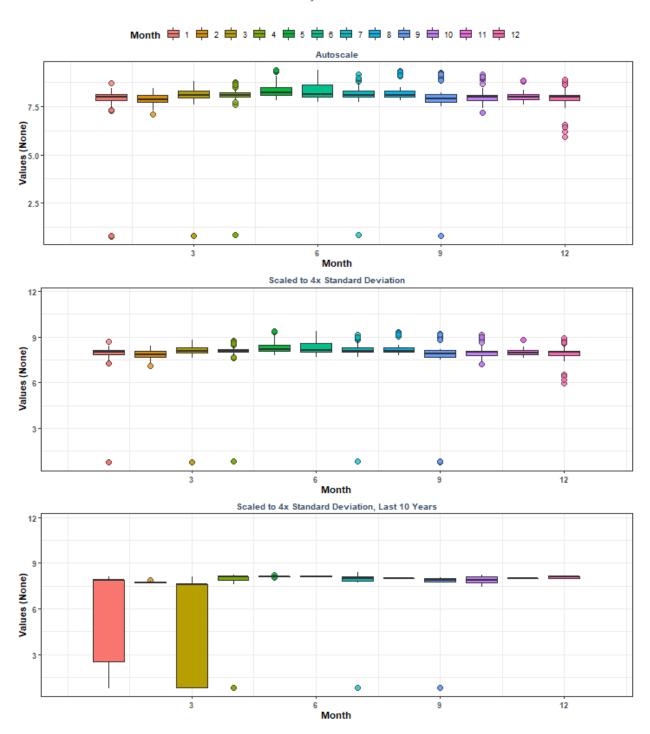
## Summary Box Plots for Indian River-Vero Beach to Ft. Pierce Aquatic Preserve $$_{\rm By\ Year}$$



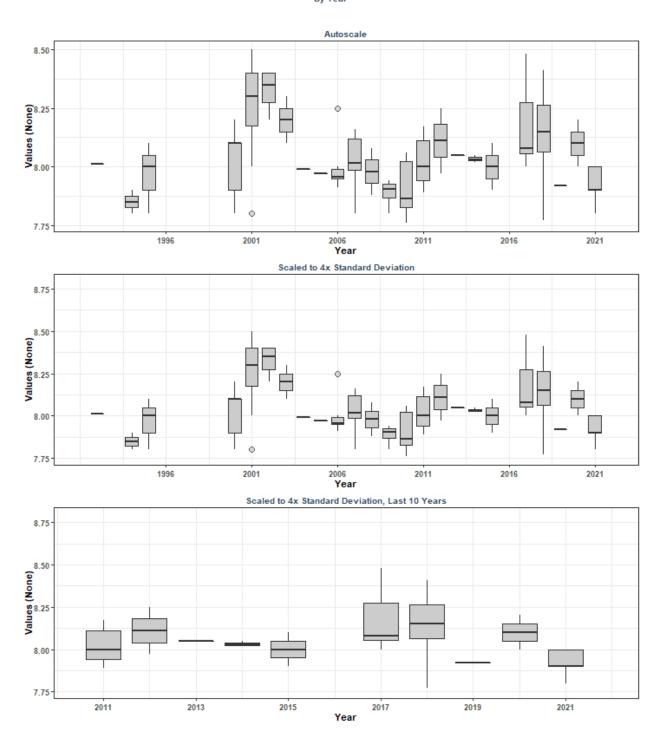
#### Summary Box Plots for Indian River-Vero Beach to Ft. Pierce Aquatic Preserve By Year & Month



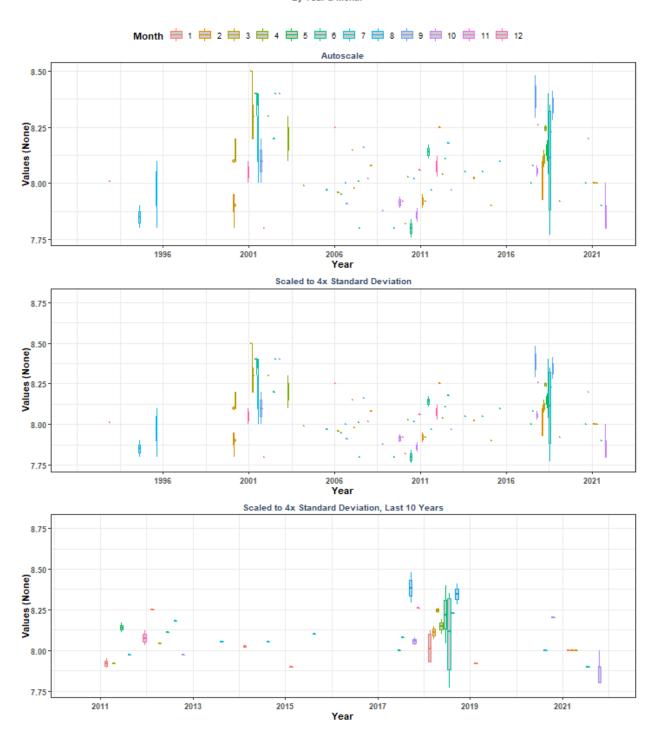
## Summary Box Plots for Indian River-Vero Beach to Ft. Pierce Aquatic Preserve $${\rm By\ Month}$$



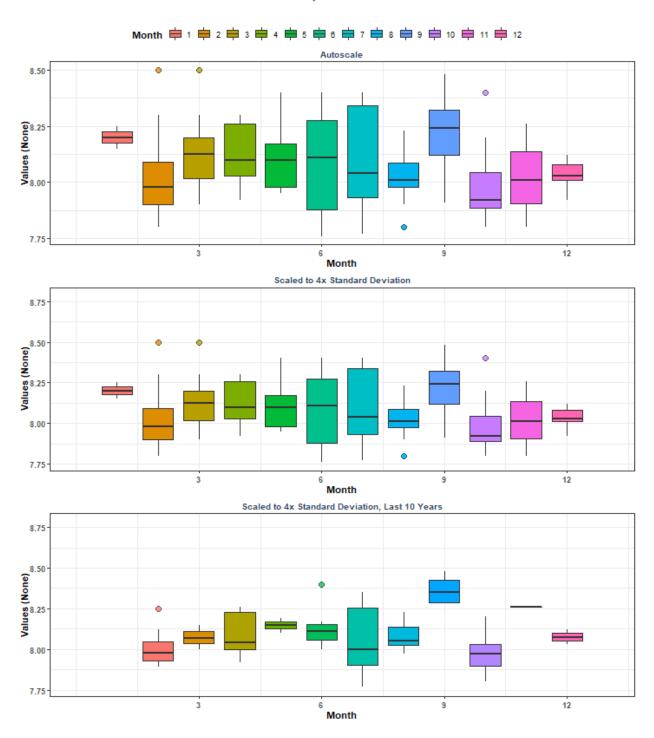
## Summary Box Plots for Jensen Beach to Jupiter Inlet Aquatic Preserve



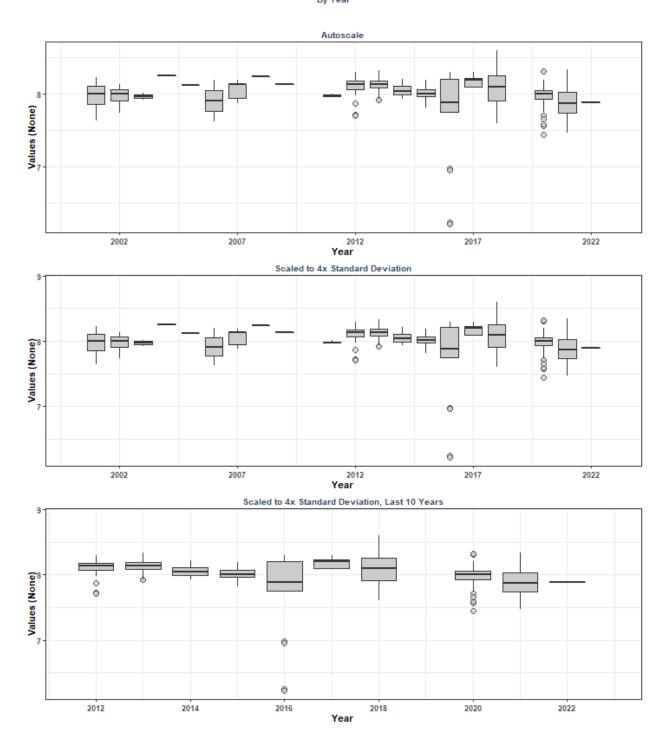
## Summary Box Plots for Jensen Beach to Jupiter Inlet Aquatic Preserve By Year & Month



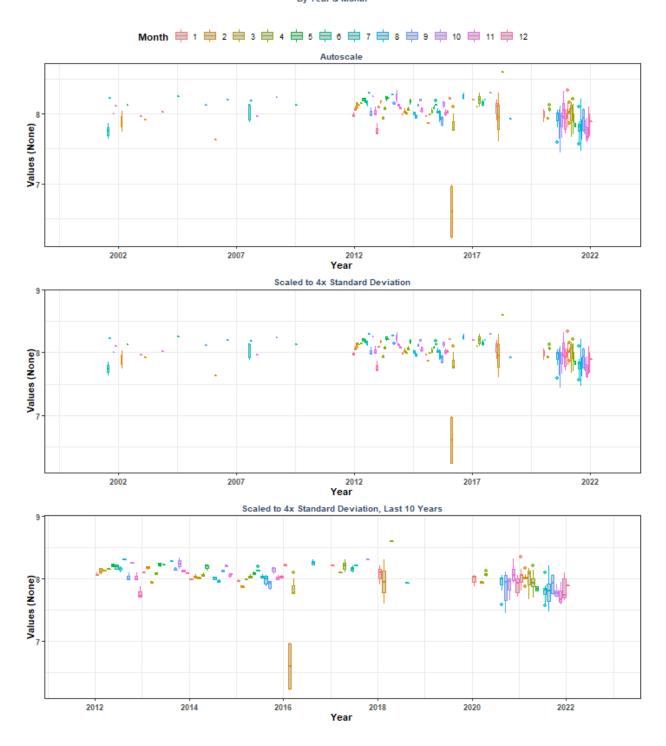
## Summary Box Plots for Jensen Beach to Jupiter Inlet Aquatic Preserve By Month



