

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

Last compiled on 07 June, 2022

## Contents

<b>Important Notes</b>	<b>1</b>
<b>Libraries</b>	<b>2</b>
<b>File Import</b>	<b>2</b>
<b>Data Filtering and Data Impacted by Specific Value Qualifiers</b>	<b>2</b>
<b>Managed Area Statistics</b>	<b>6</b>
<b>Monitoring Location Statistics</b>	<b>7</b>
<b>Seasonal Kendall Tau Analysis</b>	<b>8</b>
<b>Appendix I: Scatter Plot of Entire Dataset</b>	<b>11</b>
<b>Appendix II: Dataset Summary Box Plots</b>	<b>14</b>
<b>Appendix III: Excluded Managed Areas</b>	<b>19</b>
<b>Appendix IV: Managed Area Trendlines</b>	<b>28</b>
<b>Appendix V: Managed Area Summary Box Plots</b>	<b>56</b>

## Important Notes

All scripts and outputs can be found on the SEACAR GitHub repository:

[https://github.com/FloridaSEACAR/SEACAR\\_Panzik](https://github.com/FloridaSEACAR/SEACAR_Panzik)

Note: The top 2% of data is excluded when computing mean and standard deviations in plotting sections solely for the purpose of getting y-axis scales. The exclusion of the top 2% is not used in any statistics that are exported.

## 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)
library(kableExtra)
windowsFonts(`Segoe UI` = windowsFont('Segoe UI'))
options(scipen=999)
opts_chunk$set(warning=FALSE, message=FALSE, dpi=200)
```

## 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 <- fread(file_in, sep="|", header=TRUE, stringsAsFactors=FALSE,
              select=c("ManagedAreaName", "ProgramID", "ProgramName",
                      "ProgramLocationID", "SampleDate", "Year", "Month",
                      "RelativeDepth", "ActivityType", "ParameterName",
                      "ResultValue", "ParameterUnits", "ValueQualifier",
                      "SEACAR_QAQCFlagCode", "Include"), na.strings="")

parameter <- unique(data$ParameterName)
unit <- unique(data$ParameterUnits)
data$SampleDate <- as.Date(data$SampleDate)
data$YearMonth <- paste0(data$Month, "- ", data$Year)
data$YearMonthDec <- data$Year + ((data$Month-0.5) / 12)
data$DecDate <- decimal_date(data$SampleDate)
```

## 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 process if it is from the correct depth and has an `Include` value of 1. The script also only looks at data of the desired `ActivityType` which indicates whether it was measured in the field (`Field`) or in the lab (`Sample`).

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

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

```

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

data$Include <- as.logical(data$Include)
data$Include[grep("H", data$ValueQualifier[data$ProgramID==476])] <- TRUE
data <- data[!is.na(data$ResultValue),]

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

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

if(param_name=="Chlorophyll_a_uncorrected_for_pheophytin" |
  param_name=="Salinity" | param_name=="Turbidity"){
  data <- data[grep(activity, data$ActivityType[!is.na(data$ActivityType)]),]
}

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")],
                         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")],
                         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$Include)
pass_filter <- length(data$Include[data$Include==TRUE])

count_H <- length(grep("H", data$ValueQualifier[data$ProgramID==476]))
perc_H <- 100*count_H/length(data$ValueQualifier)

count_I <- length(grep("I", data$ValueQualifier))
perc_I <- 100*count_I/length(data$ValueQualifier)

count_Q <- length(grep("Q", data$ValueQualifier))
perc_Q <- 100*count_Q/length(data$ValueQualifier)

count_S <- length(grep("S", data$ValueQualifier))
perc_S <- 100*count_S/length(data$ValueQualifier)

count_U <- length(grep("U", data$ValueQualifier))
perc_U <- 100*count_U/length(data$ValueQualifier)

data$VQ_Plot <- data$ValueQualifier

inc_H <- ifelse(param_name=="pH" | param_name=="Dissolved_Oxygen" |
                  param_name=="Dissolved_Oxygen_Saturation", TRUE, FALSE)

if (inc_H==TRUE){
  data$VQ_Plot <- gsub("[^HU]+", "", data$VQ_Plot)
  data$VQ_Plot <- gsub("UH", "HU", data$VQ_Plot)
  data$VQ_Plot[na.omit(data$ProgramID!=476)] <- gsub("[^U]+", "",
                                                       data$VQ_Plot[na.omit(data$ProgramID!=476)])
  data$VQ_Plot[data$VQ_Plot==""] <- NA}

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$ValueQualifier))
  perc_S <- 100*count_S/length(data$ValueQualifier)
  data$VQ_Plot <- gsub("[^SU]+", "", data$VQ_Plot)
  data$VQ_Plot <- gsub("US", "SU", data$VQ_Plot)
  data$VQ_Plot[data$VQ_Plot==""] <- NA
  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)
  data$VQ_Plot[data$VQ_Plot==""] <- NA
  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: 39997, Number Passed Filter: 39890
## I Codes: 1192 (2.980224%)
## Q Codes: 91 (0.227517%)
## U Codes: 1132 (2.830212%)

```

```

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$ManagedAreaName), ])
fwrite(data_summ, paste0(out_dir, "/", param_name, "_", activity, "_", depth,
                       "_DataSummary.csv"), sep=",")
rm(data_summ)
MA_Include <- MA_Summ$ManagedAreaName [MA_Summ$SufficientData==TRUE &
                                         MA_Summ$N_Data<2000000]

```

```

n <- length(MA_Include)
MA_Exclude <- MA_Summ[MA_Summ$N_Years<10 & MA_Summ$N_Years>0,]
MA_Exclude <- MA_Exclude[,c("ManagedAreaName", "N_Years")]
z <- nrow(MA_Exclude)

```

## 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, paste0(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, "_",
                        "_ManagedArea_Year_Stats.txt"), sep="|")
rm(MA_Y_Stats)

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

```

## 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 `ProgramID` 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$ManagedAreaName,
                                             Mon_Stats$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,
                         stats.maxYear) {
  tau <- NULL
  tryCatch({ken <- kendallSeasonalTrendTest(
    y = data$ResultValue,
    ...)}

```

```

    season = data$Month,
    year = data$Year,
    independent.obs = independent)

tau <- ken$estimate[1]
p <- ken$p.value[2]
slope <- ken$estimate[2]
intercept <- ken$estimate[3]
trend <- trend_calculator(slope, stats.median, p)
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$AreaID),
       unique(data$ManagedAreaName),
       stats.median,
       independent,
       tau,
       p,
       slope,
       intercept,
       trend)
return(KT)
}
runStats <- function(data) {
  data$Index <- as.Date(data$SampleDate) # , "%Y-%m-%d")
  data$ResultValue <- as.numeric(data$ResultValue)
  # Calculate basic stats
  stats.median <- median(data$ResultValue, na.rm = TRUE)
  stats.minYear <- min(data$Year, na.rm = TRUE)
  stats.maxYear <- max(data$Year, na.rm = TRUE)
  # Calculate Kendall Tau and Slope stats, then update appropriate columns and table
  KT <- tauSeasonal(data, TRUE, stats.median,
                    stats.minYear, stats.maxYear)
  if (is.null(KT[9])) {

```

```

        KT <- tauSeasonal(data, FALSE, stats.median,
                            stats.minYear, stats.maxYear)
    }
    if (is.null(KT$Stats) == TRUE) {
        KT$Stats <- KT
    } else{
        KT$Stats <- rbind(KT$Stats, KT)
    }
    return(KT$Stats)
}
trend_calculator <- function(slope, median_value, p) {
    trend <-
        if (p < .05 & abs(slope) > abs(median_value) / 10.) {
            if (slope > 0) {
                2
            }
            else {
                -2
            }
        }
        else if (p < .05 & abs(slope) < abs(median_value) / 10.) {
            if (slope > 0) {
                1
            }
            else {
                -1
            }
        }
        else
            0
    return(trend)
}
KT$Stats <- NULL
# Loop that goes through each managed area.
# List of managed areas stored in MA_Years$ManagedAreaName
c_names <- c("AreaID", "ManagedAreaName", "Median", "Independent",
           "tau", "p", "SennSlope", "SennIntercept", "Trend")
if(n==0){
    KT$Stats <- data.frame(matrix(ncol=length(c_names),
                                    nrow=length(MA_Summ$ManagedAreaName)))
    colnames(KT$Stats) <- c_names
    KT$Stats[, c("AreaID", "ManagedAreaName")] <-
        MA_Summ[, c("AreaID", "ManagedAreaName")]
} else{
    for (i in 1:n) {
        x <- nrow(data[data$Use_In_Analysis == TRUE &
                        data$ManagedAreaName == MA_Include[i], ])
        if (x>0) {
            KT$Stats <- runStats(data[data$Use_In_Analysis == TRUE &
                                         data$ManagedAreaName ==
                                         MA_Include[i], ])
        }
    }
}

```

```

KT.Stats <- as.data.frame(KT.Stats)
c_names <- c("AreaID", "ManagedAreaName", "Median", "Independent",
           "tau", "p", "SennSlope", "SennIntercept", "Trend")
if(dim(KT.Stats)[2]==1){
  KT.Stats <- as.data.frame(t(KT.Stats))
}
colnames(KT.Stats) <- c_names
rownames(KT.Stats) <- seq(1:nrow(KT.Stats))
KT.Stats$tau <- round(as.numeric(KT.Stats$tau), digits=4)
KT.Stats$p <- round(as.numeric(KT.Stats$p), digits=4)
KT.Stats$SennSlope <- as.numeric(KT.Stats$SennSlope)
KT.Stats$SennIntercept <- as.numeric(KT.Stats$SennIntercept)
KT.Stats$Trend <- as.integer(KT.Stats$Trend)
}

KT.Stats <- merge.data.frame(MA_Summ, KT.Stats,
                             by=c("AreaID", "ManagedAreaName"), all=TRUE)

KT.Stats <- as.data.table(KT.Stats[order(KT.Stats$ManagedAreaName), ])

fwrite(KT.Stats, paste0(out_dir, "/", param_name, "_", activity, "_", depth,
                       "KendallTau_Stats.txt"), sep="|")
data <- data[!is.na(data$ResultValue), ]

```

## 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"),
        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.title.x = element_text(margin = margin(t = 5, r = 0,
                                                    b = 10, l = 0)),
        axis.title.y = element_text(margin = margin(t = 0, r = 10,
                                                    b = 0, l = 0)),
        axis.text=element_text(size=10),
        axis.text.x=element_text(face="bold", angle = 60, hjust = 1),
        axis.text.y=element_text(face="bold"))

year_lower <- min(data$Year)
year_upper <- max(data$Year)
min_RV <- min(data$ResultValue)
mn_RV <- mean(data$ResultValue[data$ResultValue <
                                    quantile(data$ResultValue, 0.98)])
sd_RV <- sd(data$ResultValue[data$ResultValue <
                                    quantile(data$ResultValue, 0.98)])
x_scale <- ifelse(year_upper - year_lower > 30, 10, 5)

```

```

y_scale <- mn_RV + 4 * sd_RV

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")
  }}
}

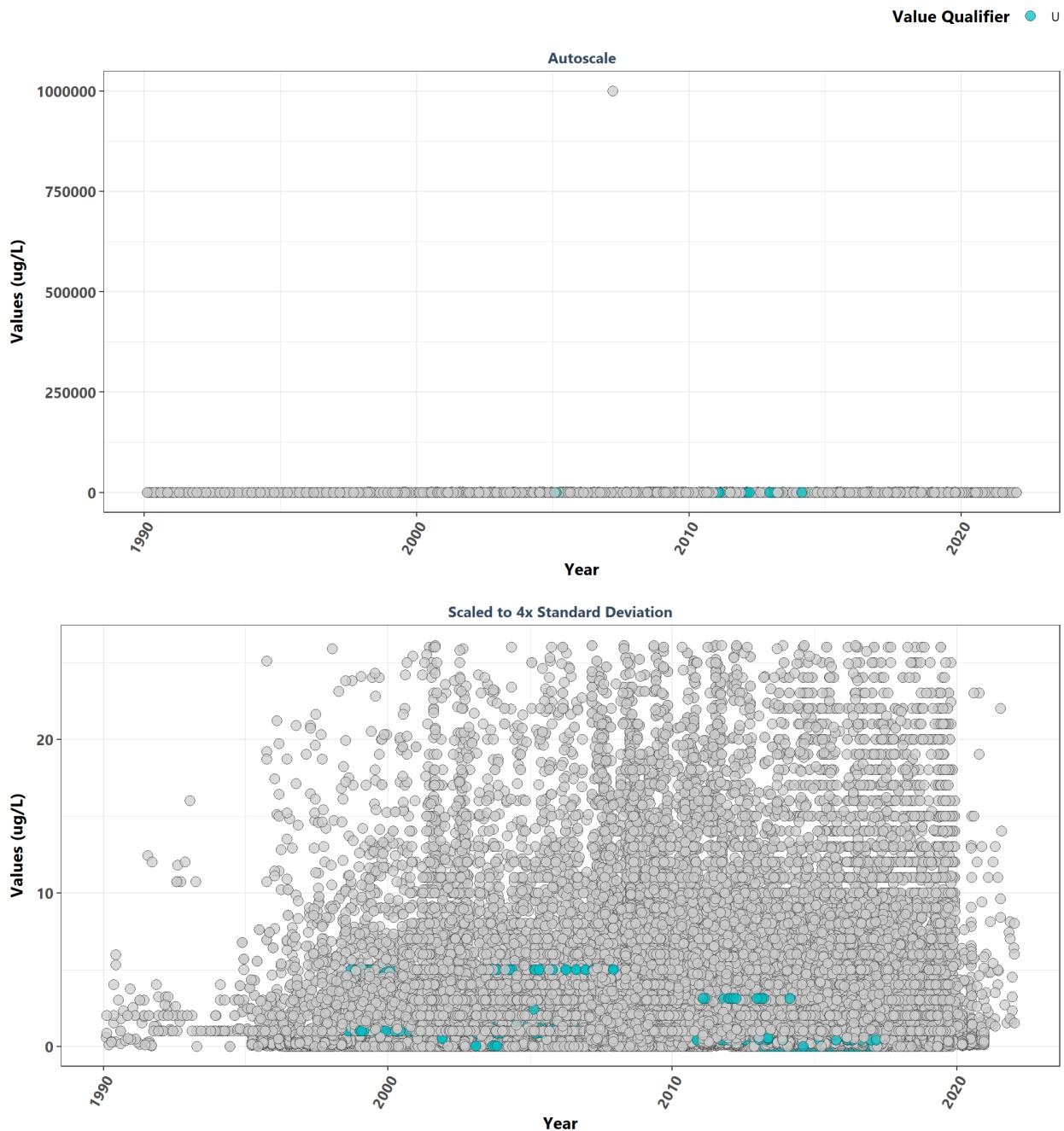
leg <- get_legend(p1)
pset <- ggarrange(leg, p1 + theme(legend.position="none"), p2,
                  ncol=1, heights=c(0.1, 1, 1))

p0 <- ggplot() + labs(title="Scatter Plot for Entire Dataset") +
  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])
mn_RV <- mean(data$ResultValue[data$Include==TRUE &
                                data$ResultValue <
                                quantile(data$ResultValue, 0.98)])
sd_RV <- sd(data$ResultValue[data$Include==TRUE &
                                data$ResultValue <
                                quantile(data$ResultValue, 0.98)])
y_scale <- mn_RV + 4 * sd_RV

p1 <- ggplot(data=data[data$Include==TRUE, ],
              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, ],
              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)

p0 <- ggplot() + labs(title="Summary Box Plots for Entire Data",
                      subtitle="By Year") + plot_theme +
theme(panel.border=element_blank(), panel.grid.major=element_blank(),
      panel.grid.minor=element_blank(), axis.line=element_blank())

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

```

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

```

p1 <- ggplot(data=data[data$Include==TRUE, ],
              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, ],
              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)
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",
                      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())

```

```

  panel.grid.minor=element_blank(), axis.line=element_blank())

```

```

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

```

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

```

p1 <- ggplot(data=data[data$Include==TRUE, ],
              aes(x=Month, y=ResultValue,
                  group=Month, fill=as.factor(Month))) +
  geom_boxplot(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, ],
              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")

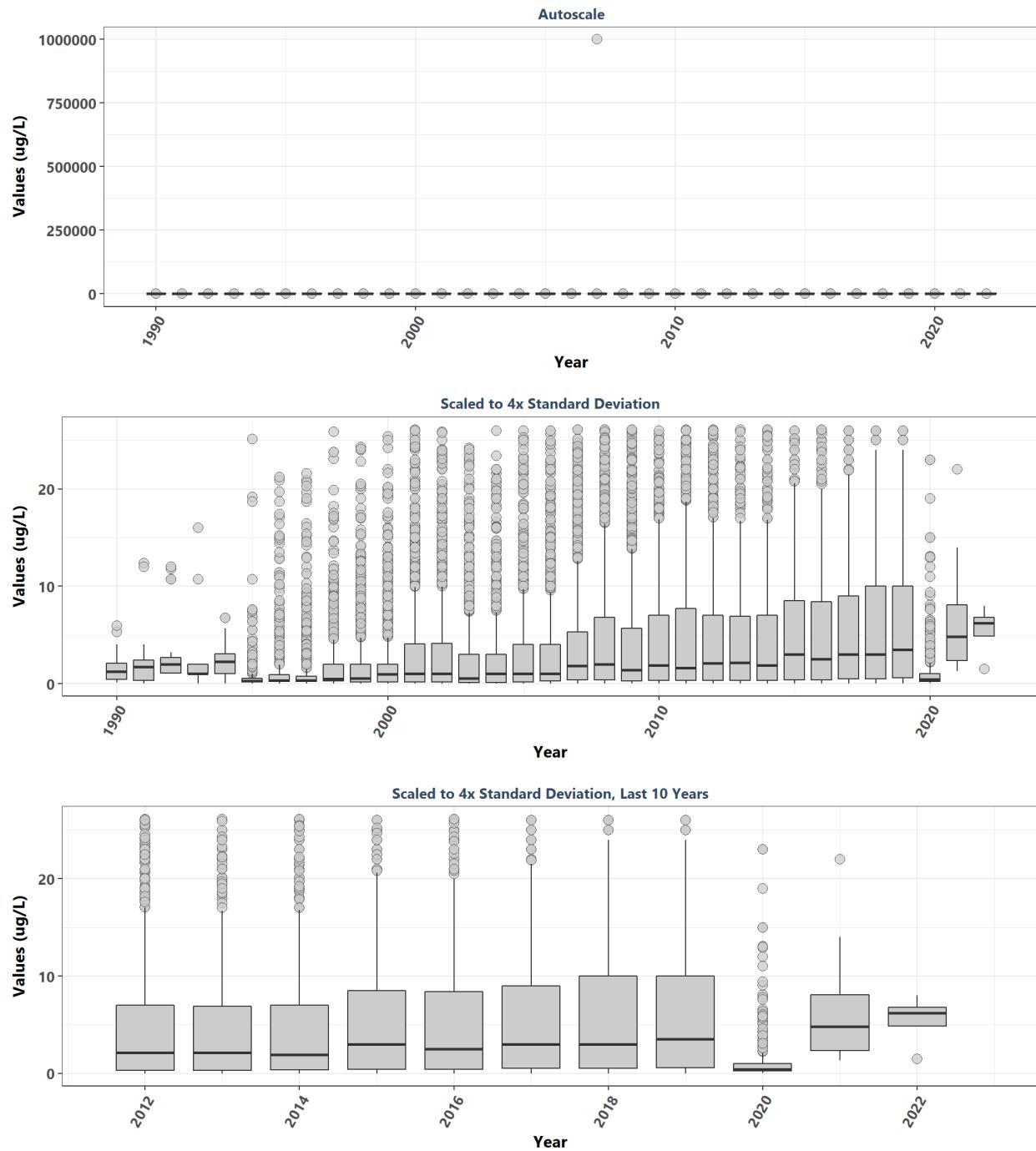
p3 <- ggplot(data=data[data$Include==TRUE &
                           data$Year >= max(data$Year) - 10, ],
              aes(x=Month, y=ResultValue,
                  group=Month, fill=as.factor(Month))) +
  geom_boxplot(color="#333333", outlier.shape=21, outlier.size=3,
               outlier.color="#333333", 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)
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",
                       subtitle="By Month") + plot_theme +
  theme(panel.border=element_blank(), panel.grid.major=element_blank(),
        panel.grid.minor=element_blank(), axis.line=element_blank())