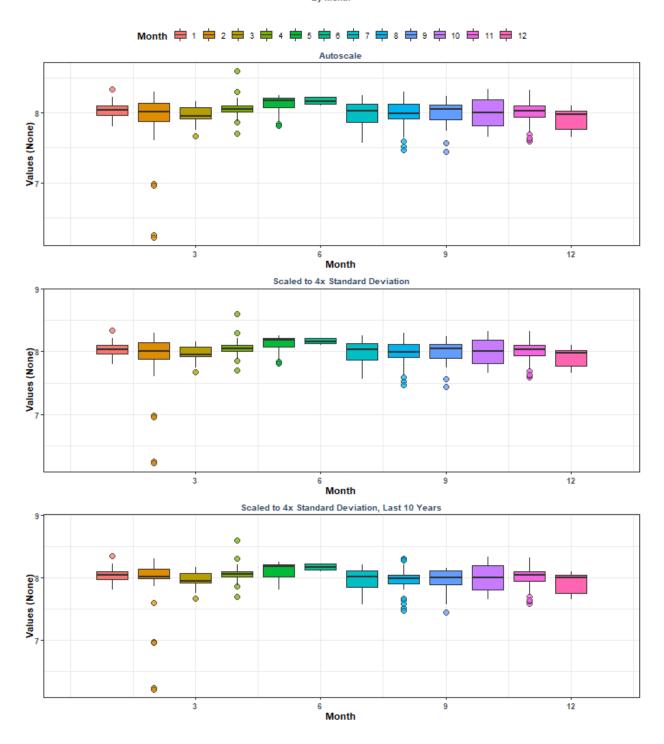
## Summary Box Plots for Lemon Bay Aquatic Preserve By Year



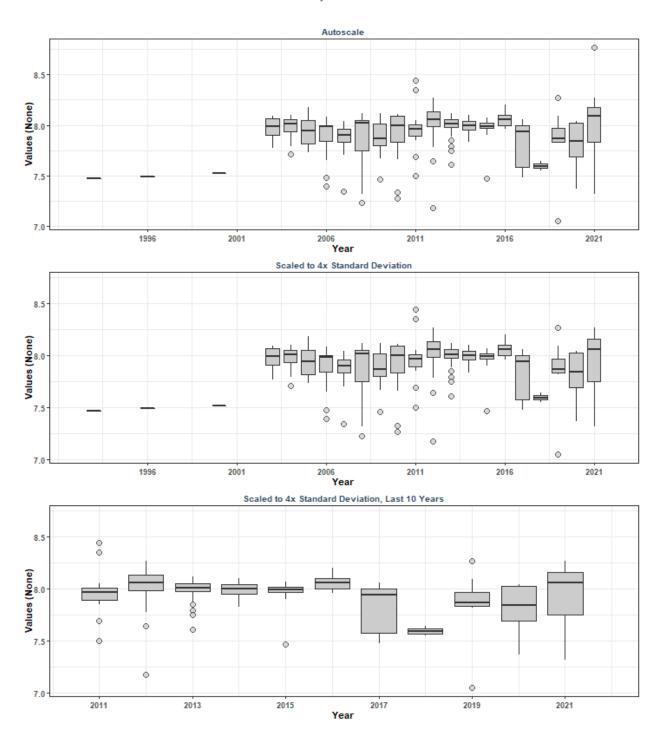
## Summary Box Plots for Lemon Bay Aquatic Preserve By Year & Month



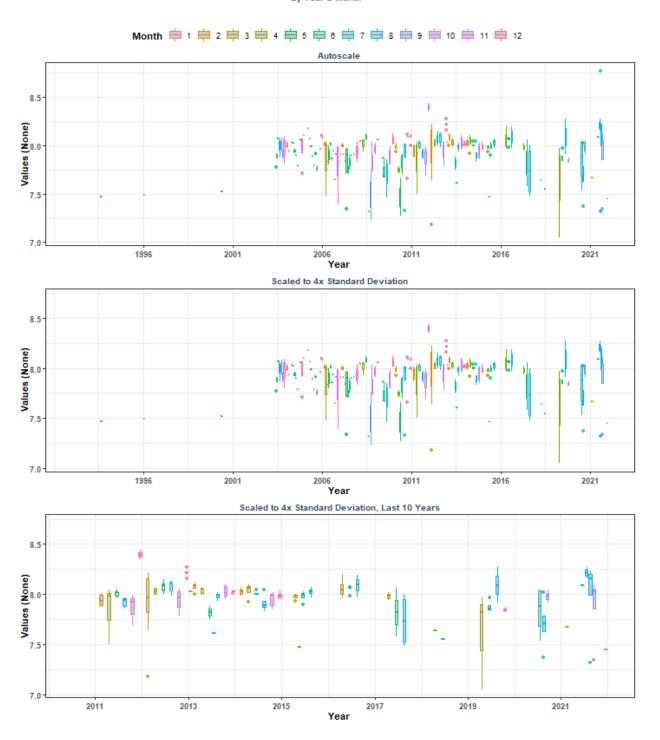
## Summary Box Plots for Lemon Bay Aquatic Preserve By Month



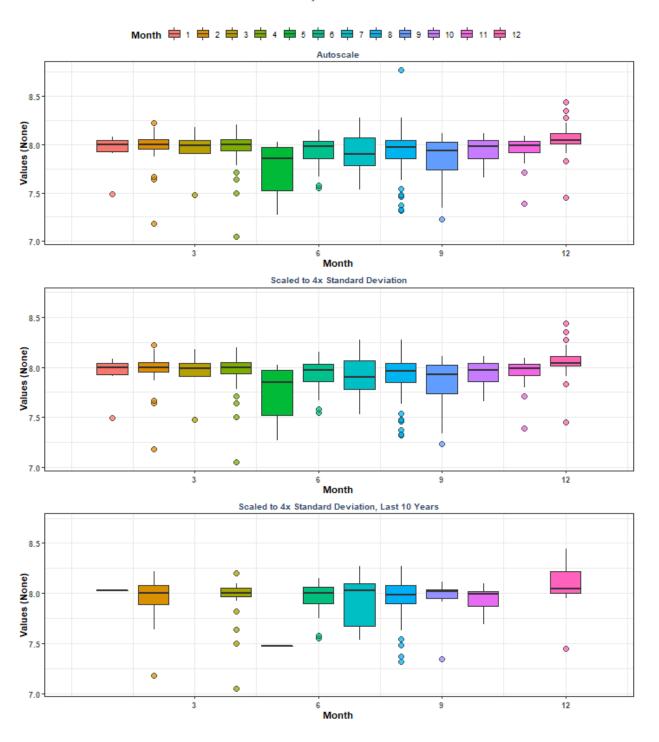
## Summary Box Plots for Loxahatchee River-Lake Worth Creek Aquatic Preserve $_{\mbox{\footnotesize By Year}}$



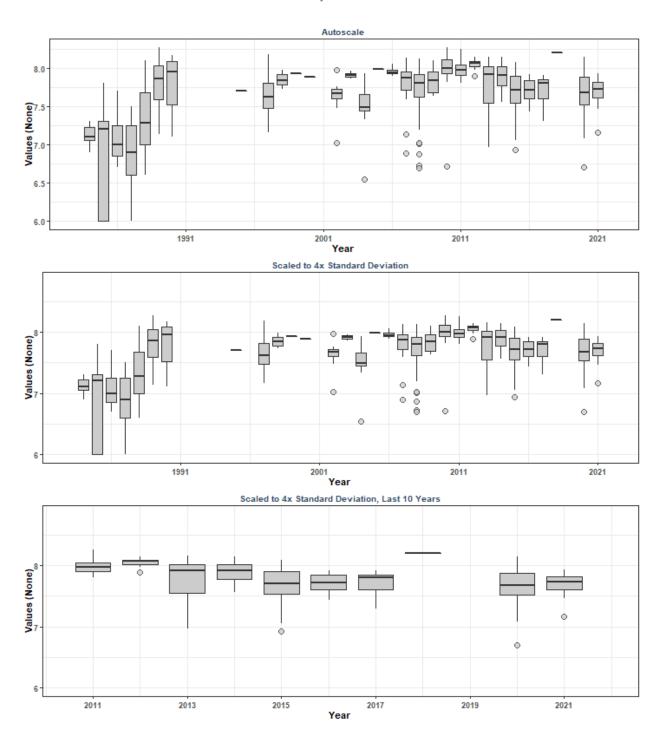
#### Summary Box Plots for Loxahatchee River-Lake Worth Creek Aquatic Preserve By Year & Month



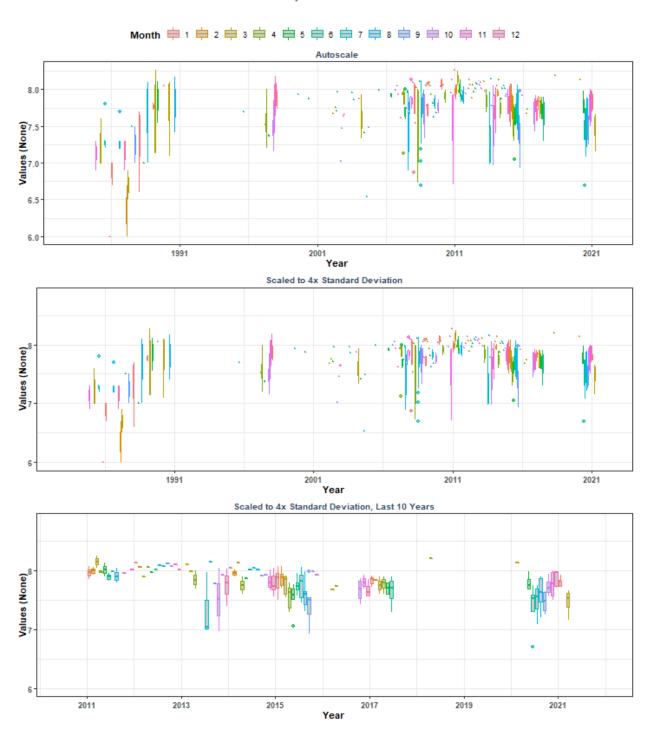
## Summary Box Plots for Loxahatchee River-Lake Worth Creek Aquatic Preserve By Month



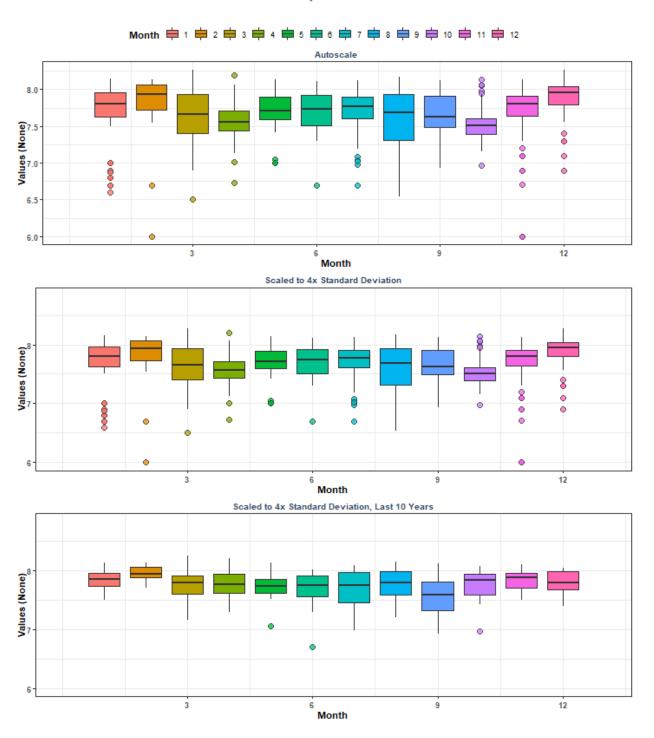
## Summary Box Plots for Nassau River-St. Johns River Marshes Aquatic Preserve $_{\rm By\,Year}$



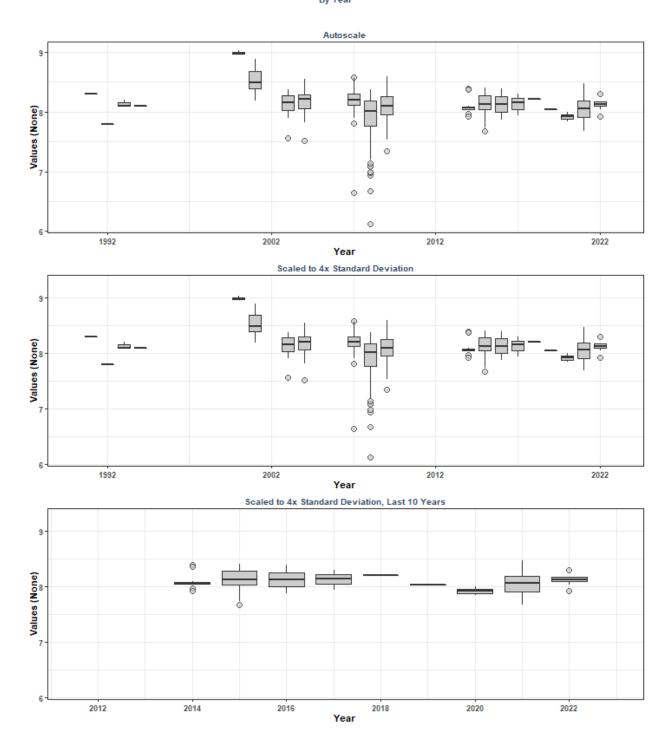
## Summary Box Plots for Nassau River-St. Johns River Marshes Aquatic Preserve By Year & Month