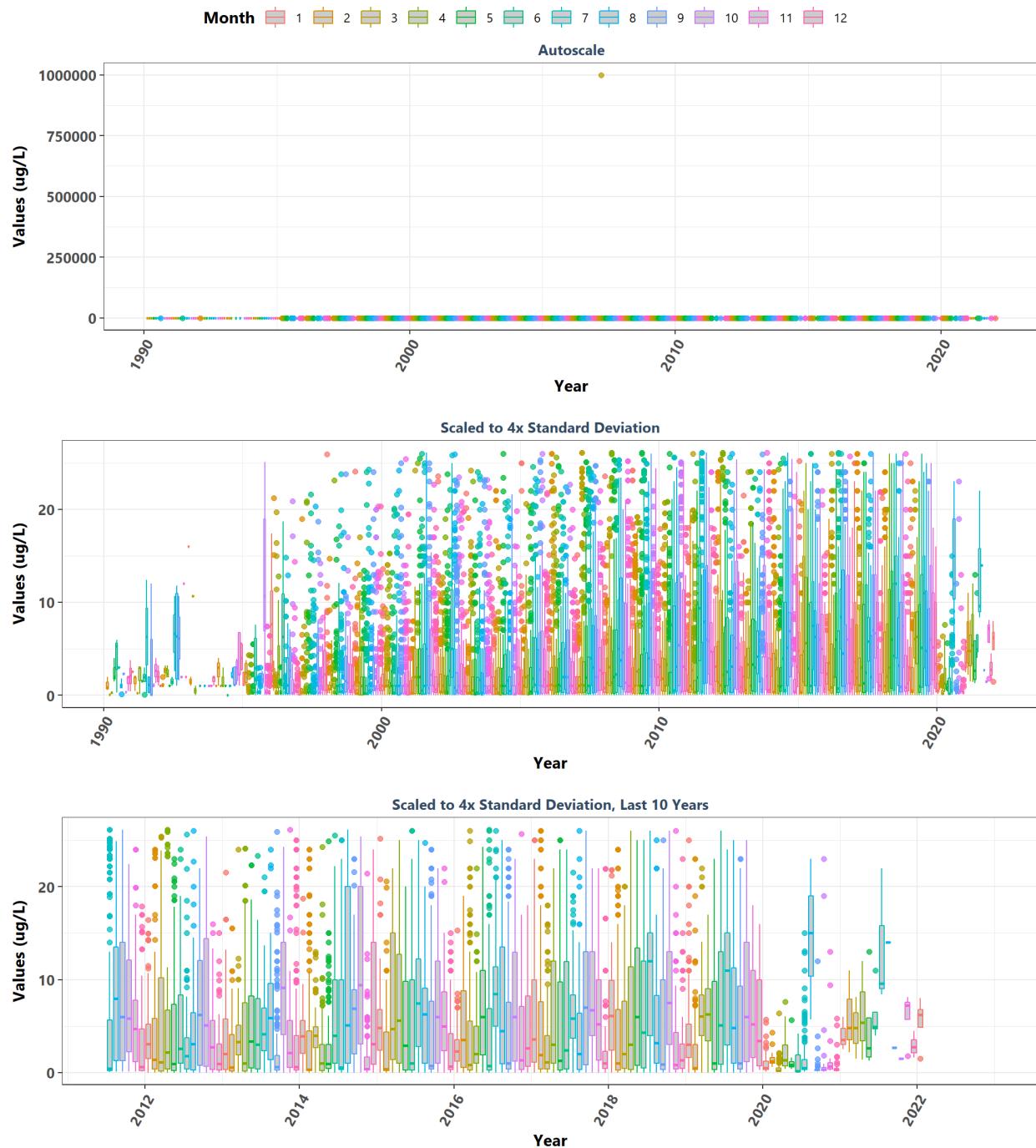
Mset <- ggarrange(p0, set, ncol=1, heights=c(0.07, 1))

```

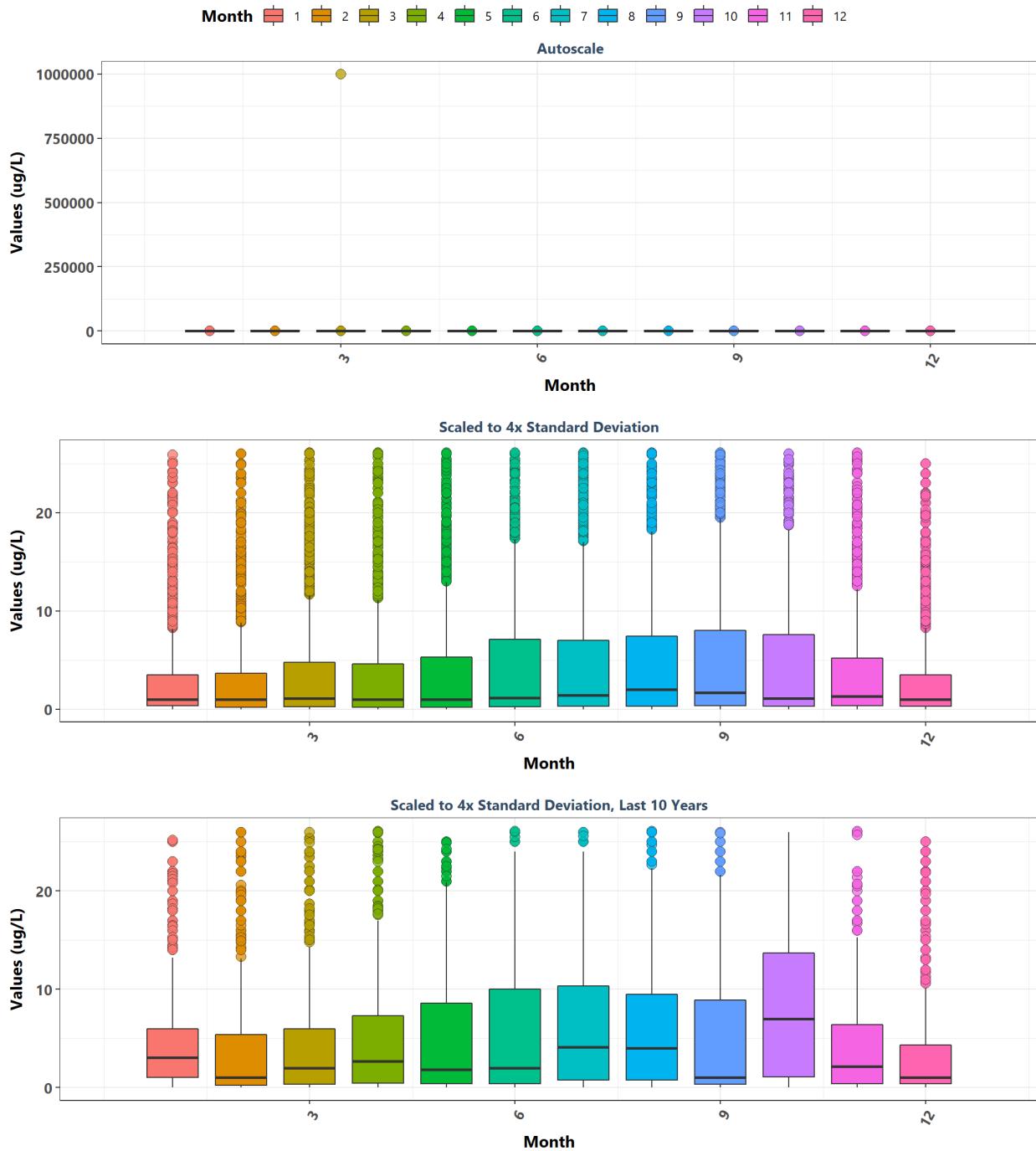
**Summary Box Plots for Entire Data**  
By Year



**Summary Box Plots for Entire Data**  
By Year & Month



### Summary Box Plots for Entire Data By Month



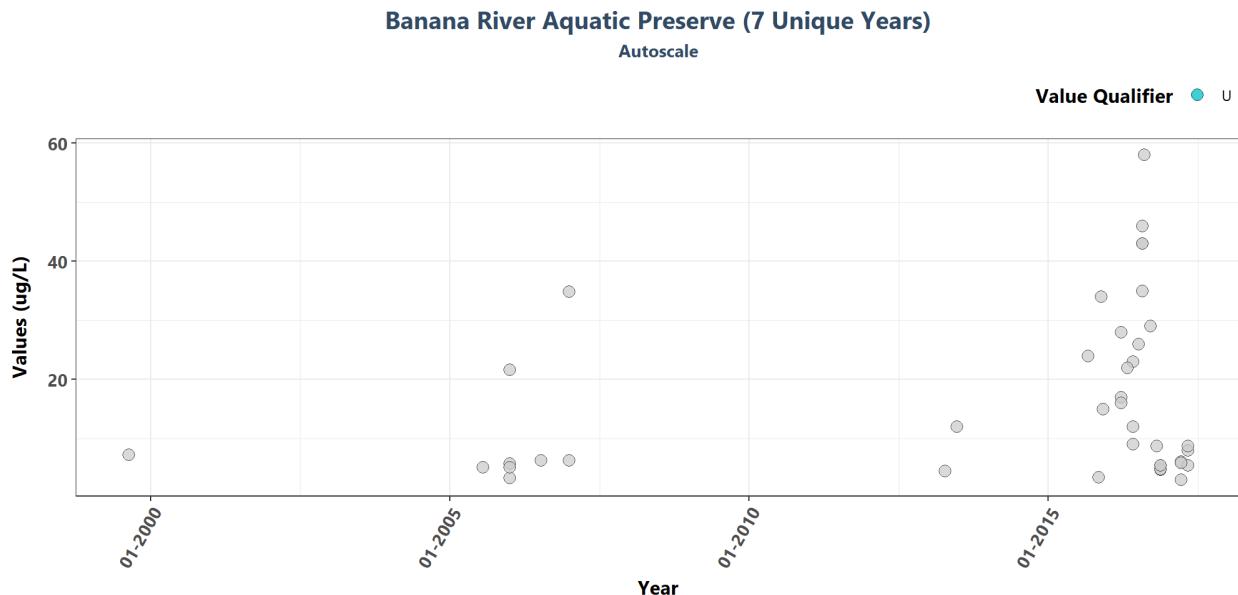
### Appendix III: Excluded Managed Areas

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){
    p1<-ggplot(data=data[data$ManagedAreaName==MA_Exclude$ManagedAreaName[i]&
      data$Include==TRUE, ],
      aes(x=SampleDate, y=ResultValue, fill=VQ_Plot)) +
      geom_point(shape=21, size=3, color="#333333", alpha=0.75) +
      labs(title=paste0(MA_Exclude$ManagedAreaName[i], " (",
        MA_Exclude$N_Years[i], " Unique Years")),
        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("%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")
      } else {
        scale_fill_manual(values=c("U"= "#00BFC4"), na.value="#cccccc")
      })
      print(p1)
    }
  }
}

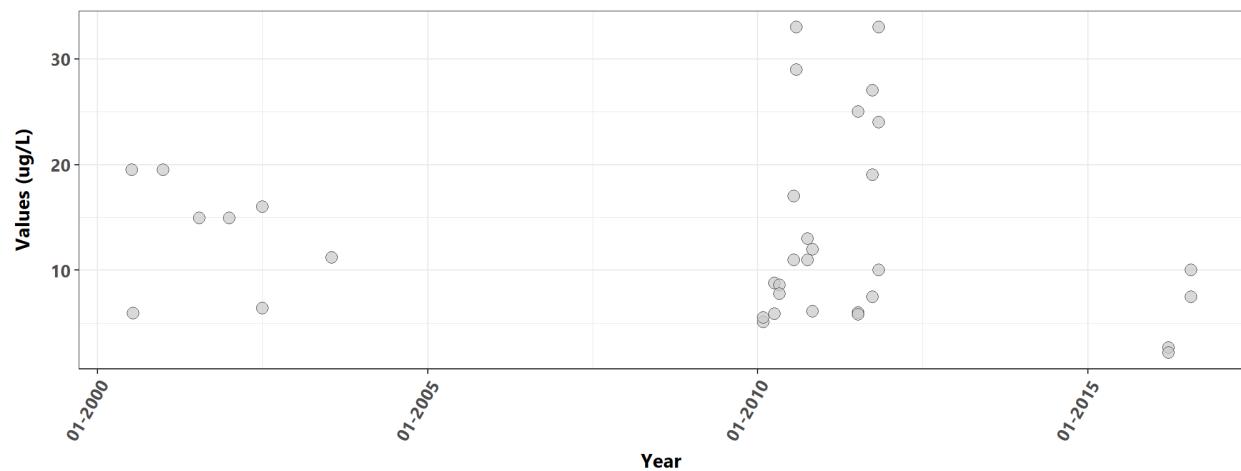
```