## Summary Box Plots for Nassau River-St. Johns River Marshes Aquatic Preserve

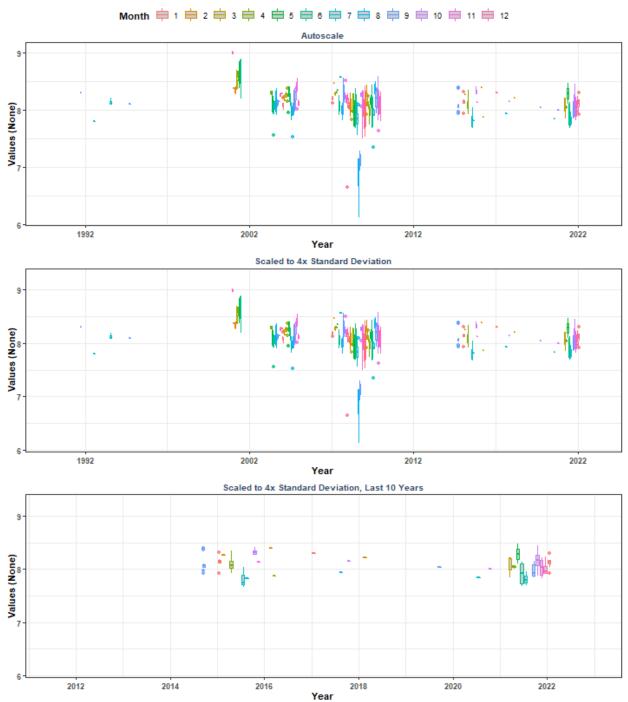


## Summary Box Plots for Nature Coast Aquatic Preserve By Year

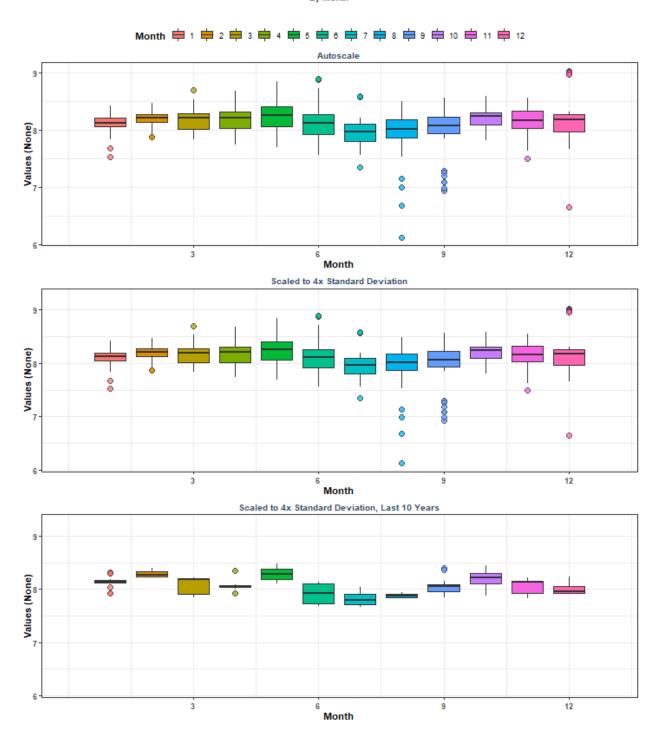


## Summary Box Plots for Nature Coast Aquatic Preserve By Year & Month

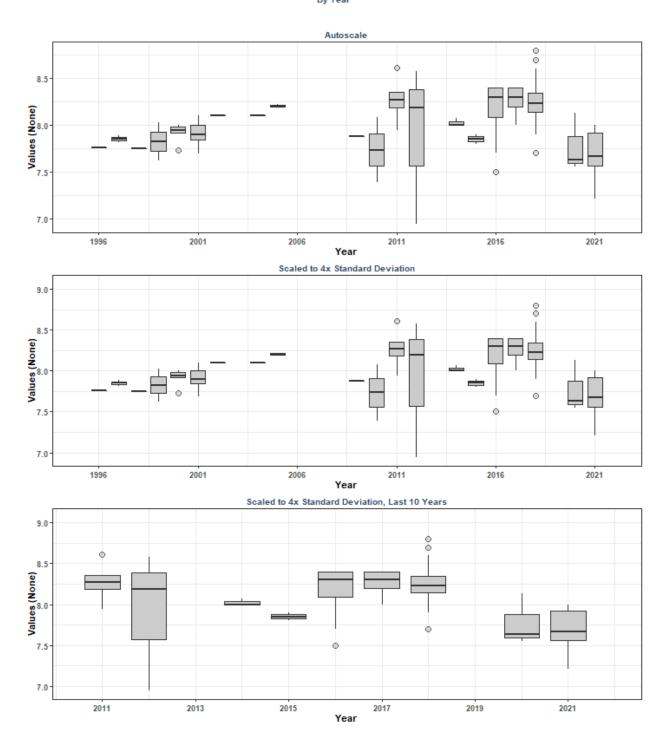




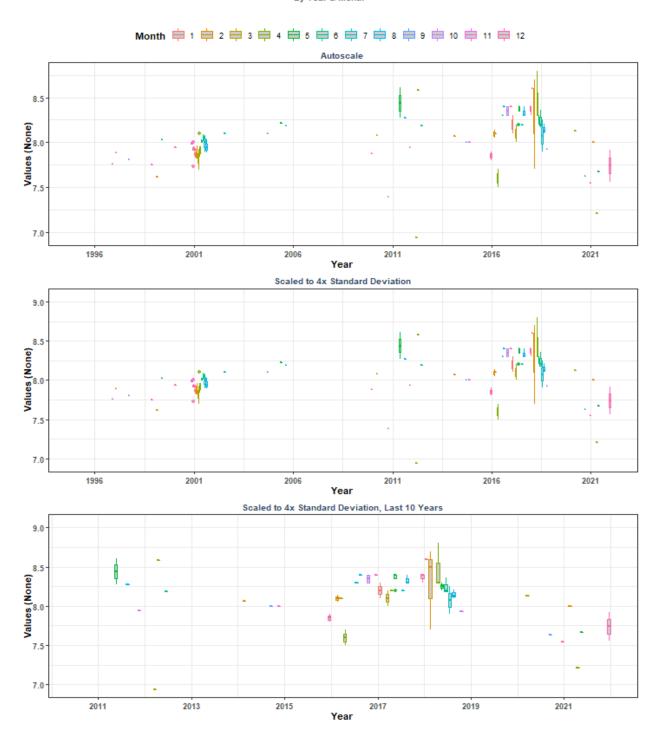
## Summary Box Plots for Nature Coast Aquatic Preserve By Month



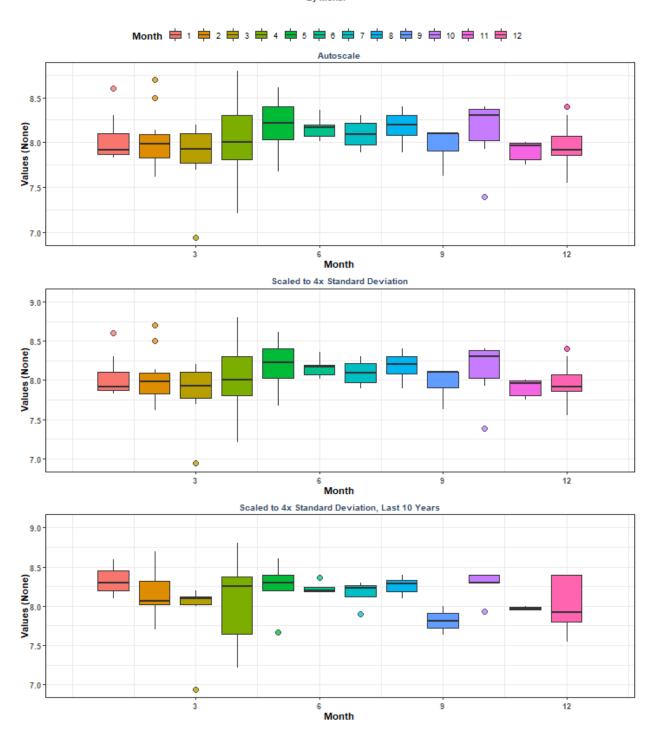
## Summary Box Plots for Pine Island Sound Aquatic Preserve $$_{\rm By\ Year}$$



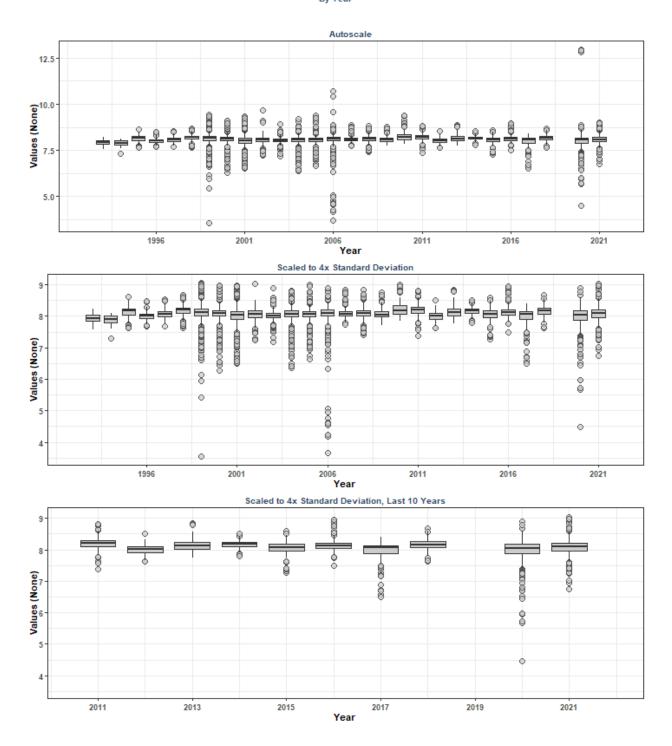
### Summary Box Plots for Pine Island Sound Aquatic Preserve By Year & Month



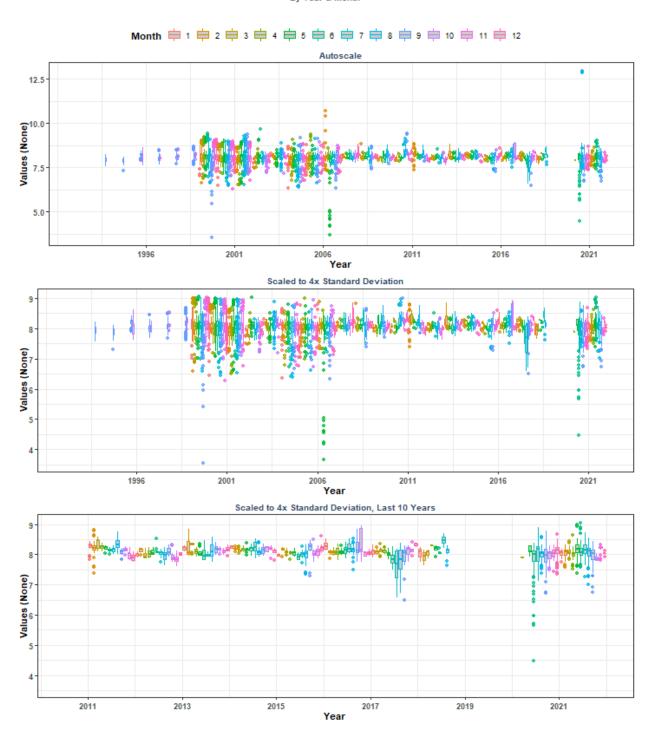
## Summary Box Plots for Pine Island Sound Aquatic Preserve By Month



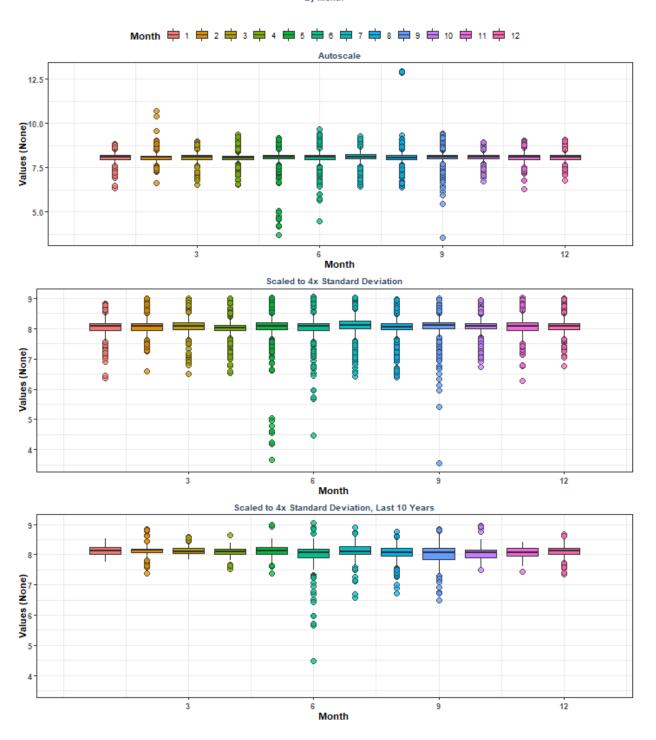
### Summary Box Plots for Pinellas County Aquatic Preserve By Year



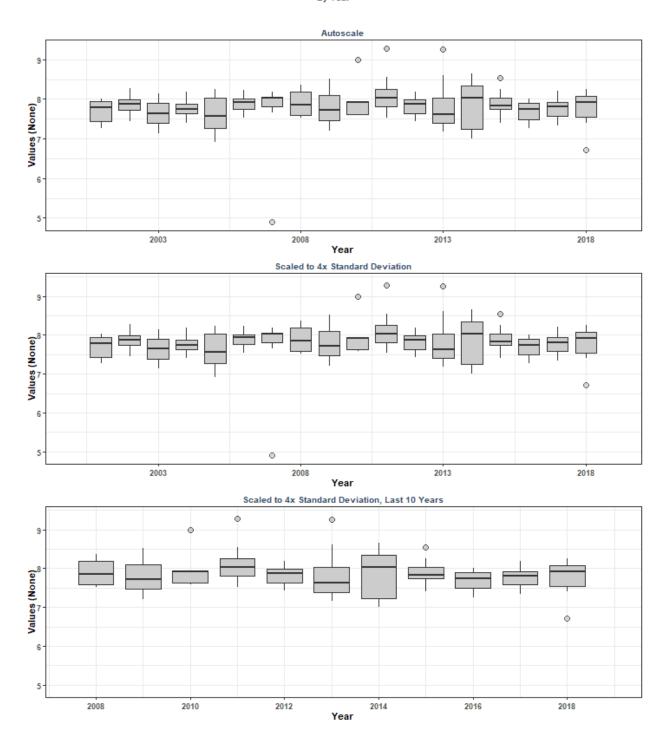
### Summary Box Plots for Pinellas County Aquatic Preserve By Year & Month



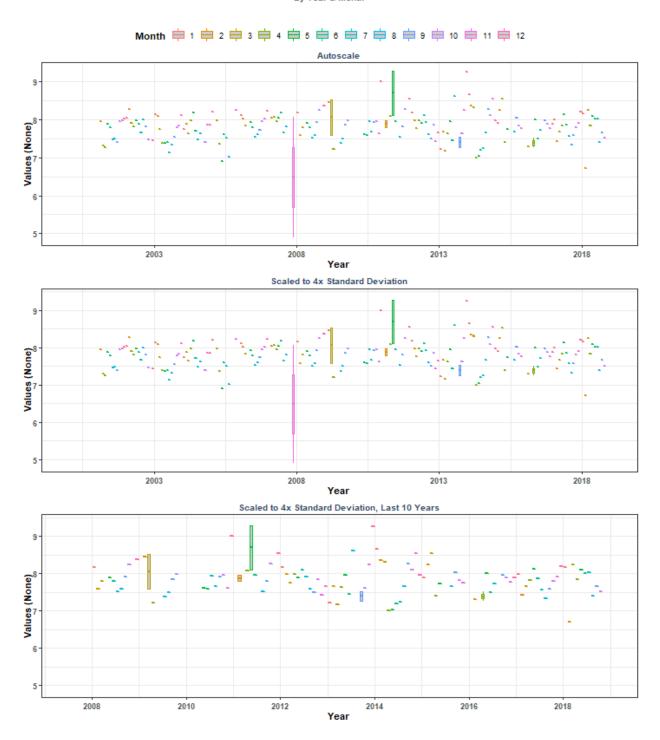
### Summary Box Plots for Pinellas County Aquatic Preserve By Month



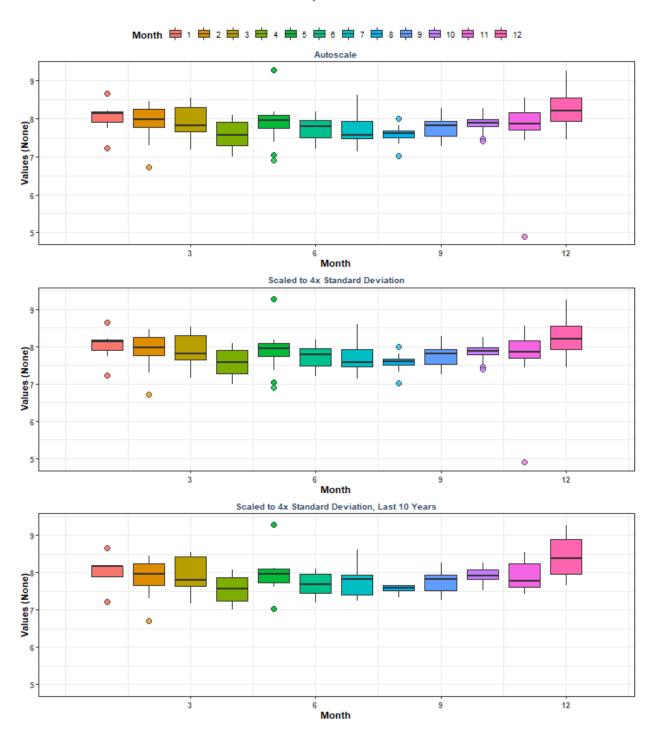
## Summary Box Plots for Rocky Bayou State Park Aquatic Preserve $_{\rm By\ Year}$



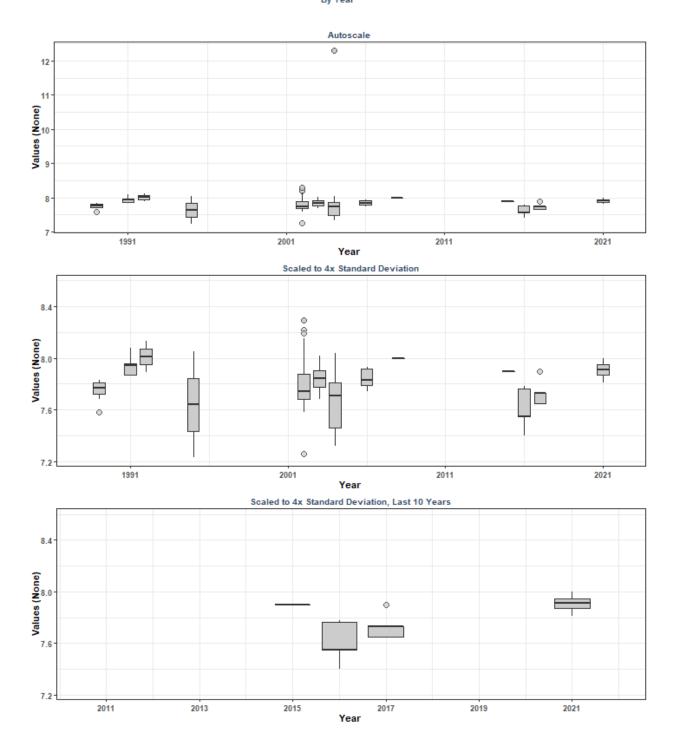
## Summary Box Plots for Rocky Bayou State Park Aquatic Preserve By Year & Month



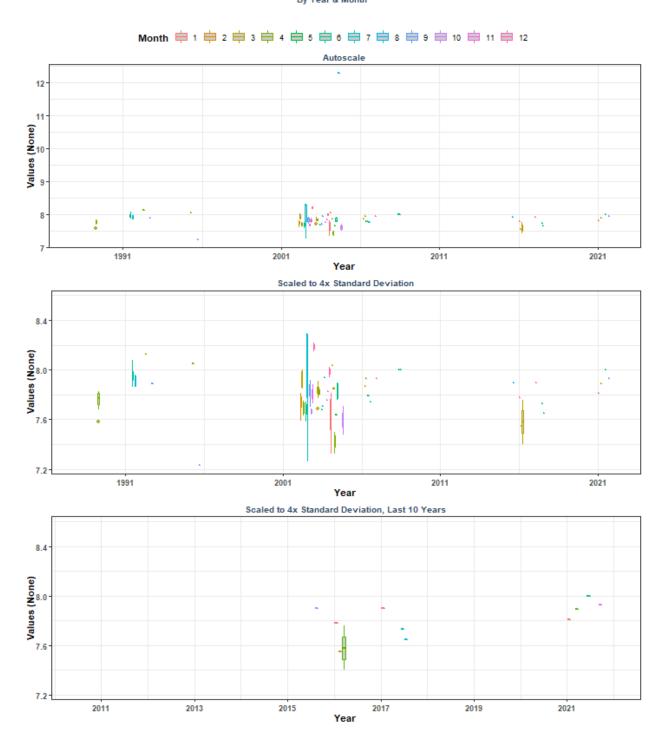
### Summary Box Plots for Rocky Bayou State Park Aquatic Preserve



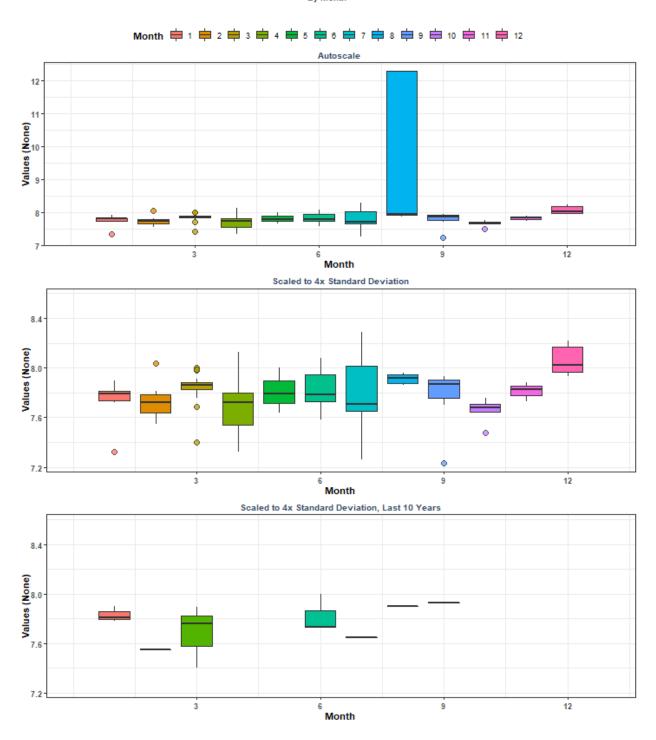
# Summary Box Plots for Rookery Bay Aquatic Preserve $$_{\rm By\ Year}$$