### Boca Ciega Bay Aquatic Preserve (7 Unique Years)

Autoscale

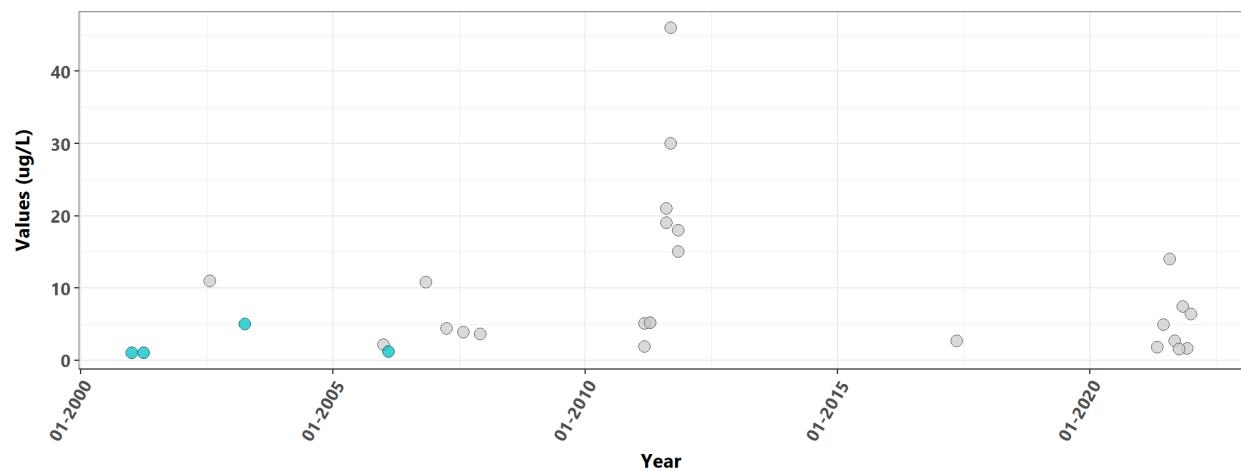
Value Qualifier ● U



### Cape Haze Aquatic Preserve (9 Unique Years)

Autoscale

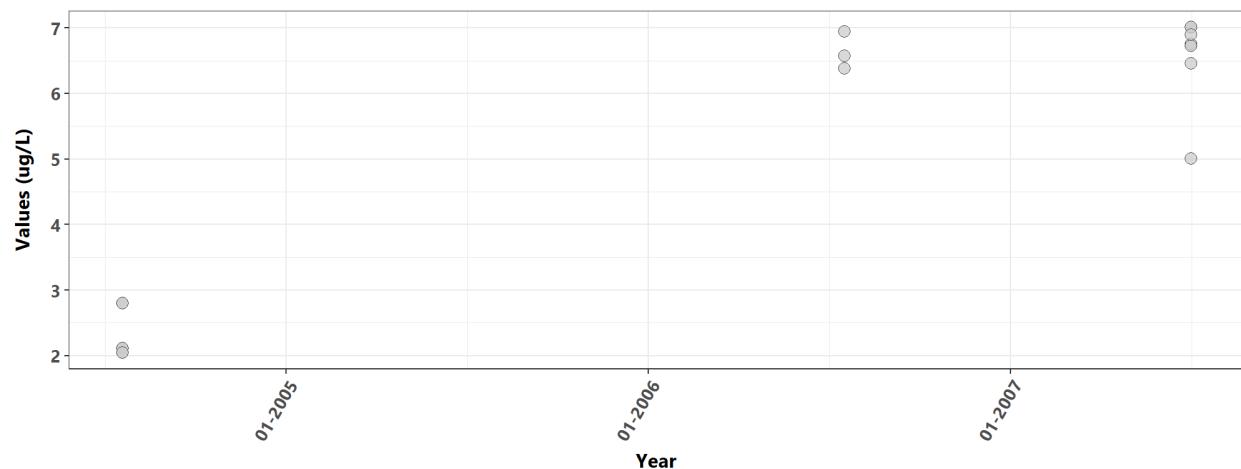
Value Qualifier ● U



### Fort Clinch State Park Aquatic Preserve (3 Unique Years)

Autoscale

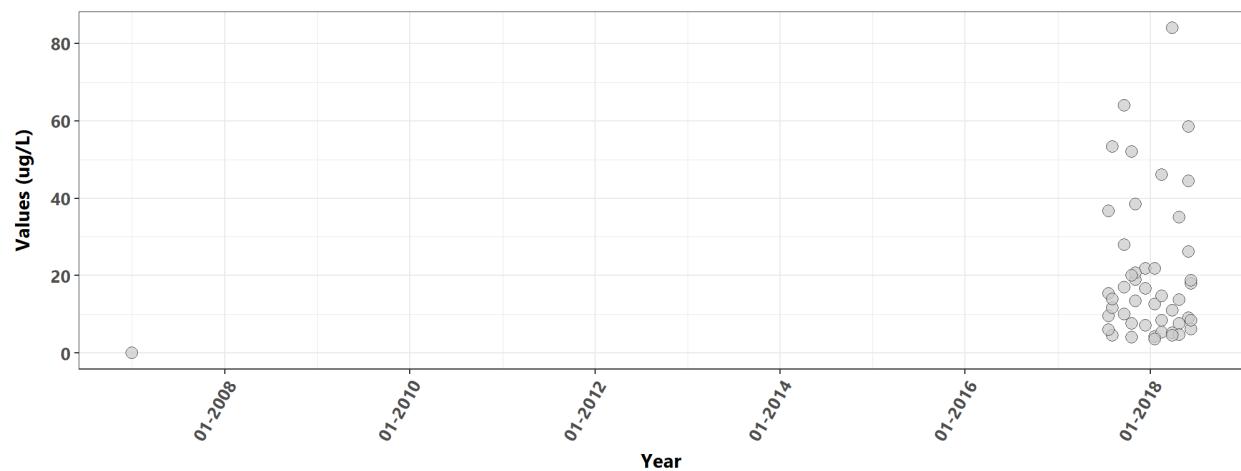
Value Qualifier ● U



### Guana River Marsh Aquatic Preserve (3 Unique Years)

Autoscale

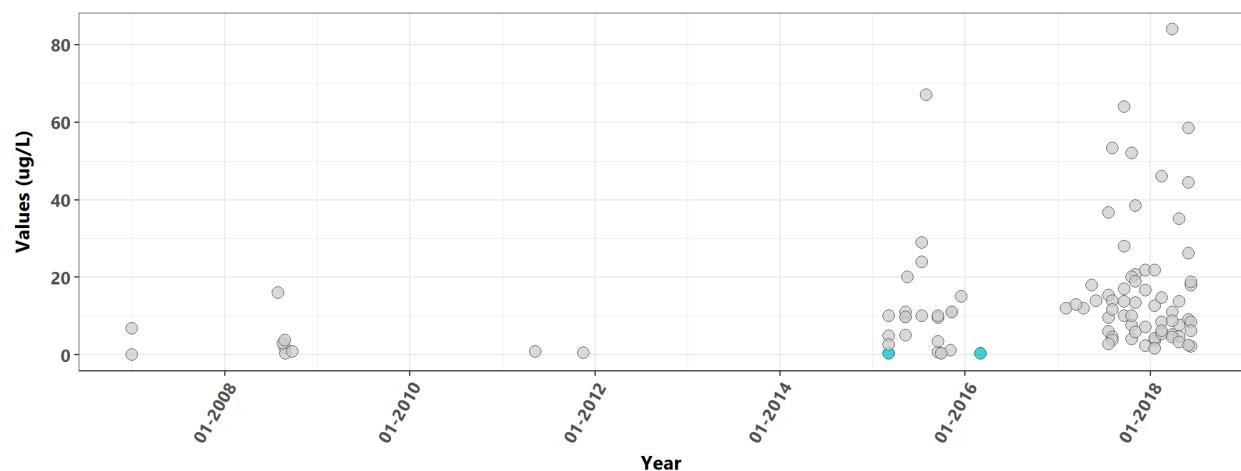
Value Qualifier ● U



### Guana Tolomato Matanzas National Estuarine Research Reserve (7 Unique Years)

Autoscale

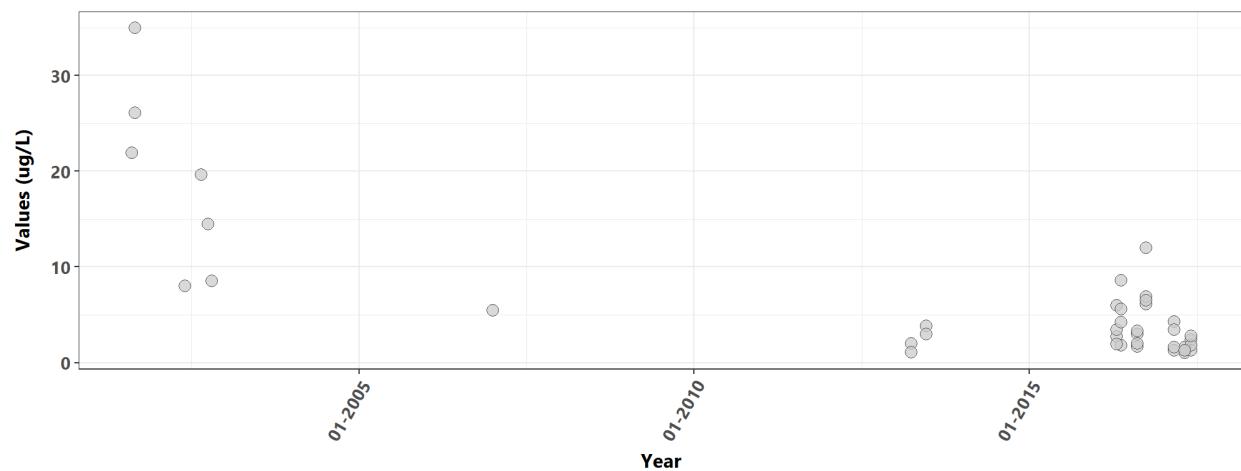
Value Qualifier ● U

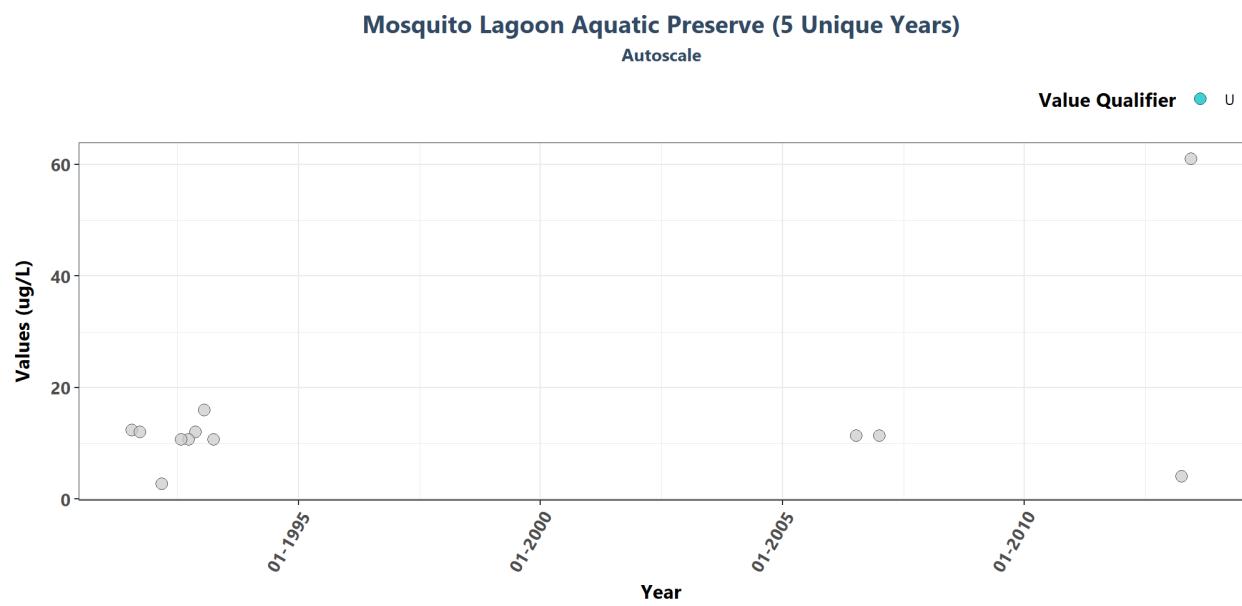
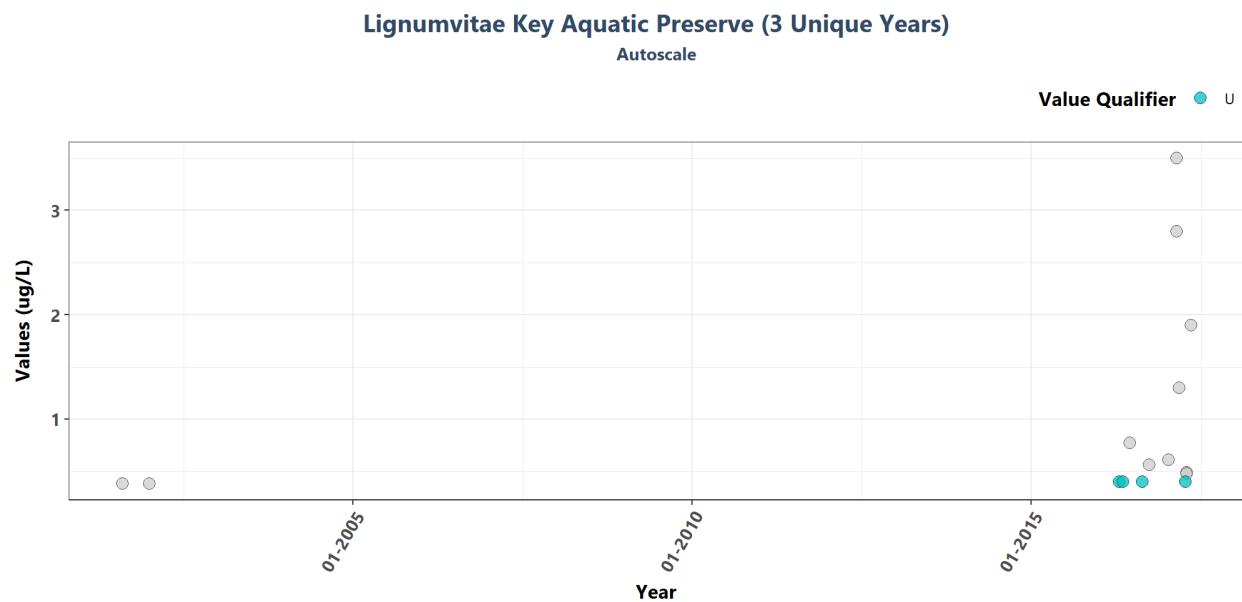