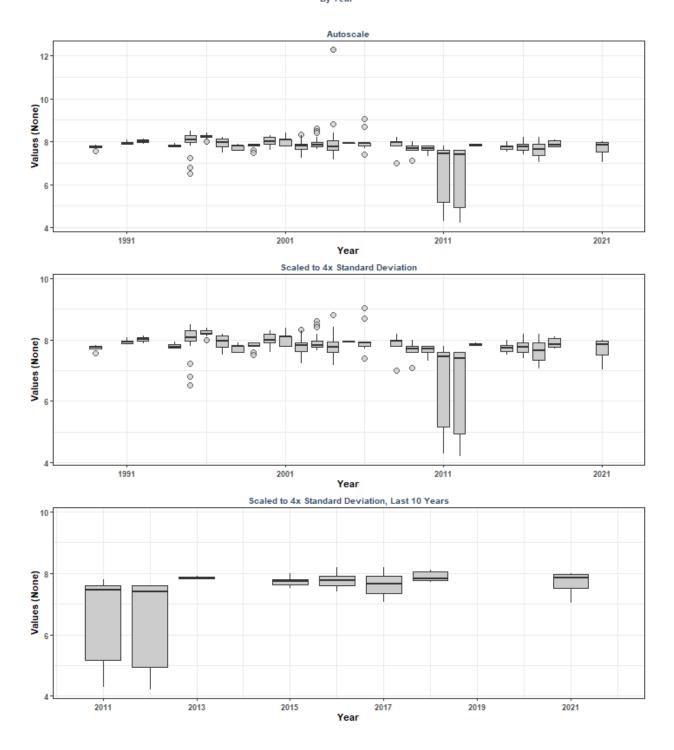
## Summary Box Plots for Rookery Bay Aquatic Preserve By Year & Month



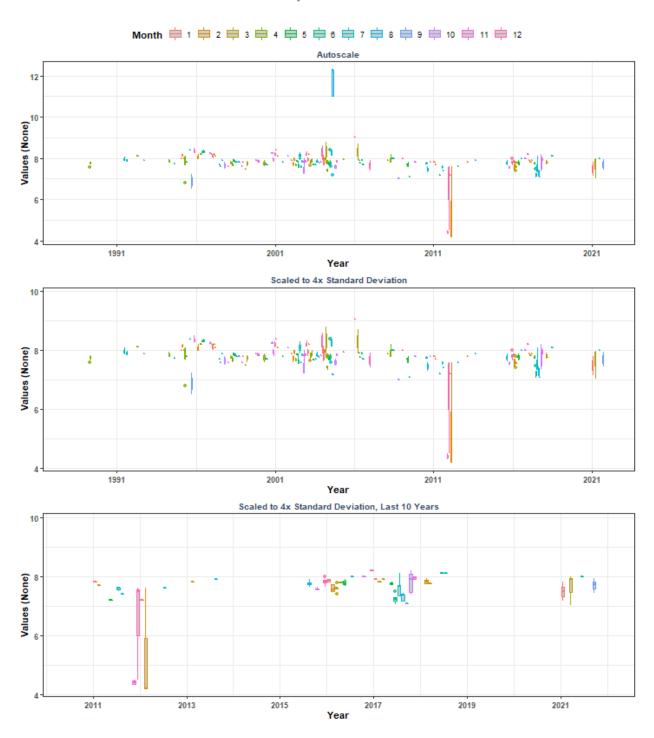
## Summary Box Plots for Rookery Bay Aquatic Preserve By Month



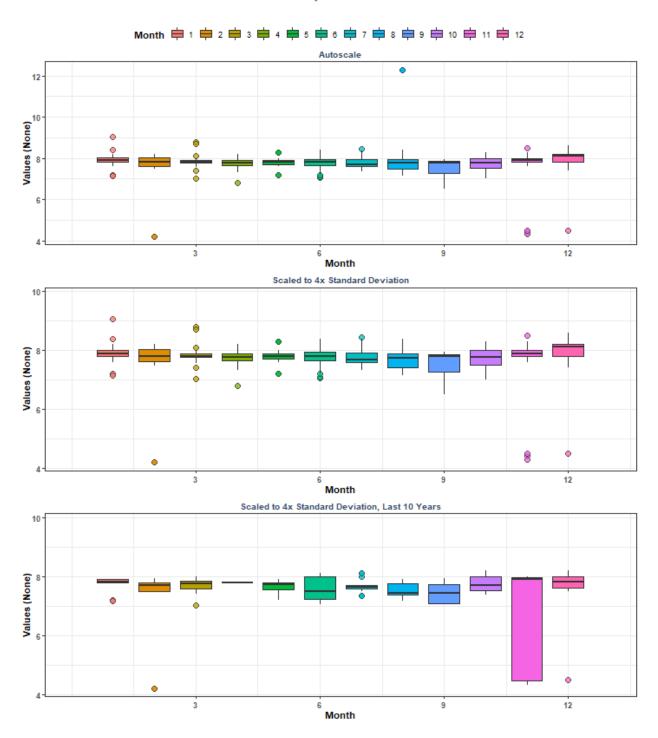
## Summary Box Plots for Rookery Bay National Estuarine Research Reserve $$_{\rm By\ Year}$$



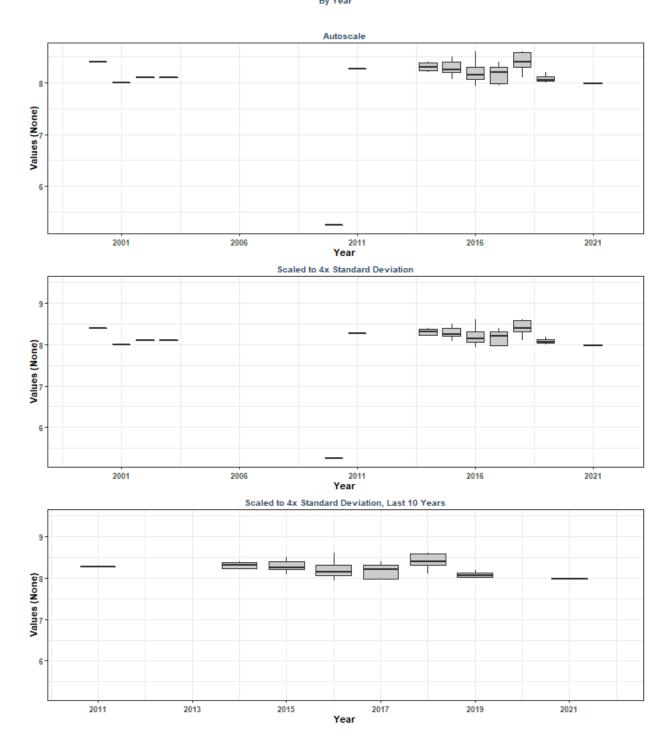
## Summary Box Plots for Rookery Bay National Estuarine Research Reserve $_{\mbox{\scriptsize By Year}~\&~\mbox{\scriptsize Month}}$



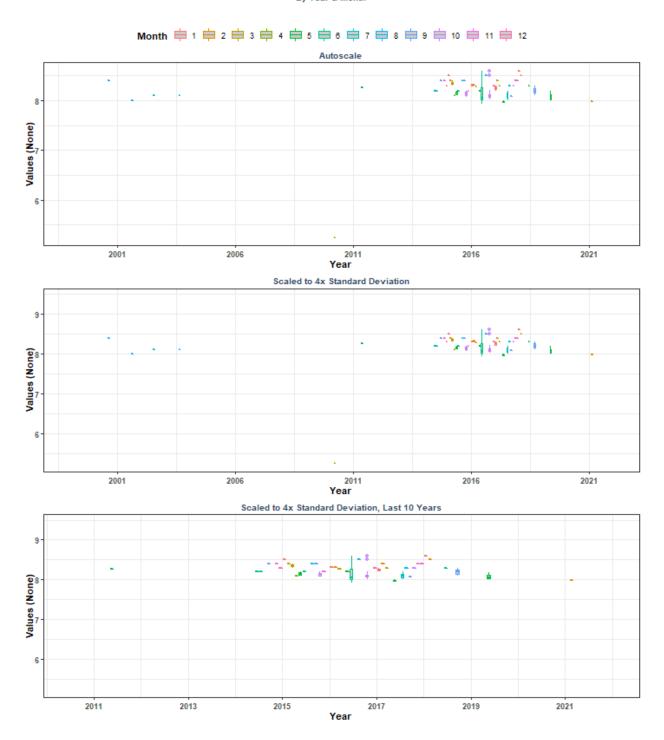
## Summary Box Plots for Rookery Bay National Estuarine Research Reserve



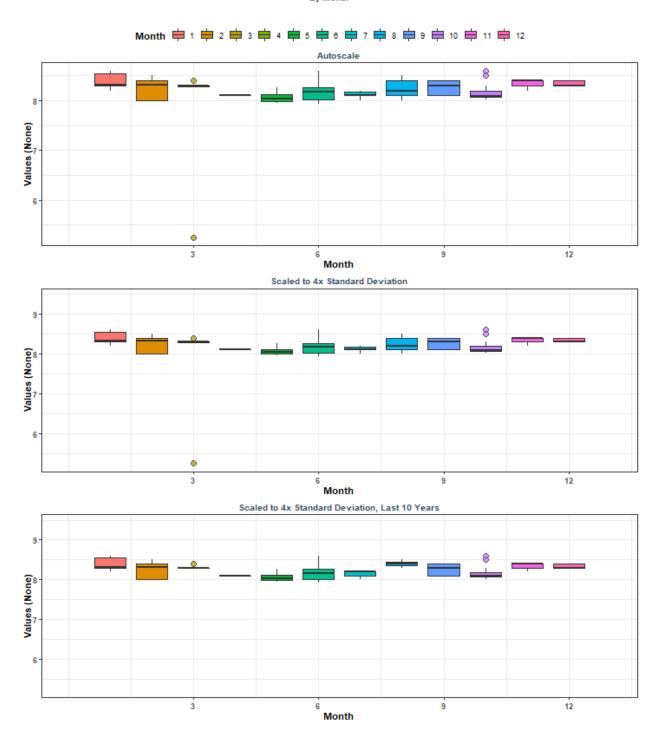
## Summary Box Plots for St. Andrews State Park Aquatic Preserve $$_{\rm By\ Year}$$



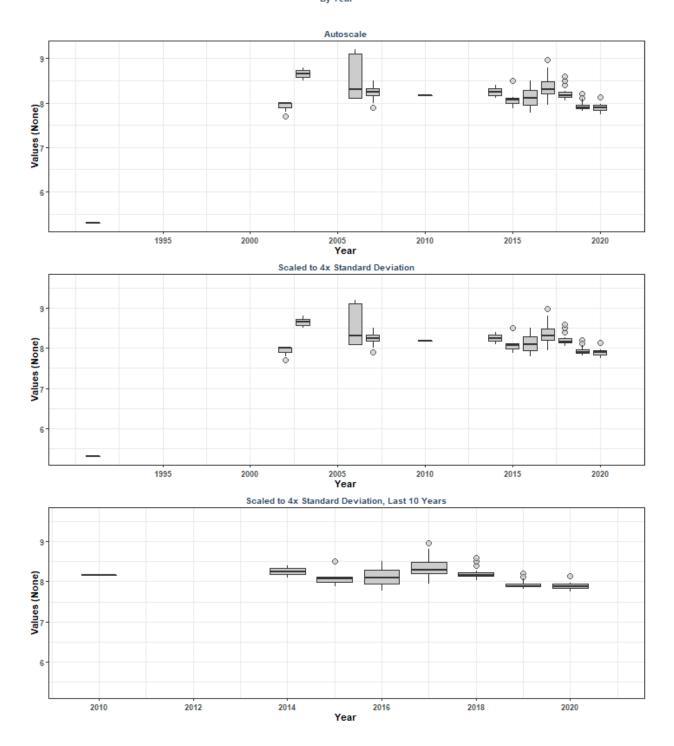
#### Summary Box Plots for St. Andrews State Park Aquatic Preserve By Year & Month



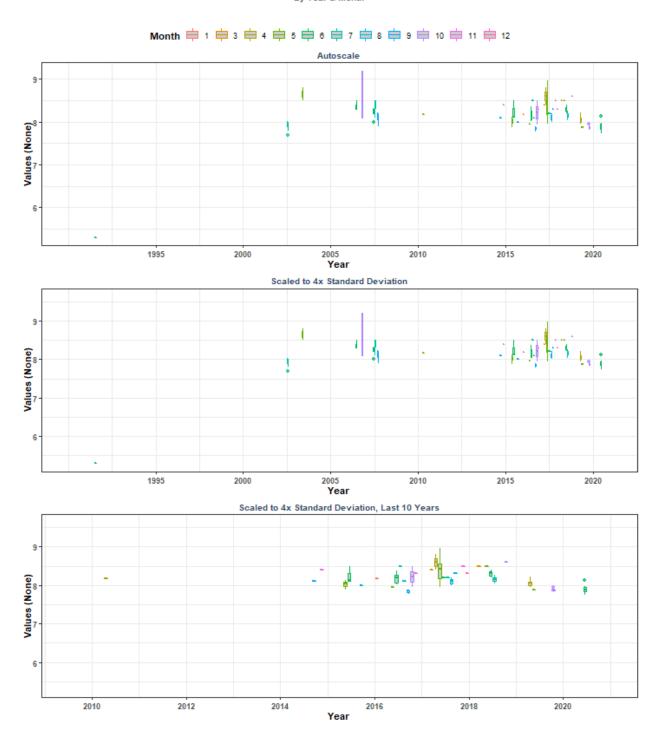
### Summary Box Plots for St. Andrews State Park Aquatic Preserve



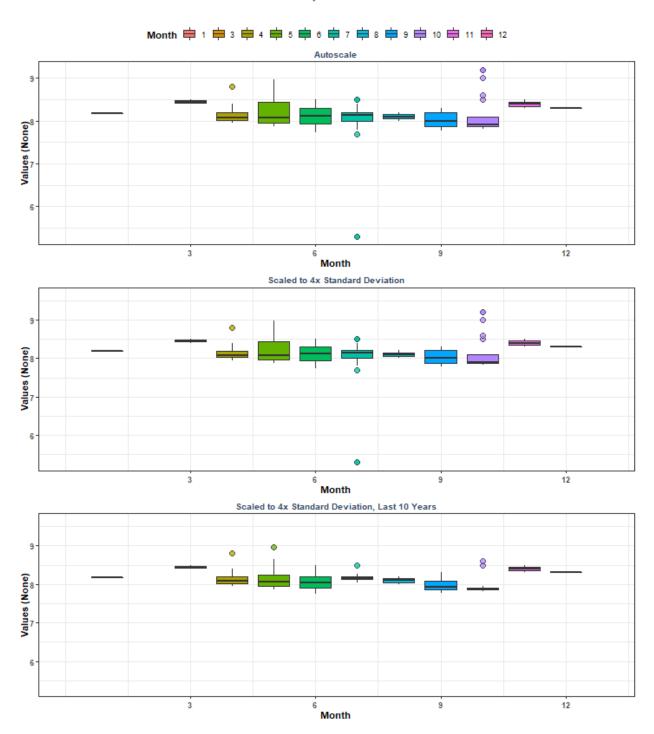
## Summary Box Plots for St. Joseph Bay Aquatic Preserve $$_{\rm By\ Year}$$



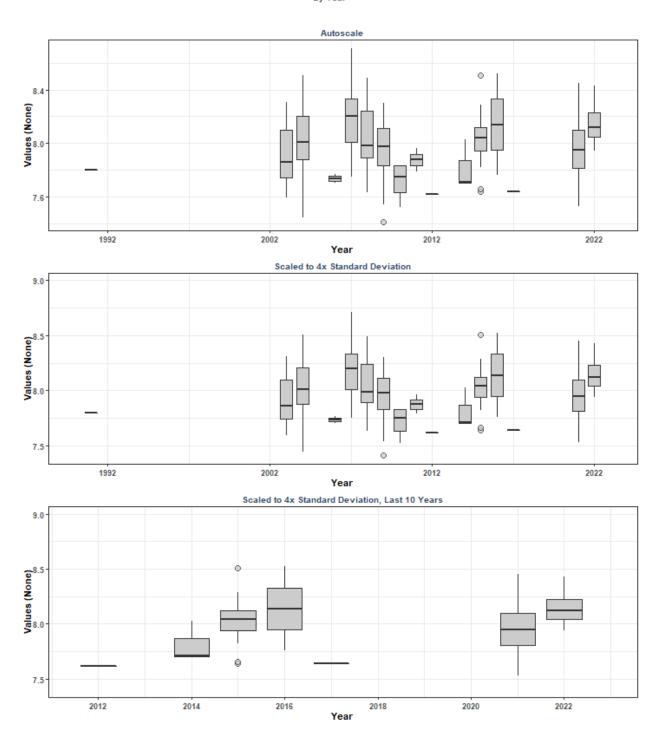
### Summary Box Plots for St. Joseph Bay Aquatic Preserve By Year & Month



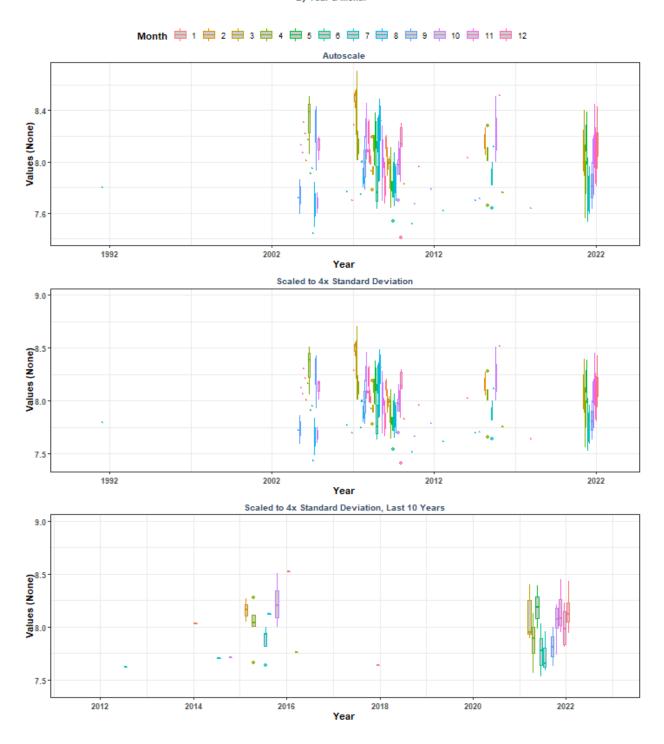
### Summary Box Plots for St. Joseph Bay Aquatic Preserve By Month