### Indian River-Vero Beach to Ft. Pierce Aquatic Preserve (6 Unique Years)

Autoscale

Value Qualifier ● U

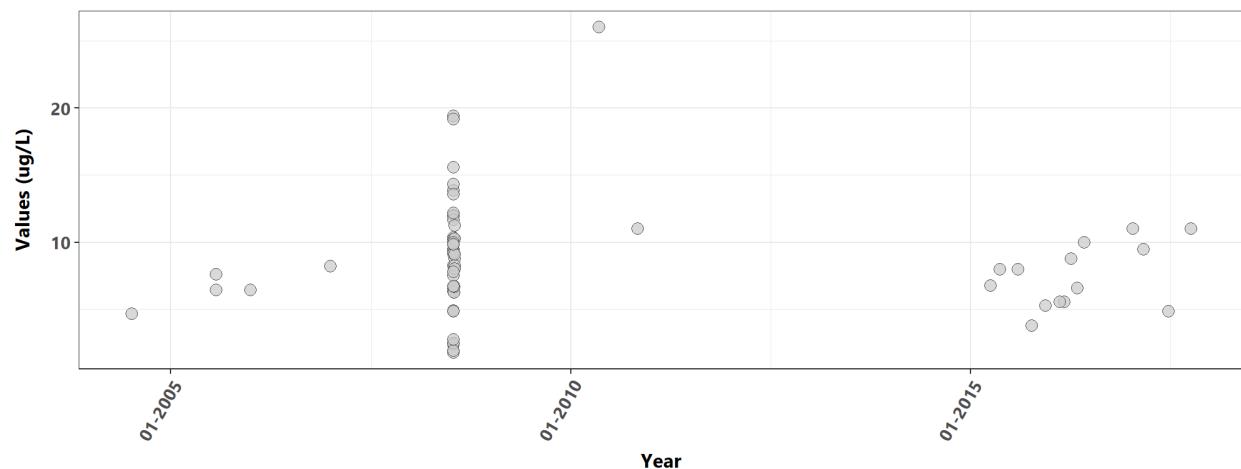




### Nassau River-St. Johns River Marshes Aquatic Preserve (8 Unique Years)

Autoscale

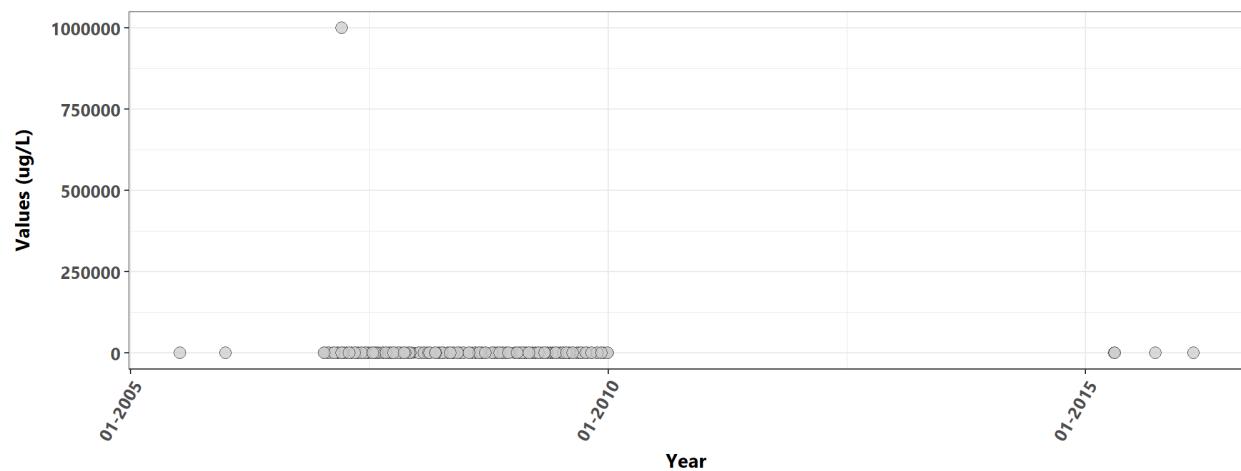
Value Qualifier ● U



### Nature Coast Aquatic Preserve (6 Unique Years)

Autoscale

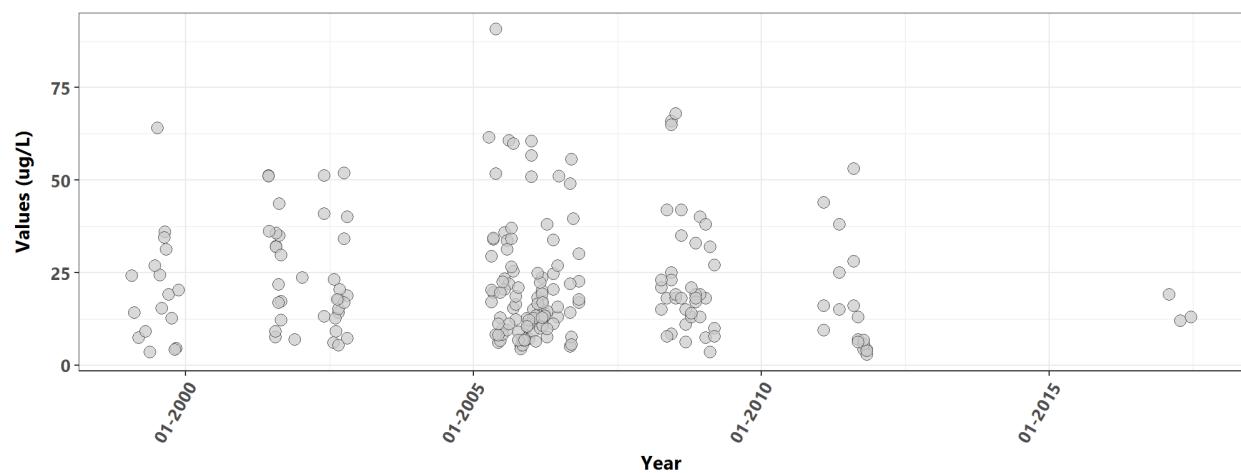
Value Qualifier ● U



### North Fork St. Lucie Aquatic Preserve (9 Unique Years)

Autoscale

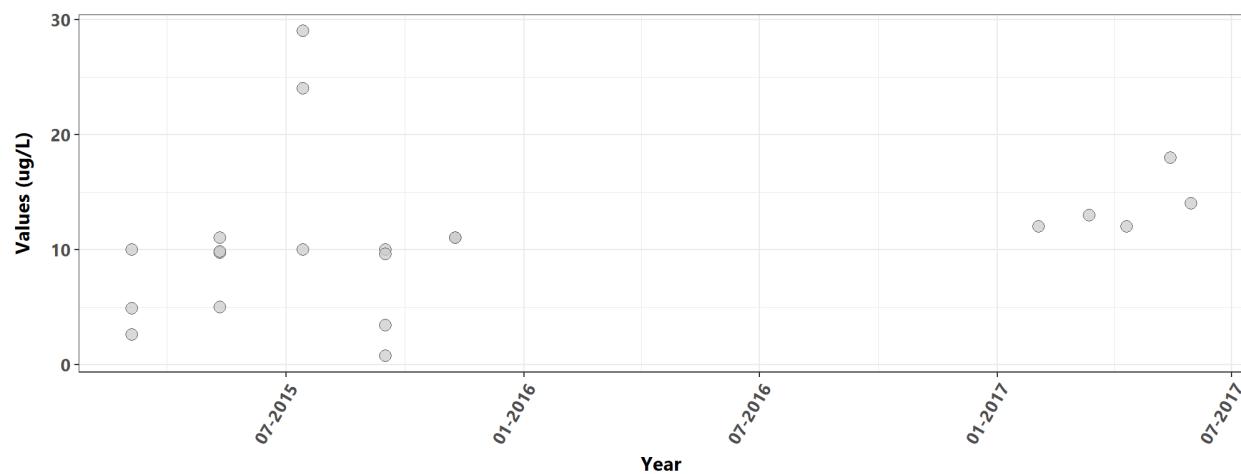
Value Qualifier ● U



### Pellicer Creek Aquatic Preserve (2 Unique Years)

Autoscale

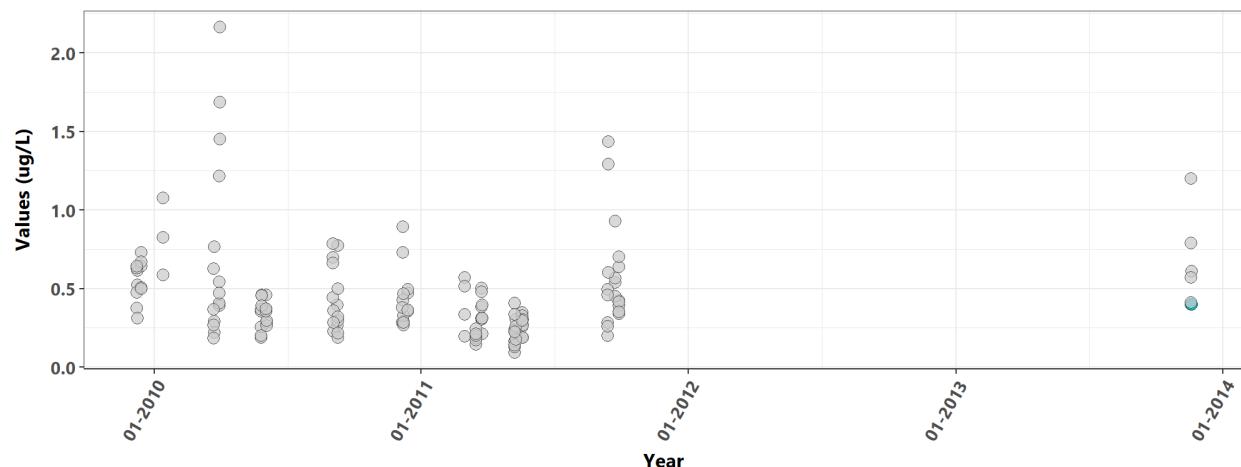
Value Qualifier ● U



### Southeast Florida Coral Reef Ecosystem Conservation Area (4 Unique Years)

Autoscale

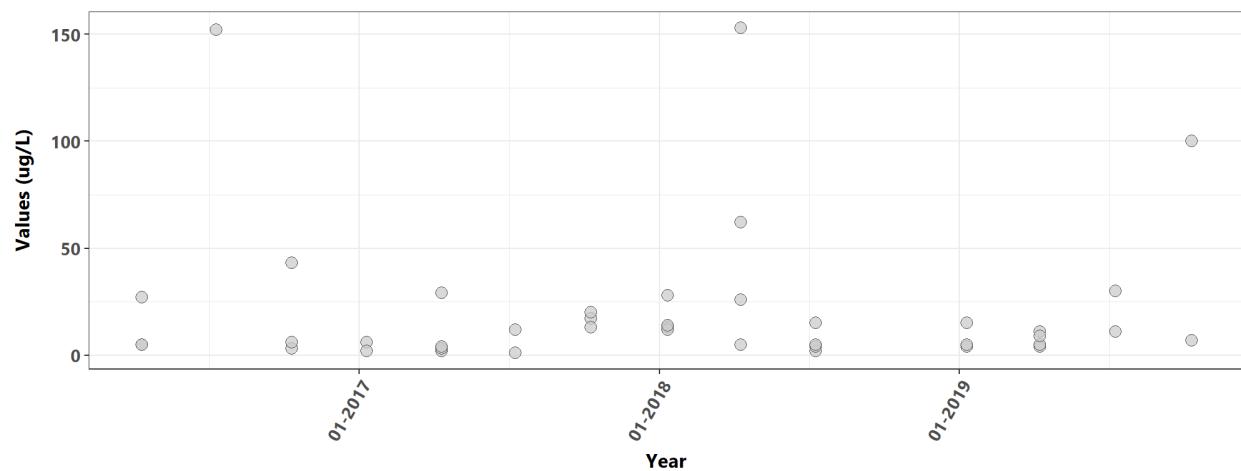
Value Qualifier ● U

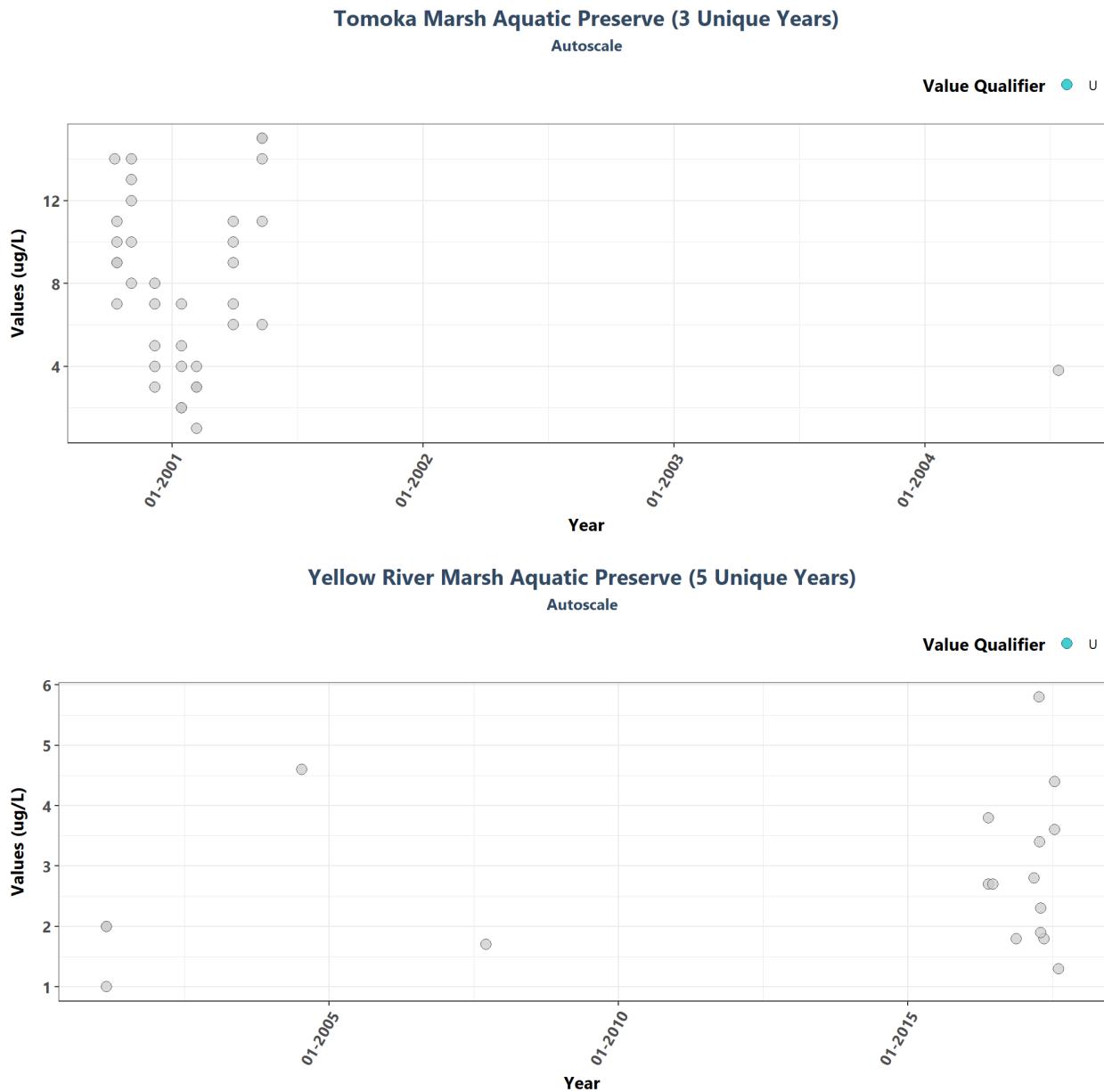


### St. Joseph Bay State Buffer Preserve (4 Unique Years)

Autoscale

Value Qualifier ● U





## 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

```

if(n==0){
  print("There are no managed areas that qualify.")
} else {
  for (i in 1:n) {
    plot_data <- data[data$SufficientData==TRUE &
                      data$ManagedAreaName==MA_Include[i],]
    year_lower <- min(plot_data$Year)
    year_upper <- max(plot_data$Year)
    min_RV <- min(plot_data$ResultValue)
    mn_RV <- mean(plot_data$ResultValue[plot_data$ResultValue <
                                         quantile(data$ResultValue, 0.98)])
    sd_RV <- sd(plot_data$ResultValue[plot_data$ResultValue <
                                         quantile(data$ResultValue, 0.98)])
    x_scale <- ifelse(year_upper - year_lower > 30, 10, 5)
    y_scale <- 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]]
    s_int <- KT.Stats$SennIntercept[KT.Stats$ManagedAreaName==MA_Include[i]]
    trend <- KT.Stats$Trend[KT.Stats$ManagedAreaName==MA_Include[i]]
    p <- KT.Stats$p[KT.Stats$ManagedAreaName==MA_Include[i]]

    model <- lm(ResultValue ~ DecDate,
                data=plot_data)
    m_int <- coef(model)[[1]]
    m_slope <- coef(model)[[2]]
    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="#000099", size=1.2, alpha=0.7) +
      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")
      }}
}

```

```

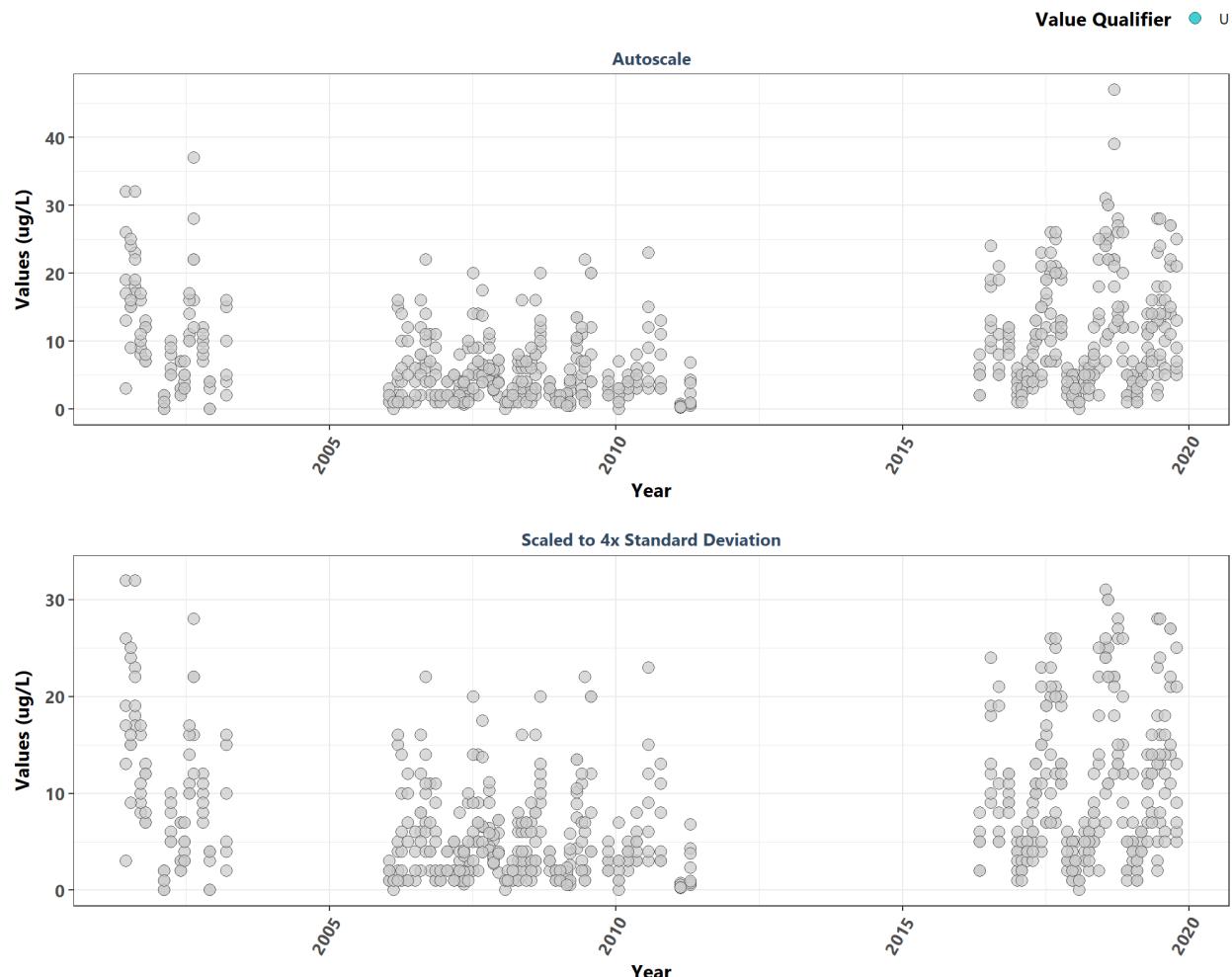
p2 <- ggplot(data=plot_data,
  aes(x=DecDate, y=ResultValue, fill=VQ_Plot)) +
  geom_point(shape=21, size=3, color="#333333", alpha=0.75) +
  geom_abline(aes(slope=s_slope, intercept=s_int),
    color="#000099", size=1.2, alpha=0.7) +
  ylim(min_RV, y_scale) +
  labs(subtitle="Scaled to 4x Standard Deviation",
    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)
  KTset <- ggarrange(leg, p1 + theme(legend.position="none"), p2,
    ncol=1, heights=c(0.1, 1, 1))

p0 <- ggplot() + labs(title=paste0(MA_Include[i])) +
  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(0.1, 1)))
cat('\n')
print(KT$Stats[KT$Stats$ManagedAreaName==MA_Include[i], ] %>%
  select(N_Data, Median, tau, SennSlope, SennIntercept, p, Trend) %>%
  kable(format="latex") %>%
  row_spec(0,bold=TRUE) %>%
  kable_styling(latex_options = "HOLD_position") %>%
  add_footnote(
    "p < 0.00005 appear as 0 due to rounding"))
cat('\n')
rm(plot_data)
rm(KTset, leg)
}
}

```

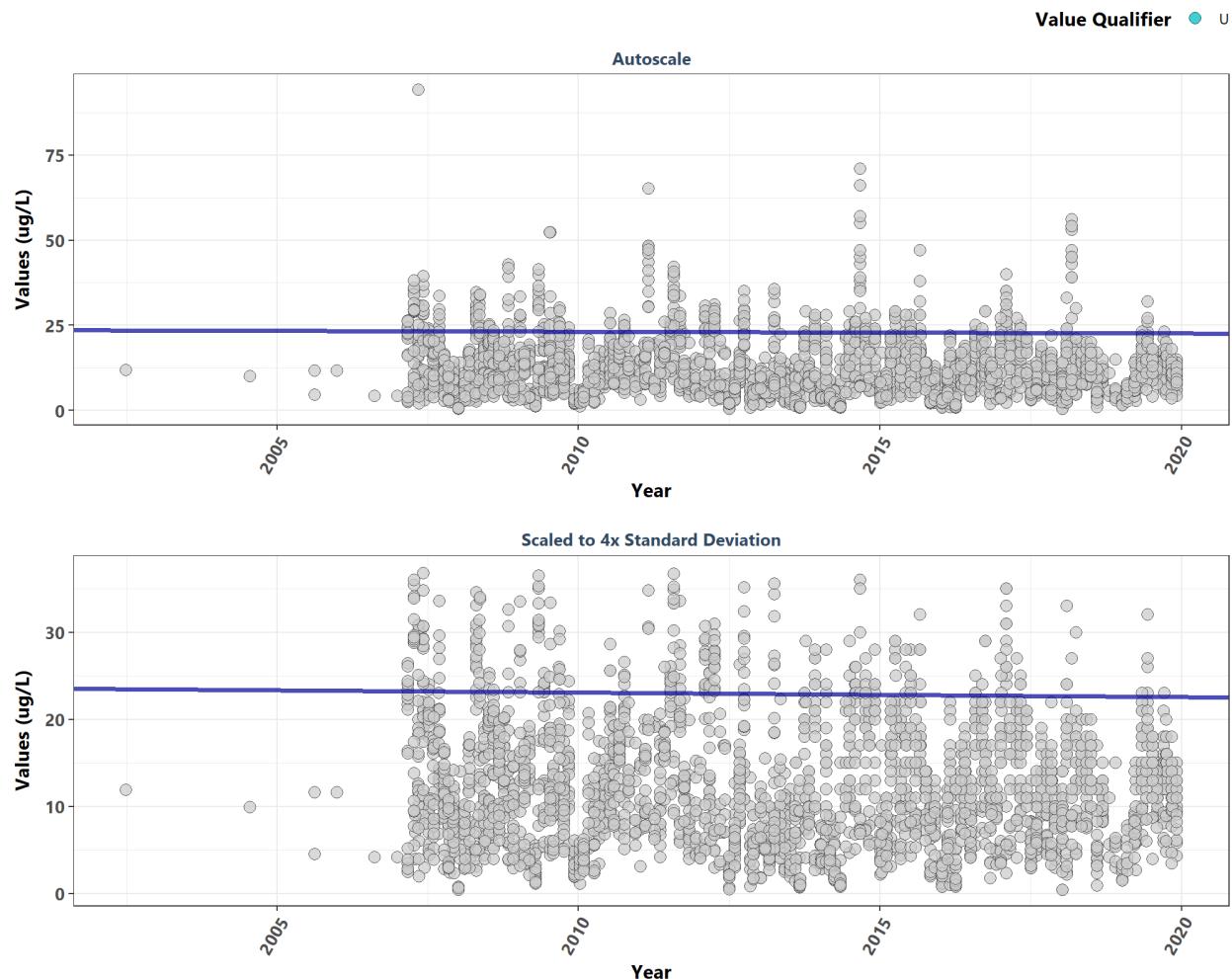
### Alligator Harbor Aquatic Preserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
678	5	0.1885	0.1538462	-339.1	0	1