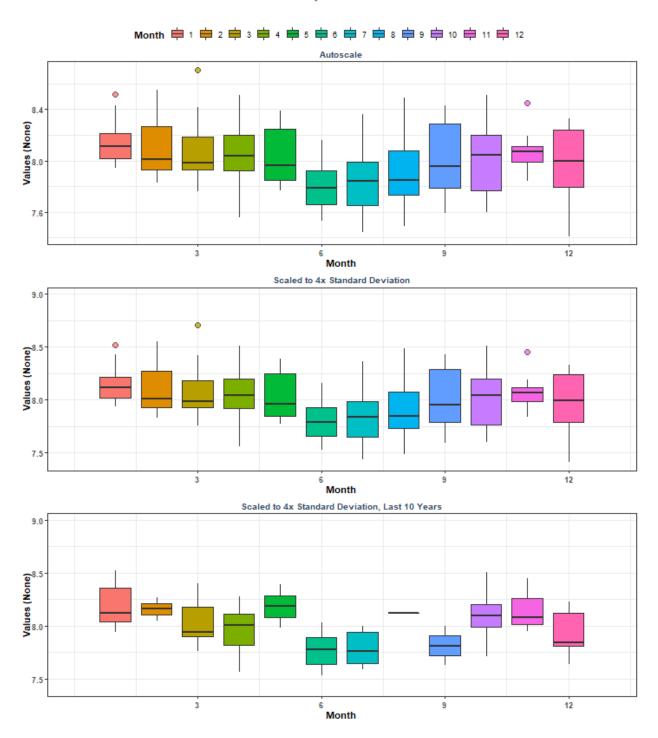
## Summary Box Plots for St. Martins Marsh Aquatic Preserve



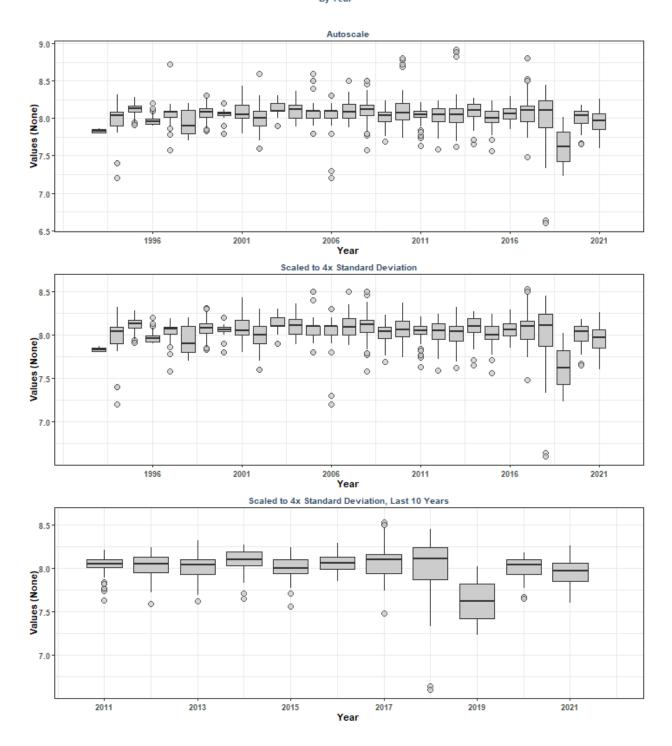
## Summary Box Plots for St. Martins Marsh Aquatic Preserve By Year & Month



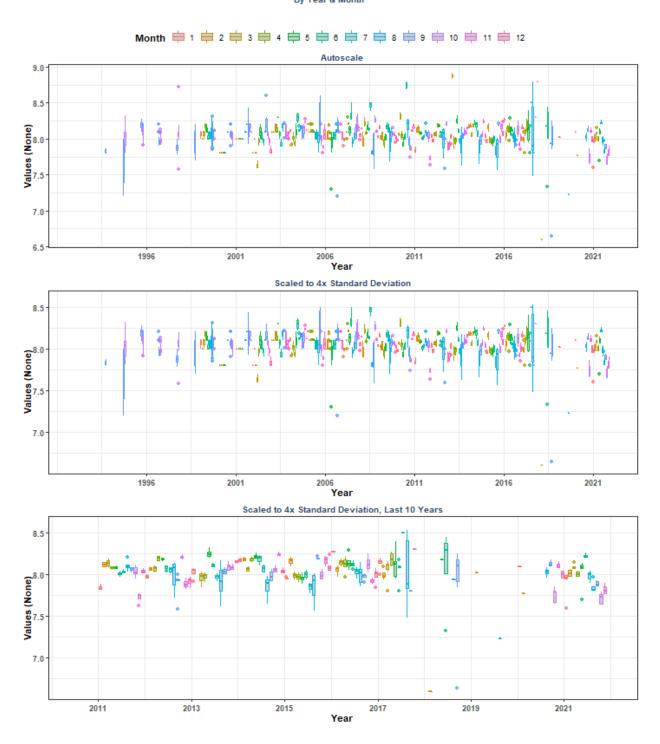
## Summary Box Plots for St. Martins Marsh Aquatic Preserve By Month



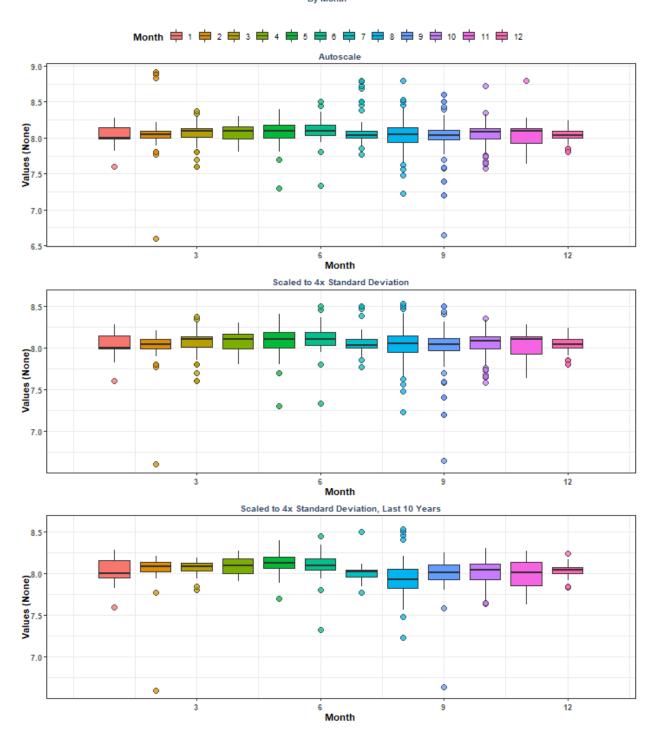
## Summary Box Plots for Terra Ceia Aquatic Preserve By Year



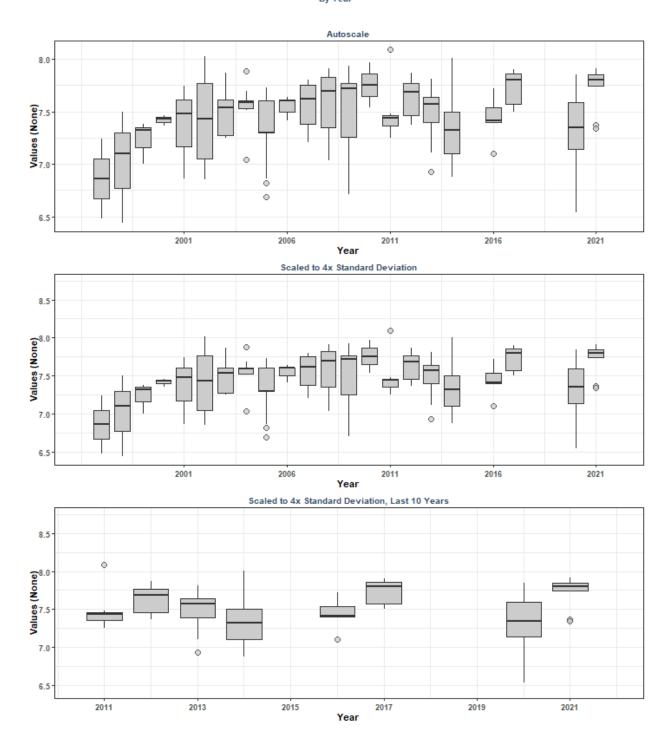
## Summary Box Plots for Terra Ceia Aquatic Preserve By Year & Month



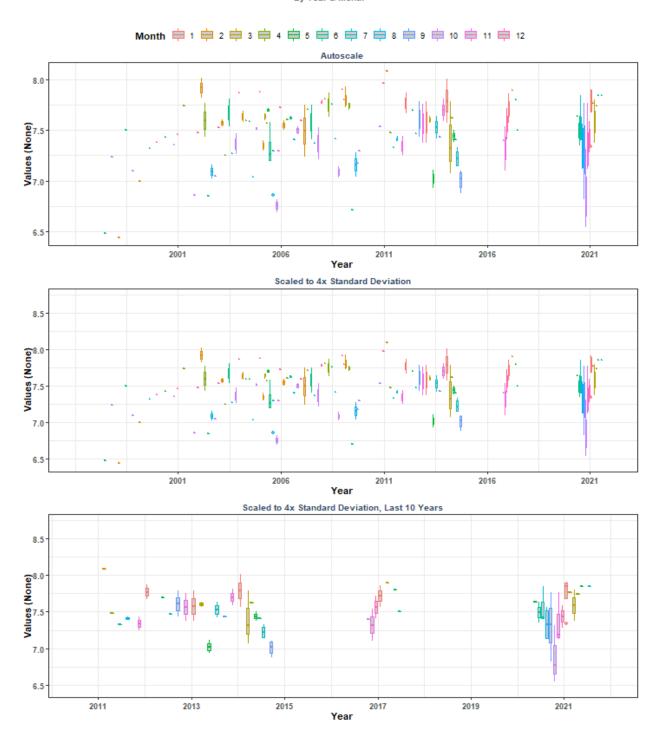
### Summary Box Plots for Terra Ceia Aquatic Preserve By Month



## Summary Box Plots for Tomoka Marsh Aquatic Preserve By Year



### Summary Box Plots for Tomoka Marsh Aquatic Preserve By Year & Month



# Summary Box Plots for Tomoka Marsh Aquatic Preserve By Month