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

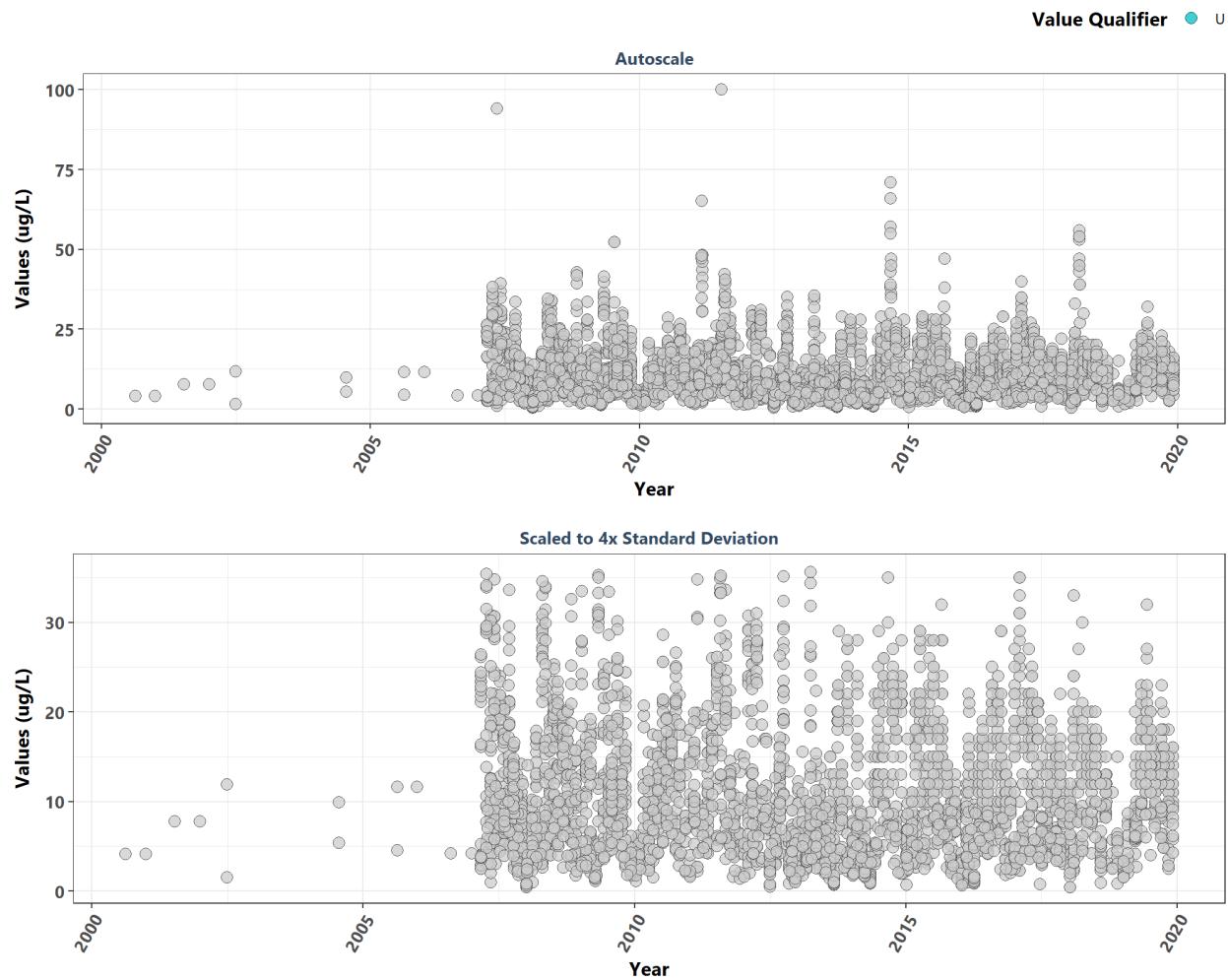
### Apalachicola Bay Aquatic Preserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
2843	11	-0.0178	-0.0518333	127.296	0.0527	0

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

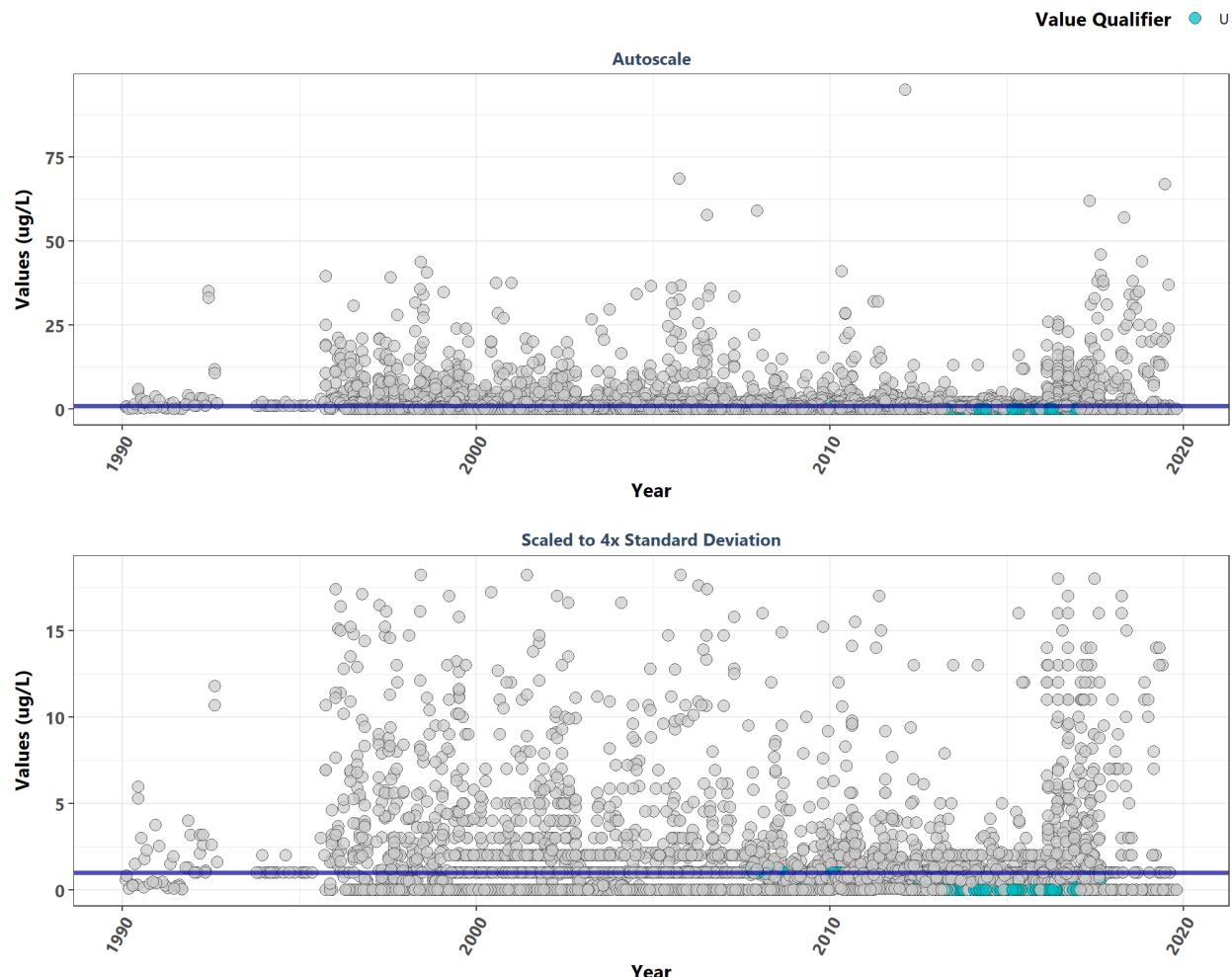
### Apalachicola National Estuarine Research Reserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
3429	9.8	0.0125	0	-14.4	0.5574	0

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

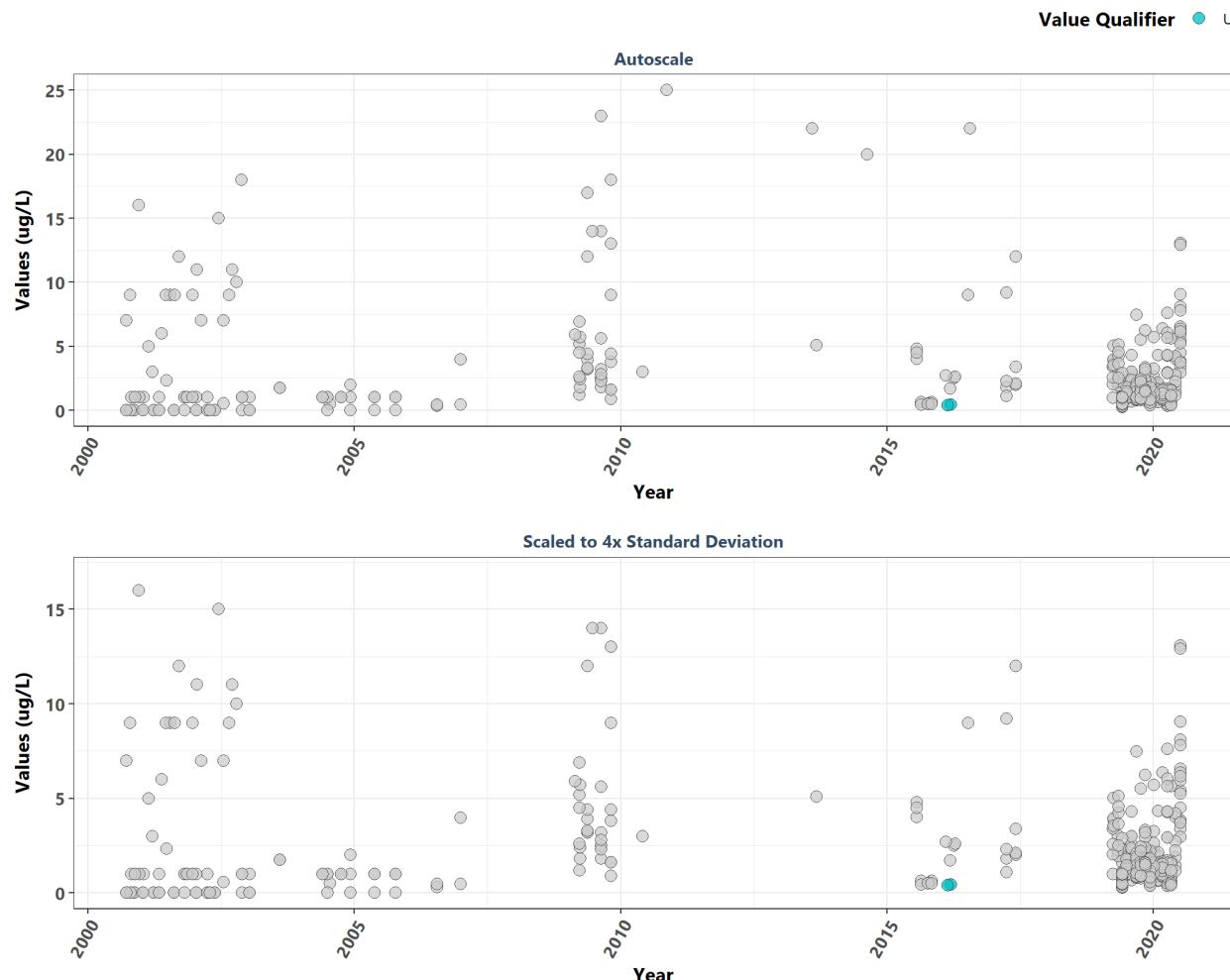
### Big Bend Seagrasses Aquatic Preserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
3651	1	-0.068	0	1	0	-1

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

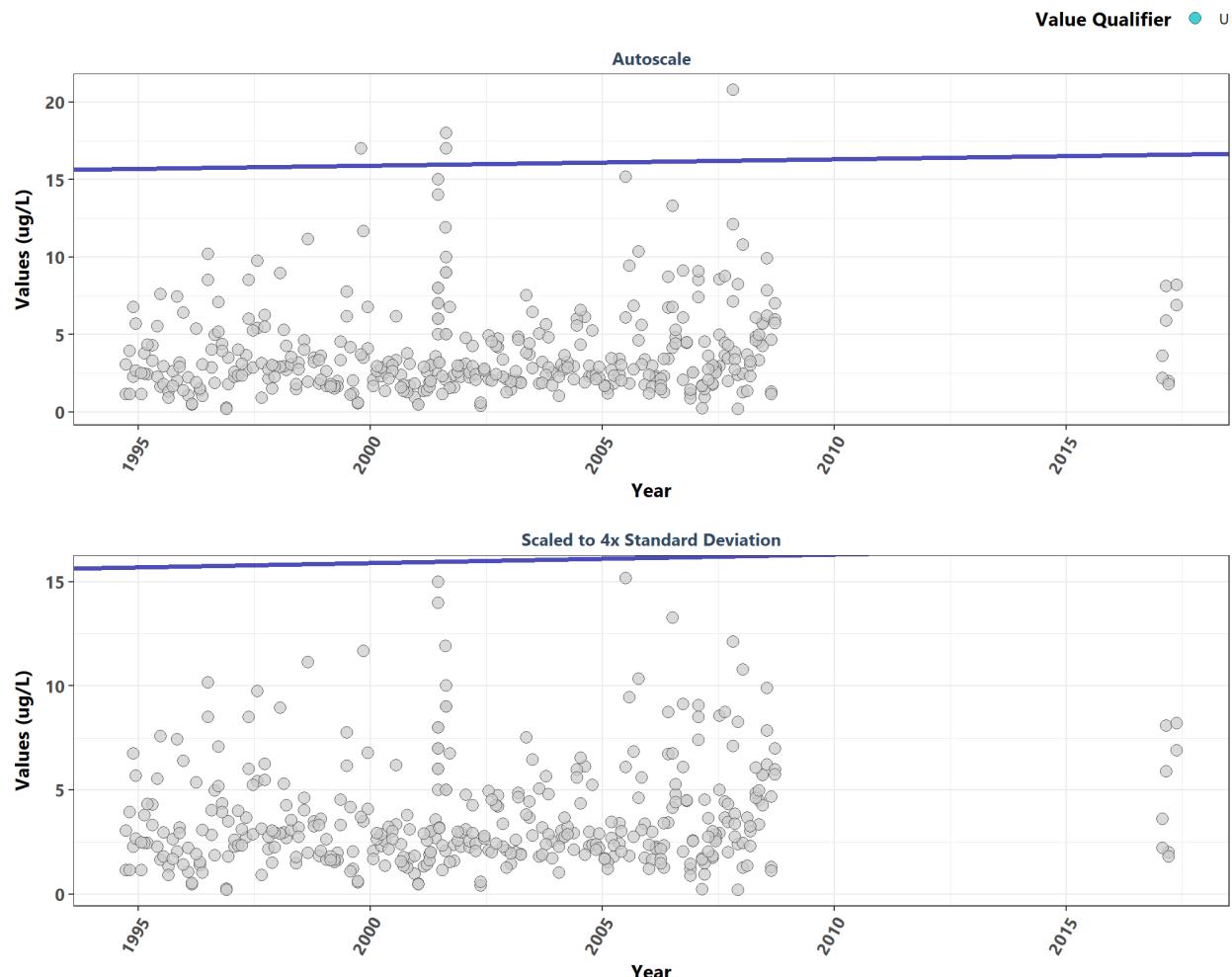
### Biscayne Bay Aquatic Preserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
452	1.315	0.0485	0.01	24.43061	0.1212	0

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

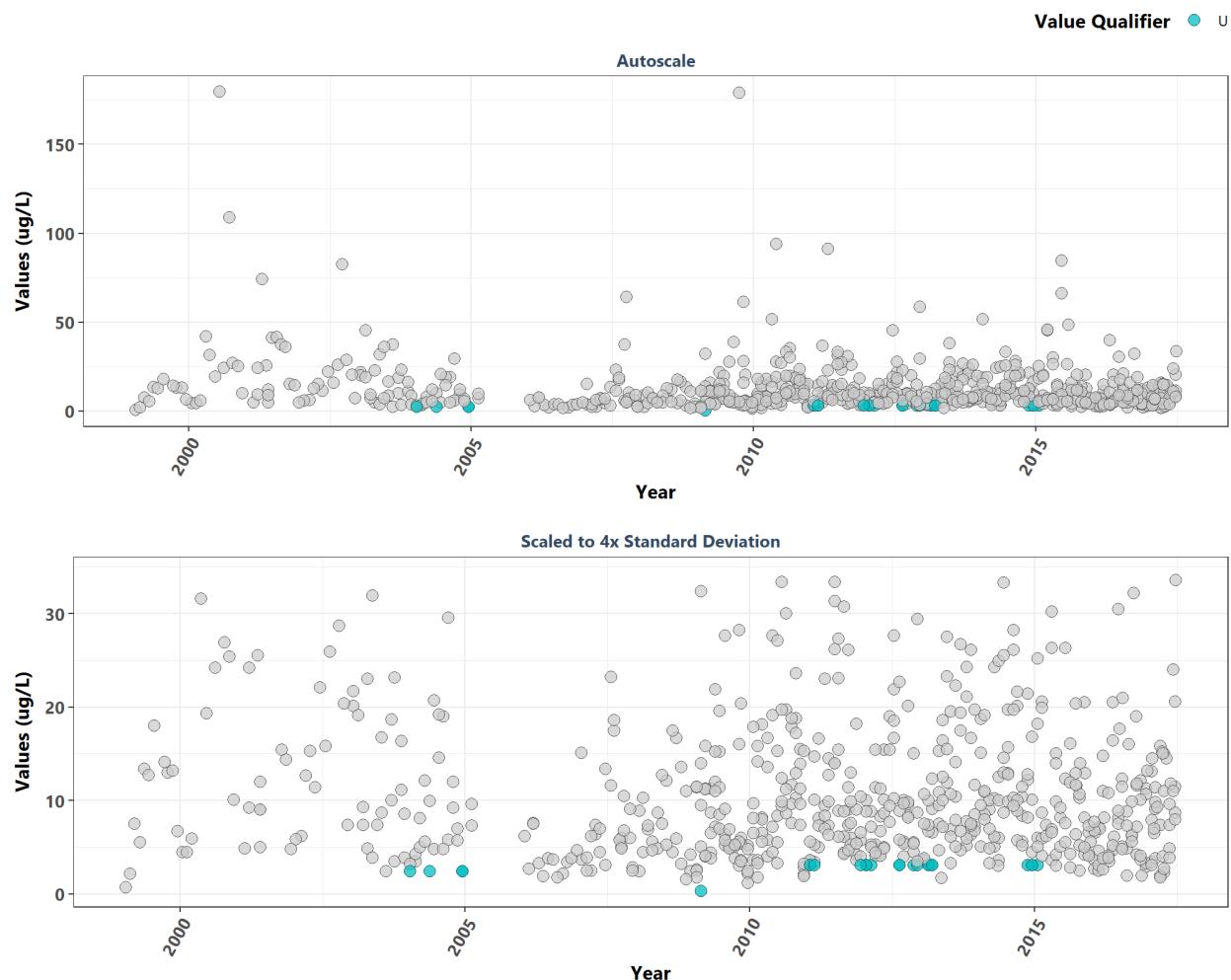
### Cape Romano-Ten Thousand Islands Aquatic Preserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
393	2.9329	0.075	0.04125	-66.60633	0.0475	1

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

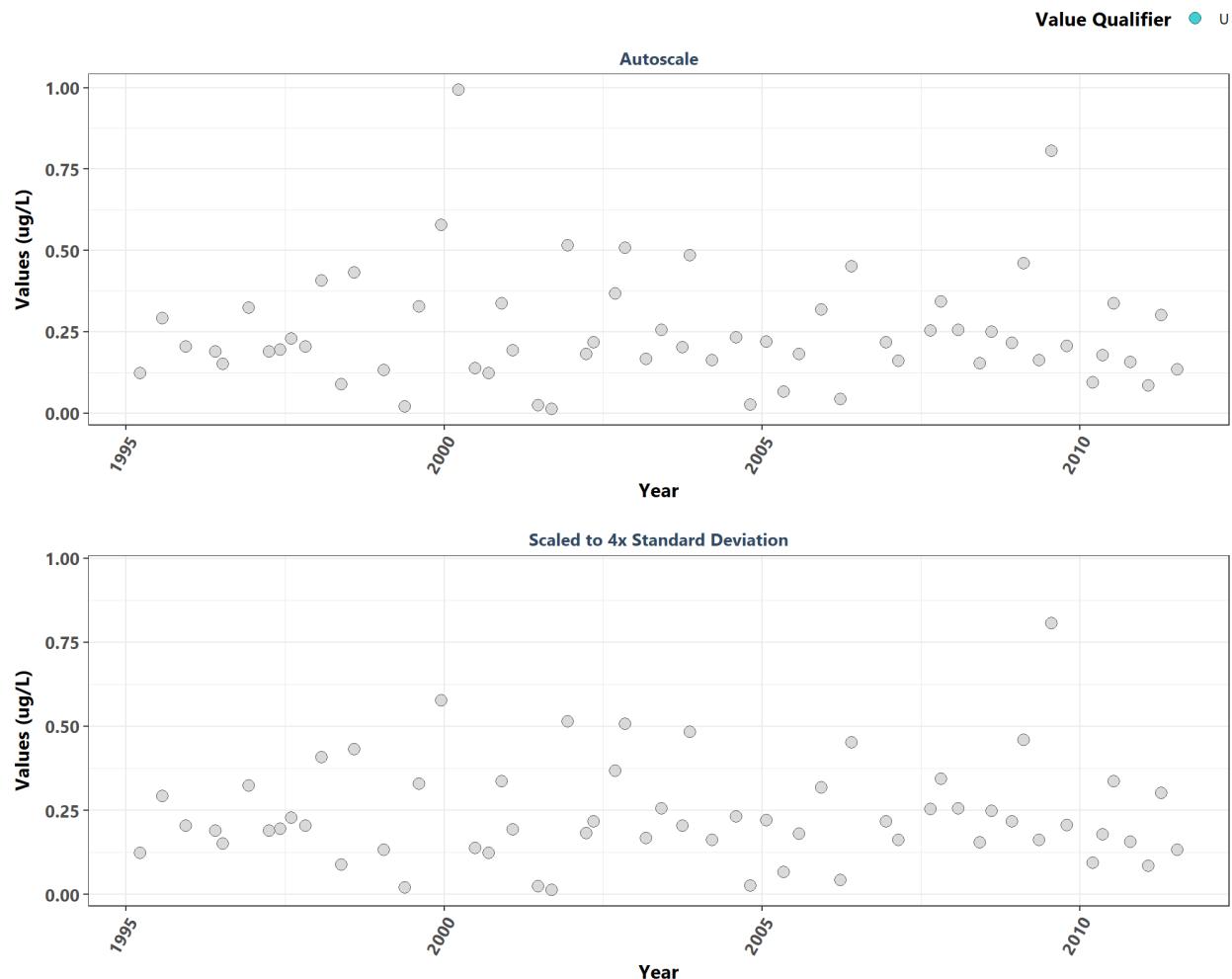
### Cockroach Bay Aquatic Preserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
675	8.7	-0.0204	-0.0333333	308.2219	0.4646	0

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

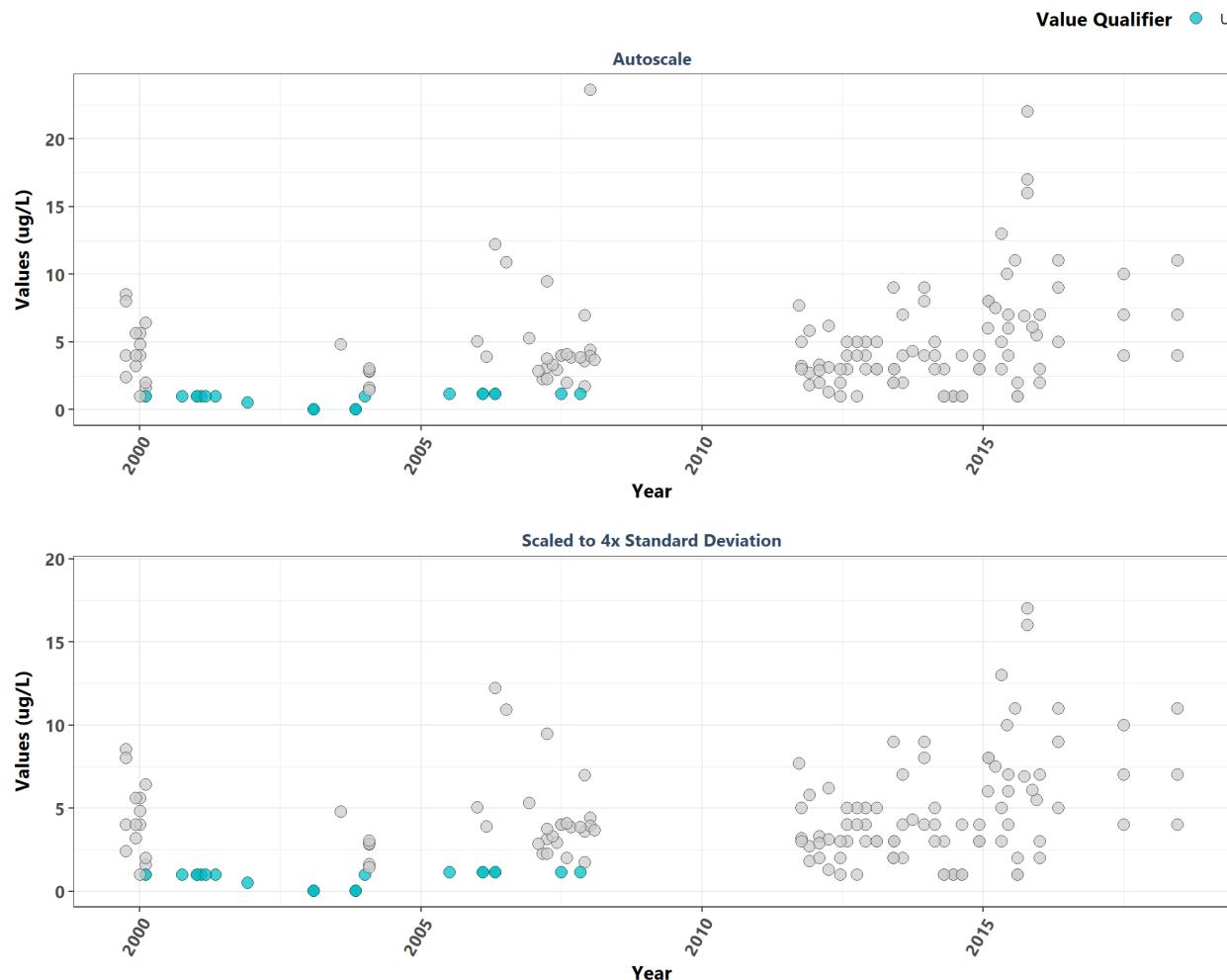
### Coupon Eight Aquatic Preserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
61	0.20439494	-0.0294	-0.0036611	1.237638	0.3537	0

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

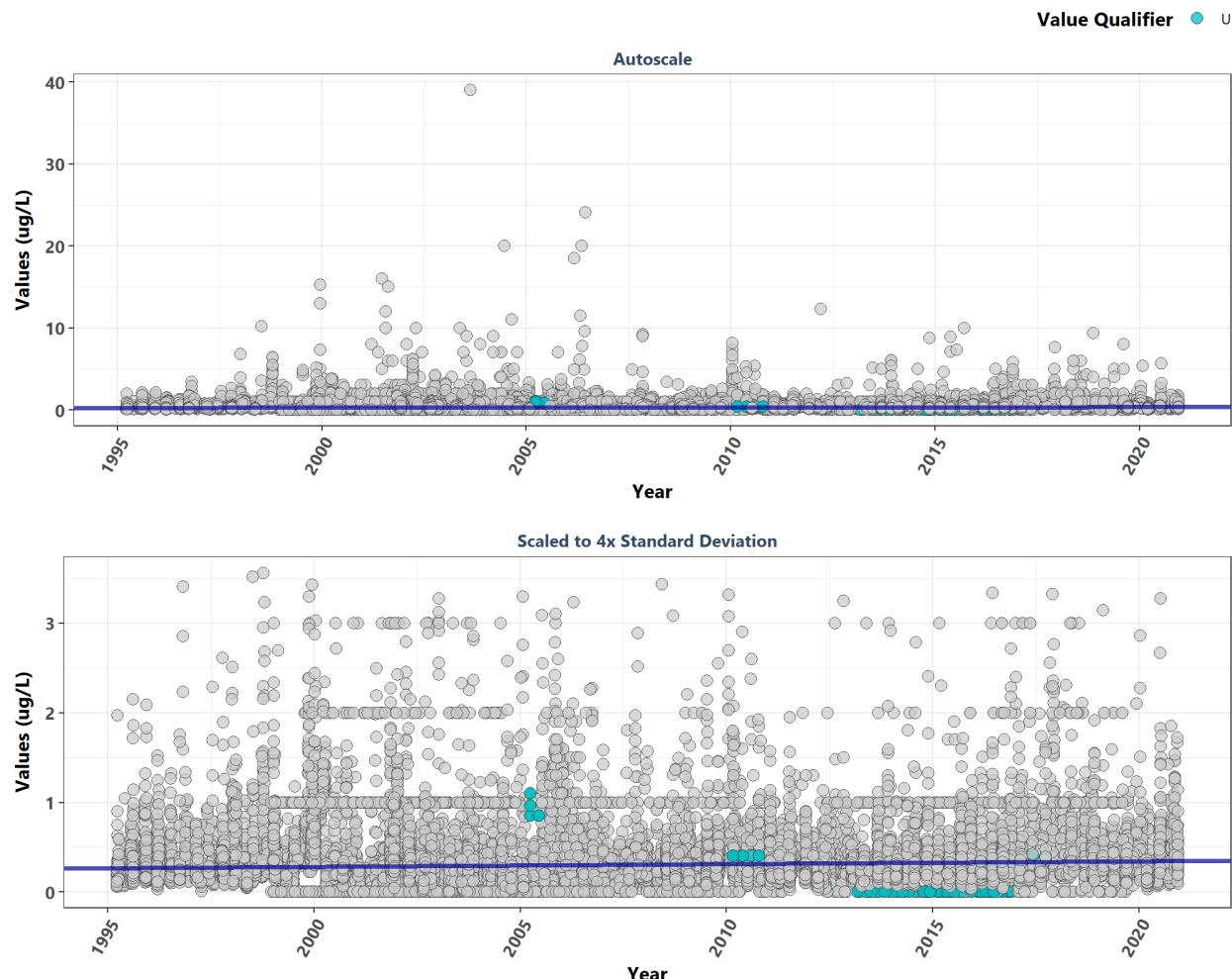
### Estero Bay Aquatic Preserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
144	3.12	0.3399	0.3076923	-697.6306	0	1

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

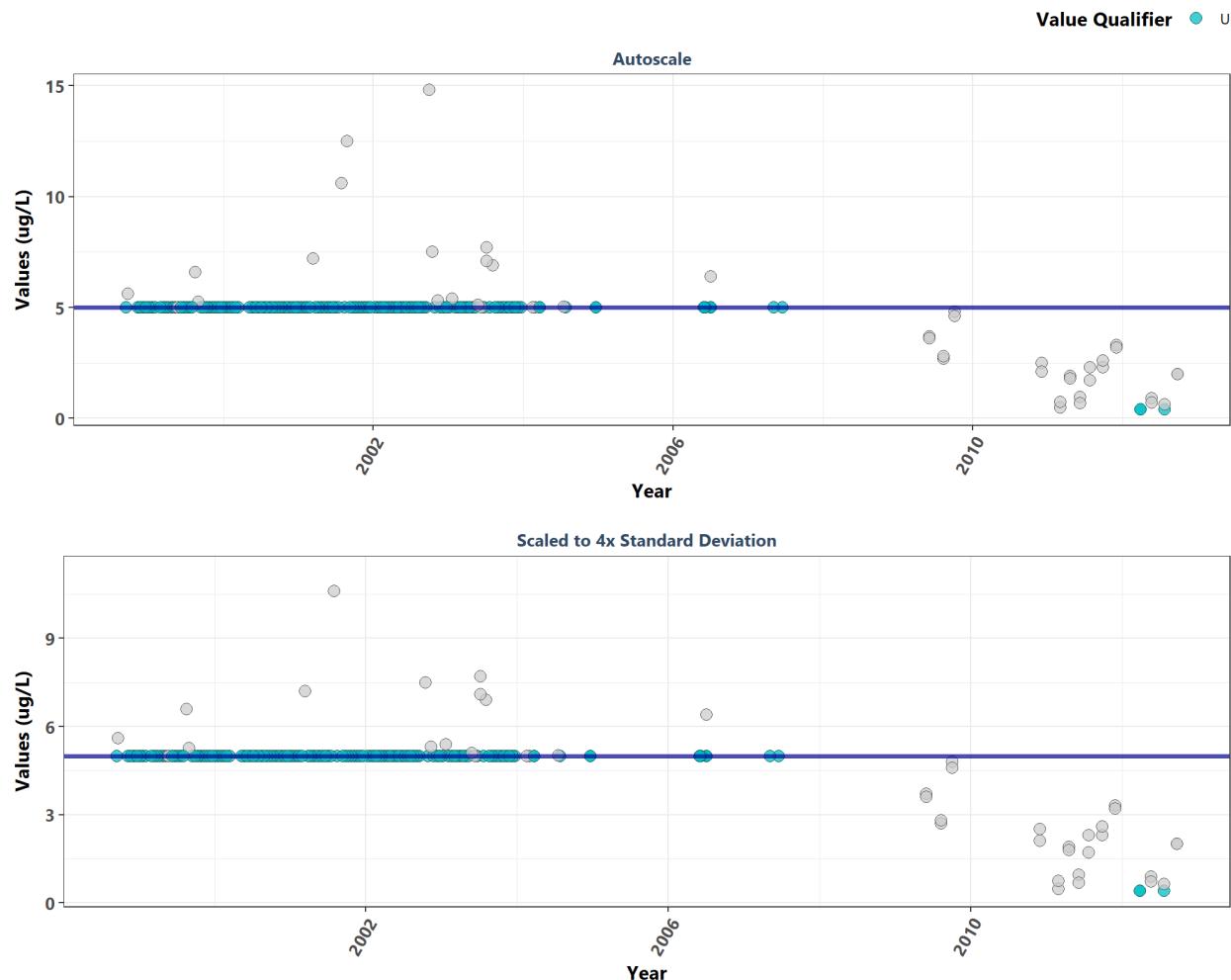
## Florida Keys National Marine Sanctuary



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
15978	0.26577653	0.0669	0.0030667	-5.852509	0	1

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

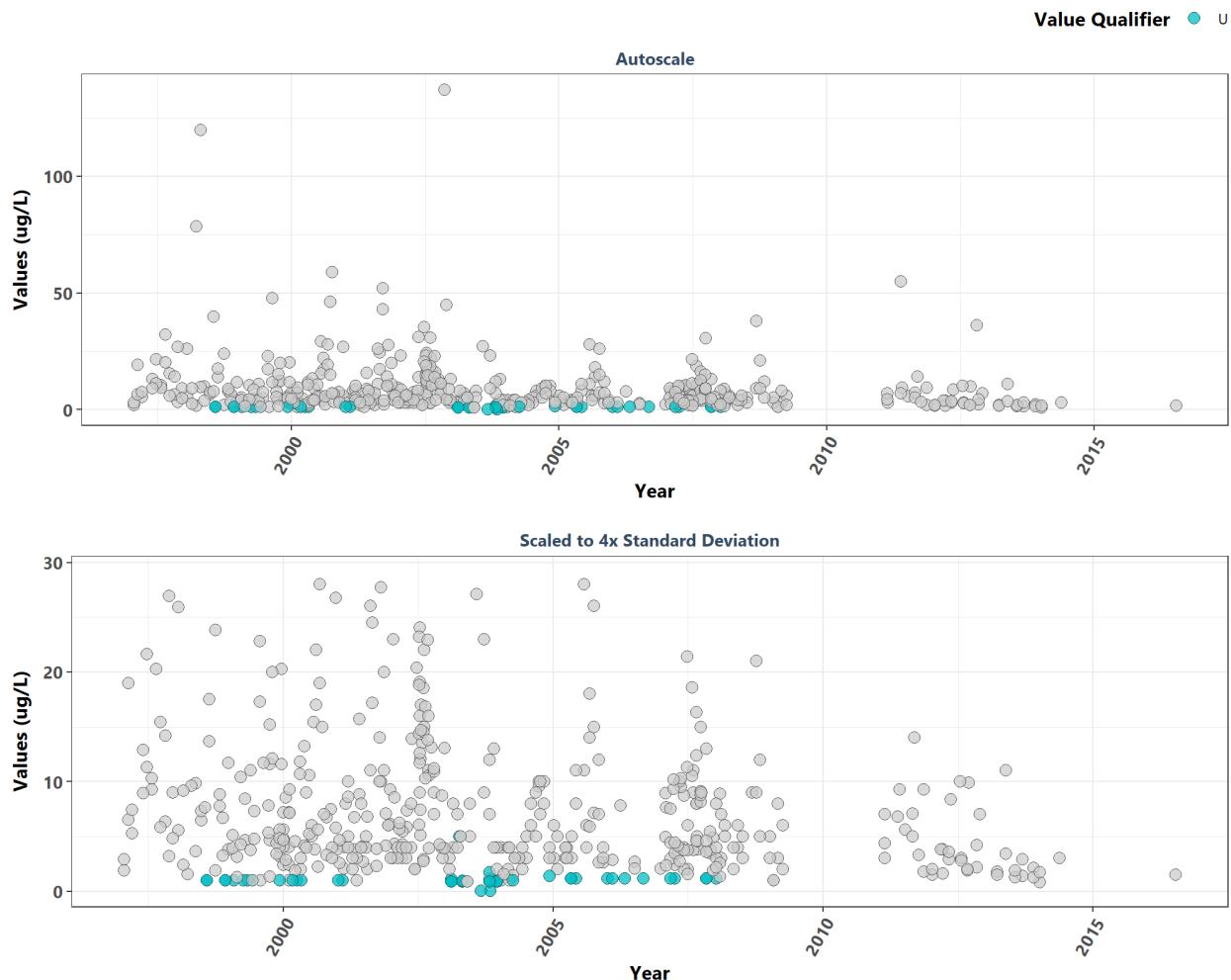
### Fort Pickens State Park Aquatic Preserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
190	5	-0.2417	0	5	0	-1

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

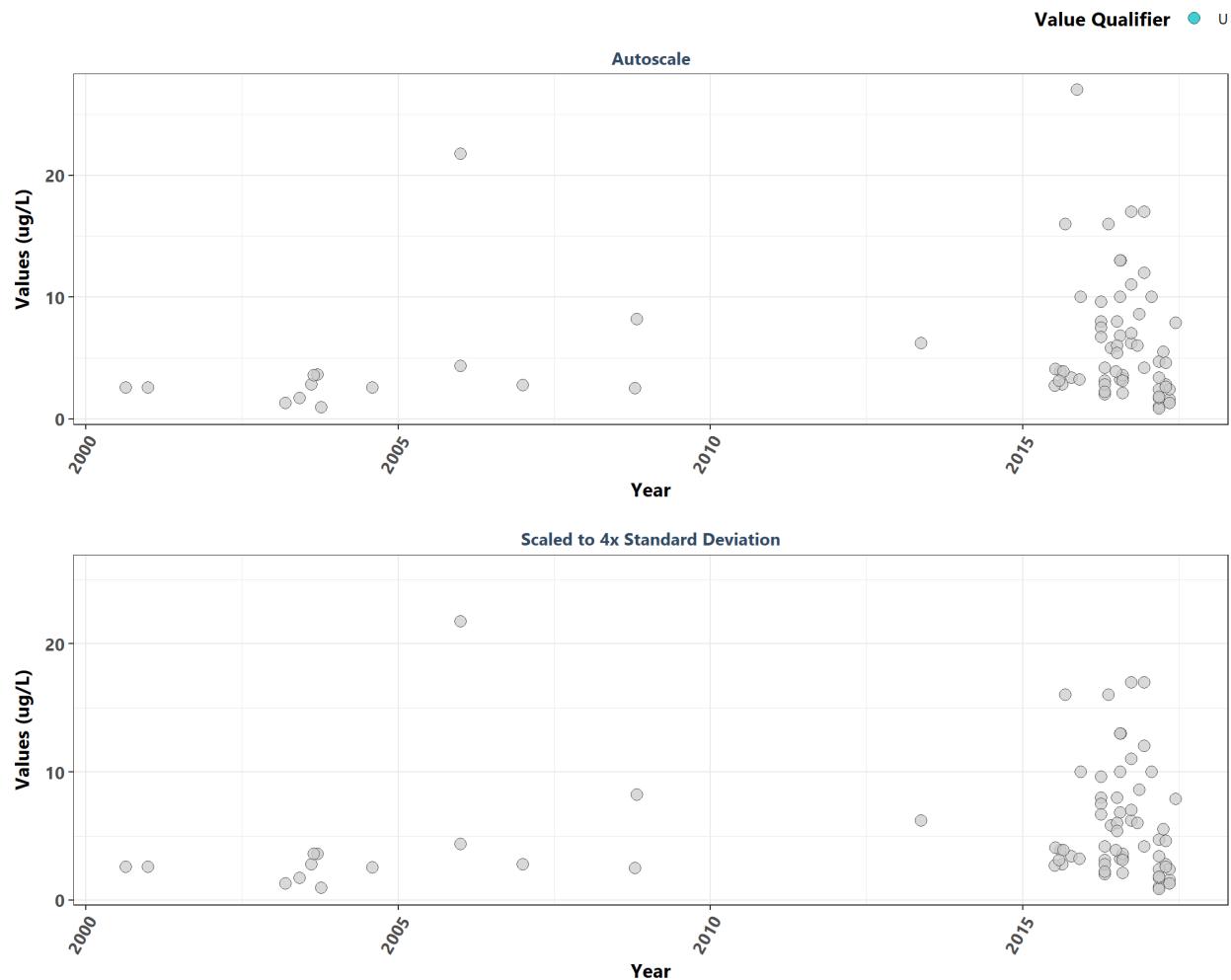
### Gasparilla Sound-Charlotte Harbor Aquatic Preserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
504	5	-0.1581	-0.25	312.8835	0	-1

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

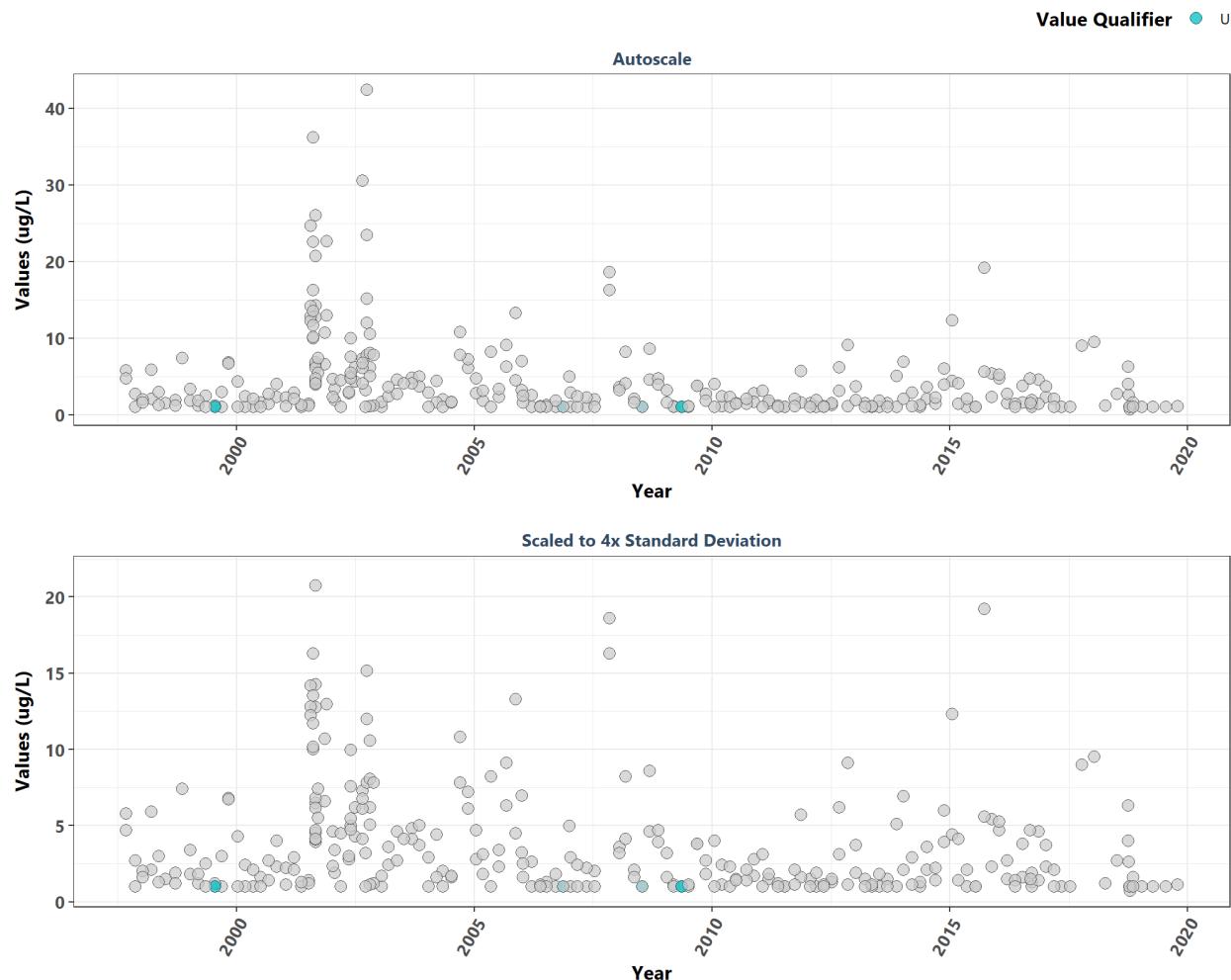
### Indian River-Malabar to Vero Beach Aquatic Preserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
77	3.61	0.1367	0.0384615	-234.1793	0.2513	0

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

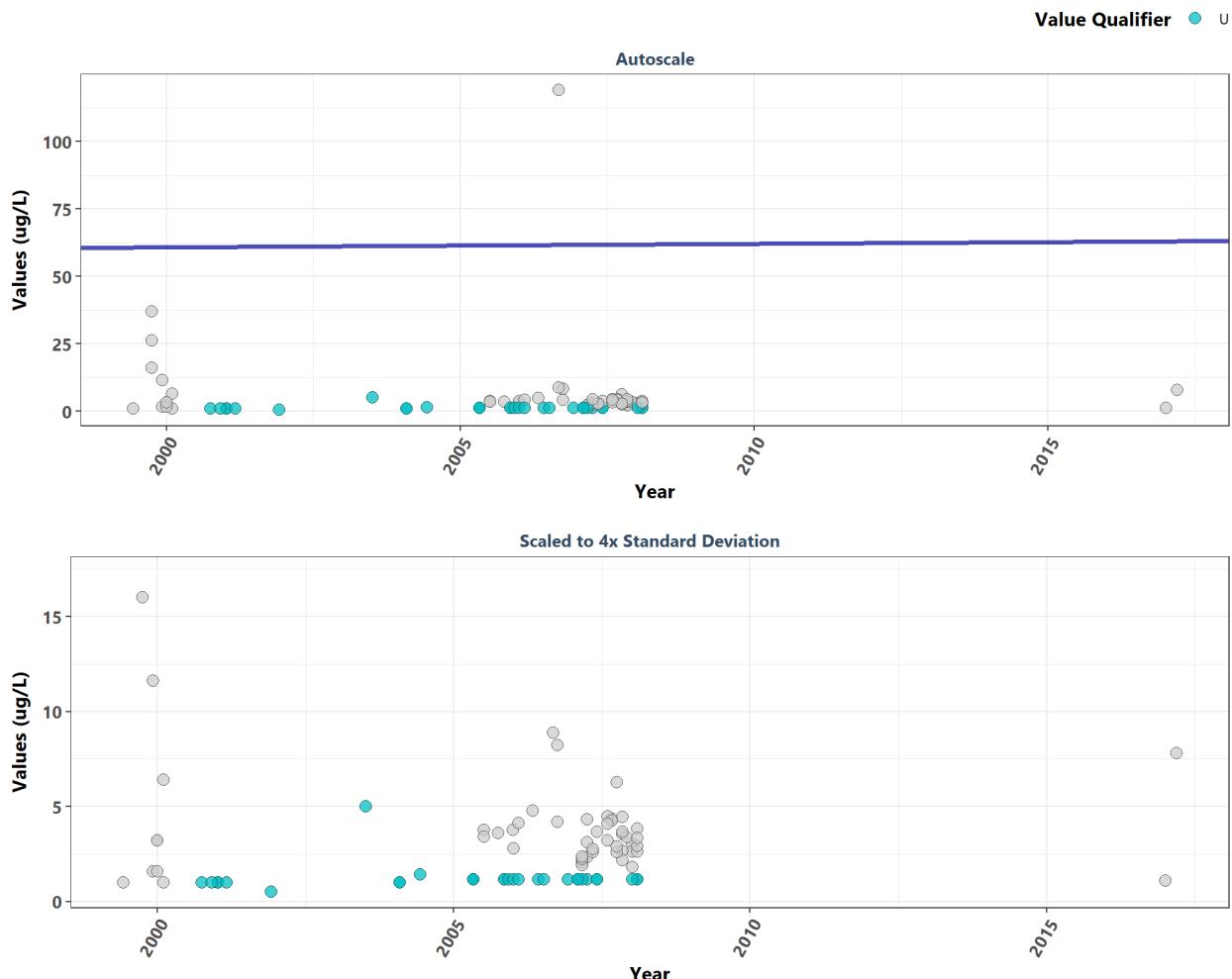
### Jensen Beach to Jupiter Inlet Aquatic Preserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
301	2.4	-0.1403	-0.0285714	164.465	0.0019	-1

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

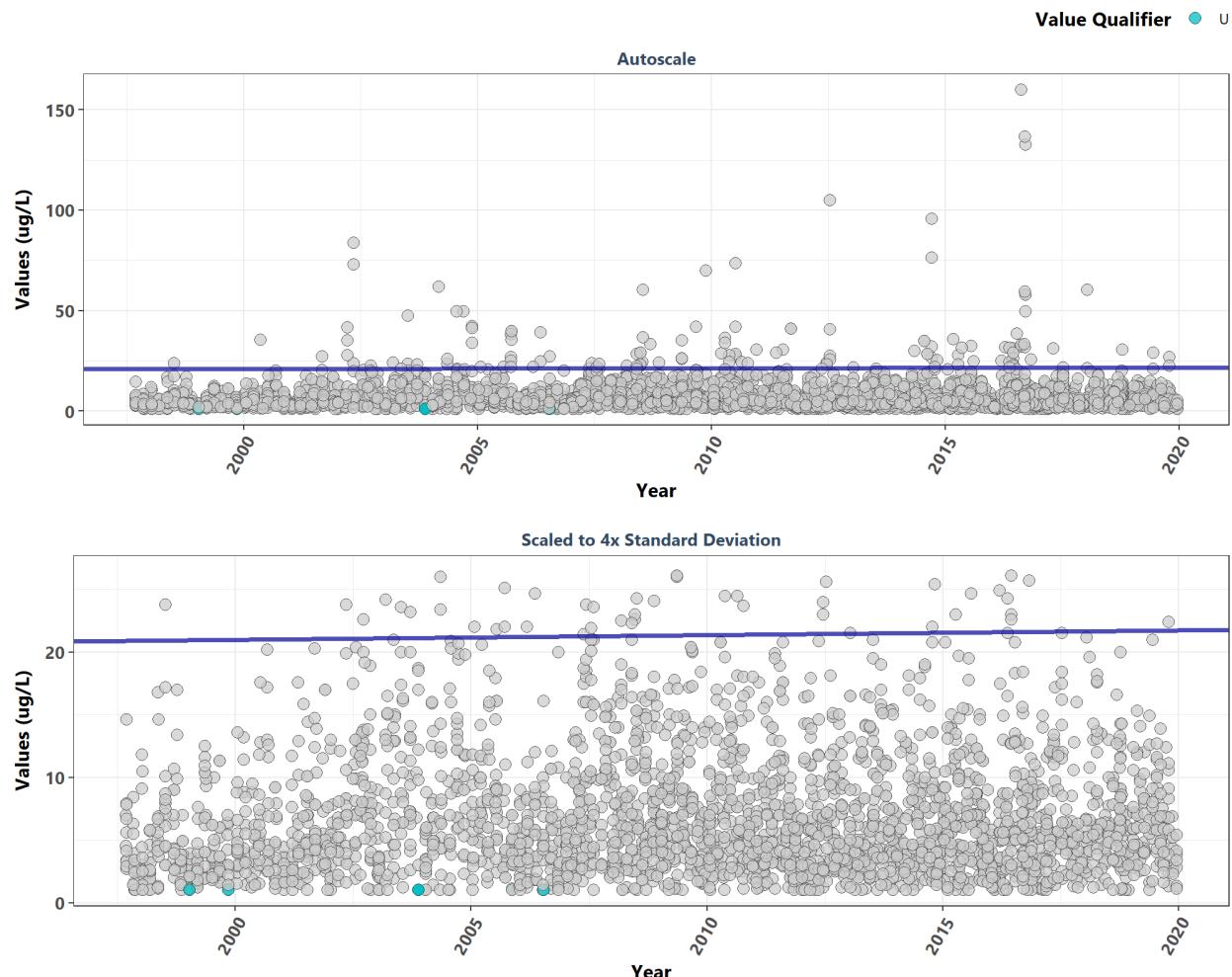
### Lemon Bay Aquatic Preserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
76	2.345	0.236	0.131	-201.2312	0.0012	1

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

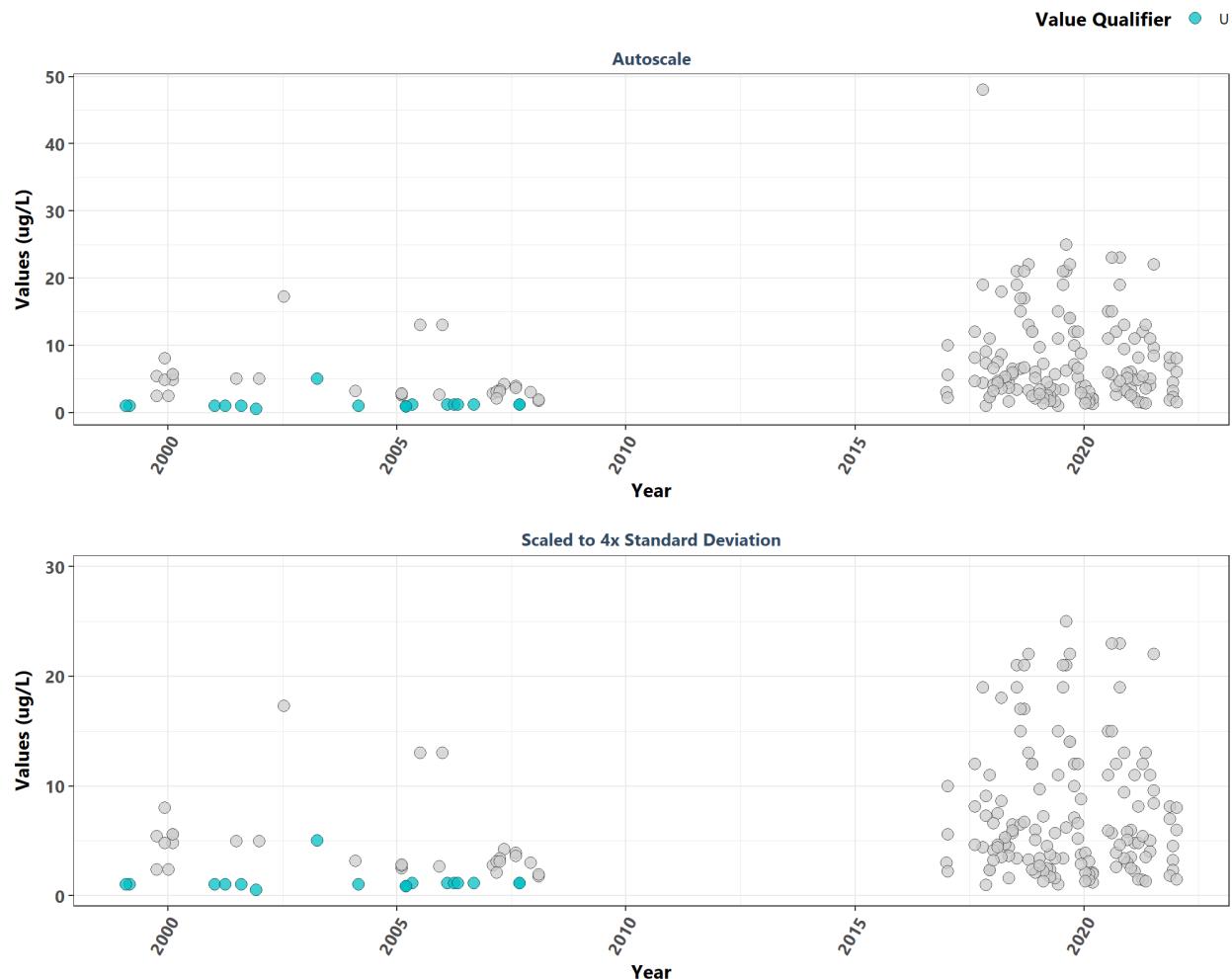
### Loxahatchee River-Lake Worth Creek Aquatic Preserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
3138	5.6	0.0286	0.0375	-54.025	0.0008	1

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

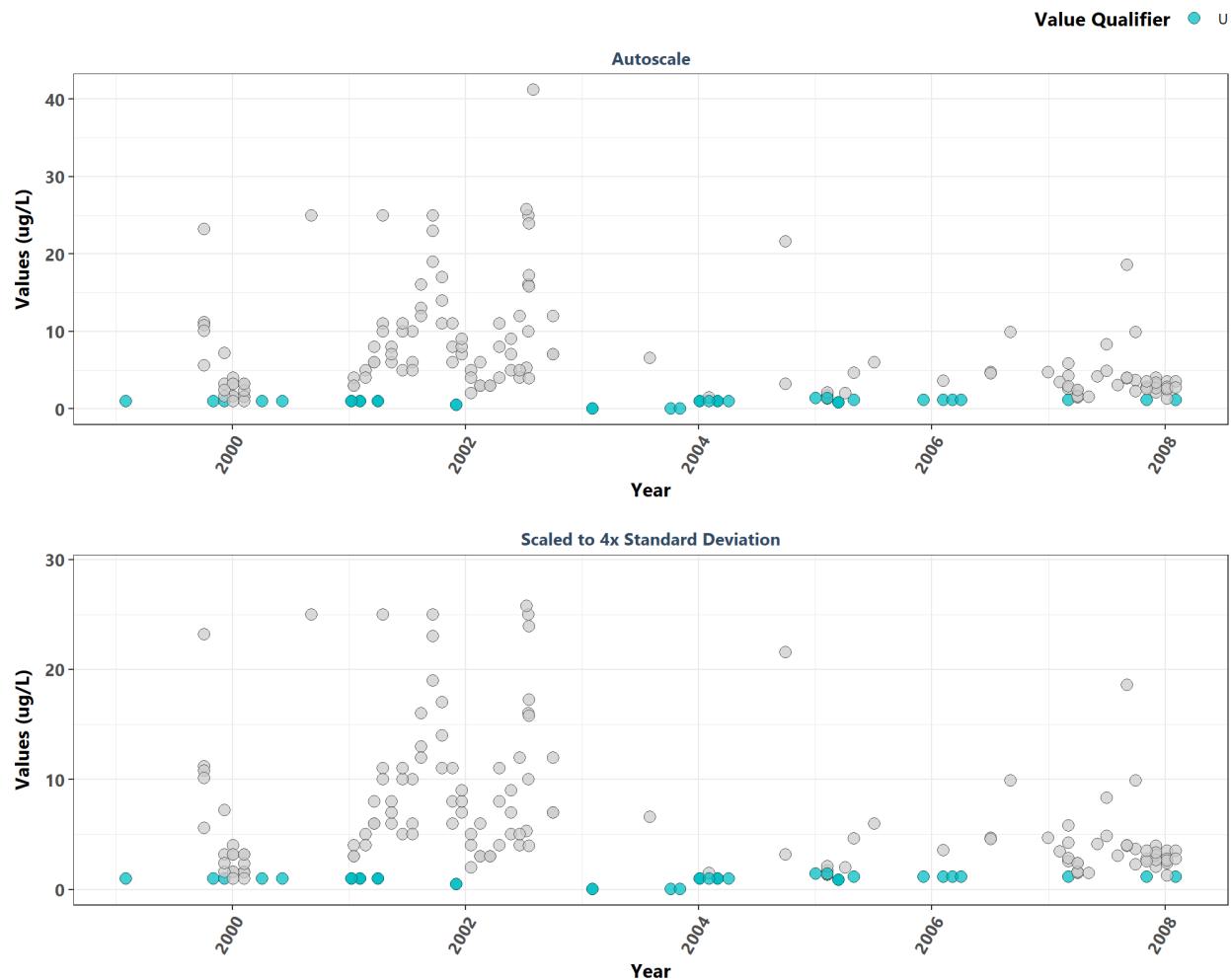
### Matlacha Pass Aquatic Preserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
183	4.6	0.1472	0.0941176	-97.4225	0.0072	1

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

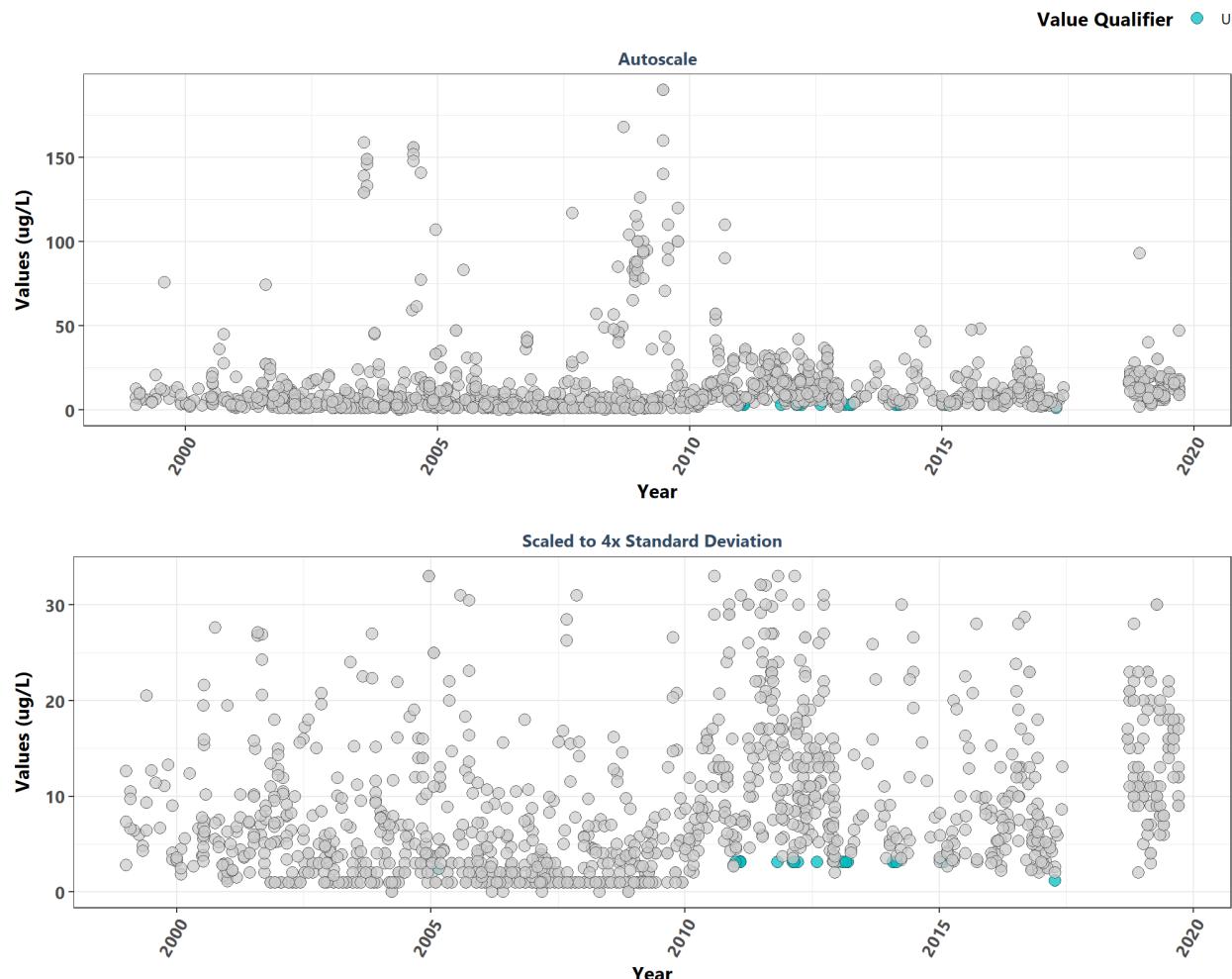
### Pine Island Sound Aquatic Preserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
160	3.485	-0.1362	-0.08	1000.832	0.2533	0

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

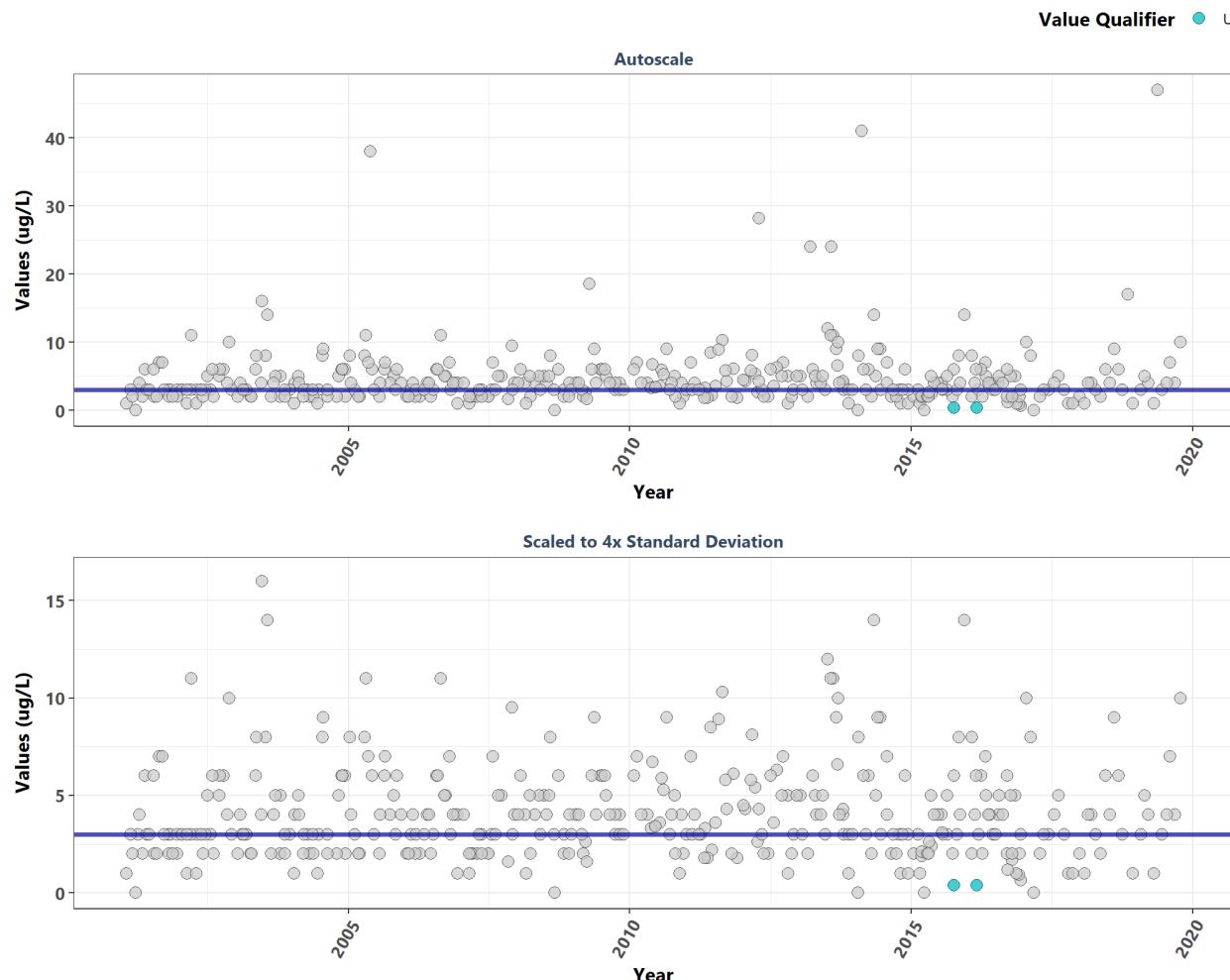
### Pinellas County Aquatic Preserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
1291	7	0.1651	0.25	-551.4989	0	1

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

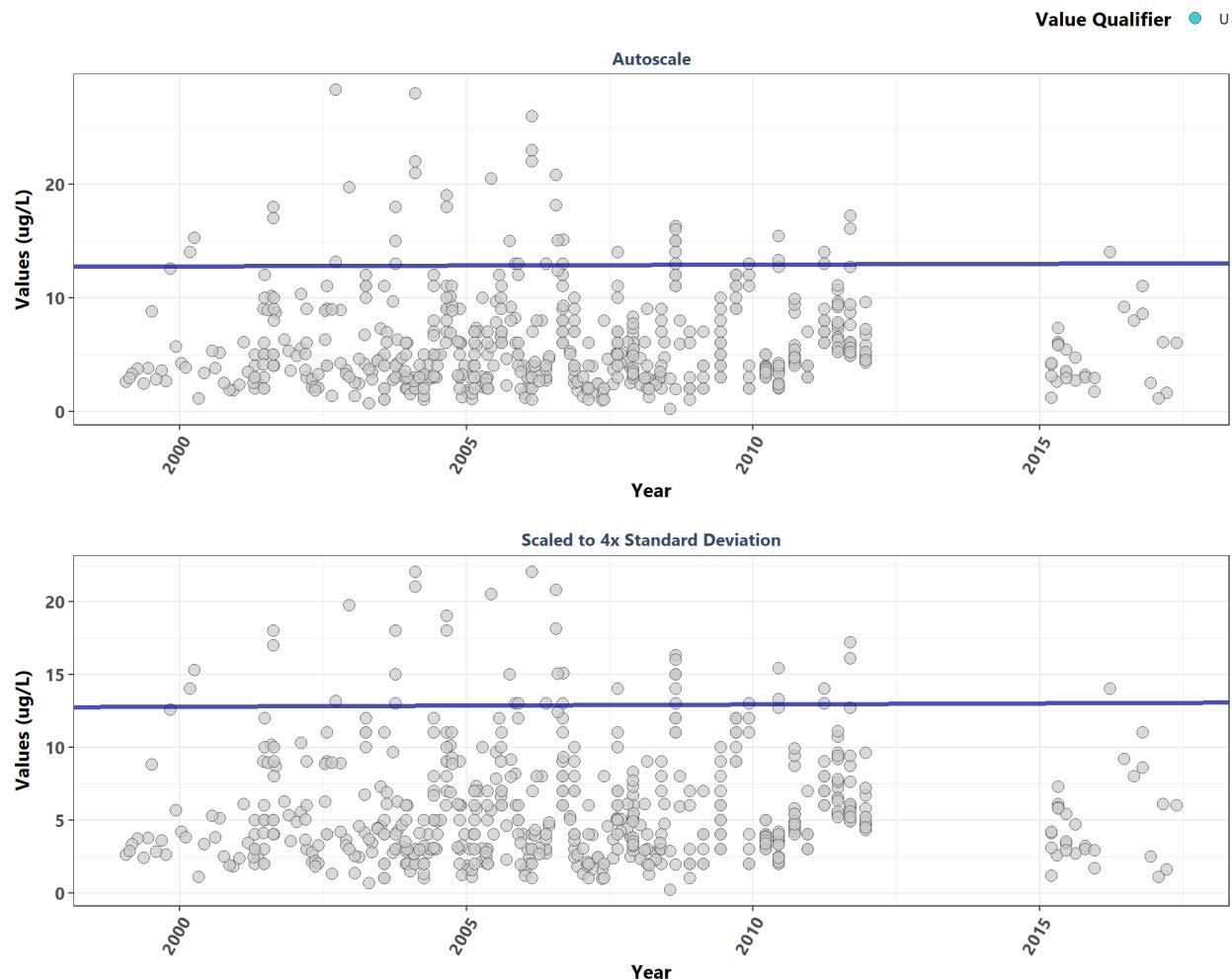
### Rocky Bayou State Park Aquatic Preserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
386	3.5	0.0235	0	3	0.4879	0

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

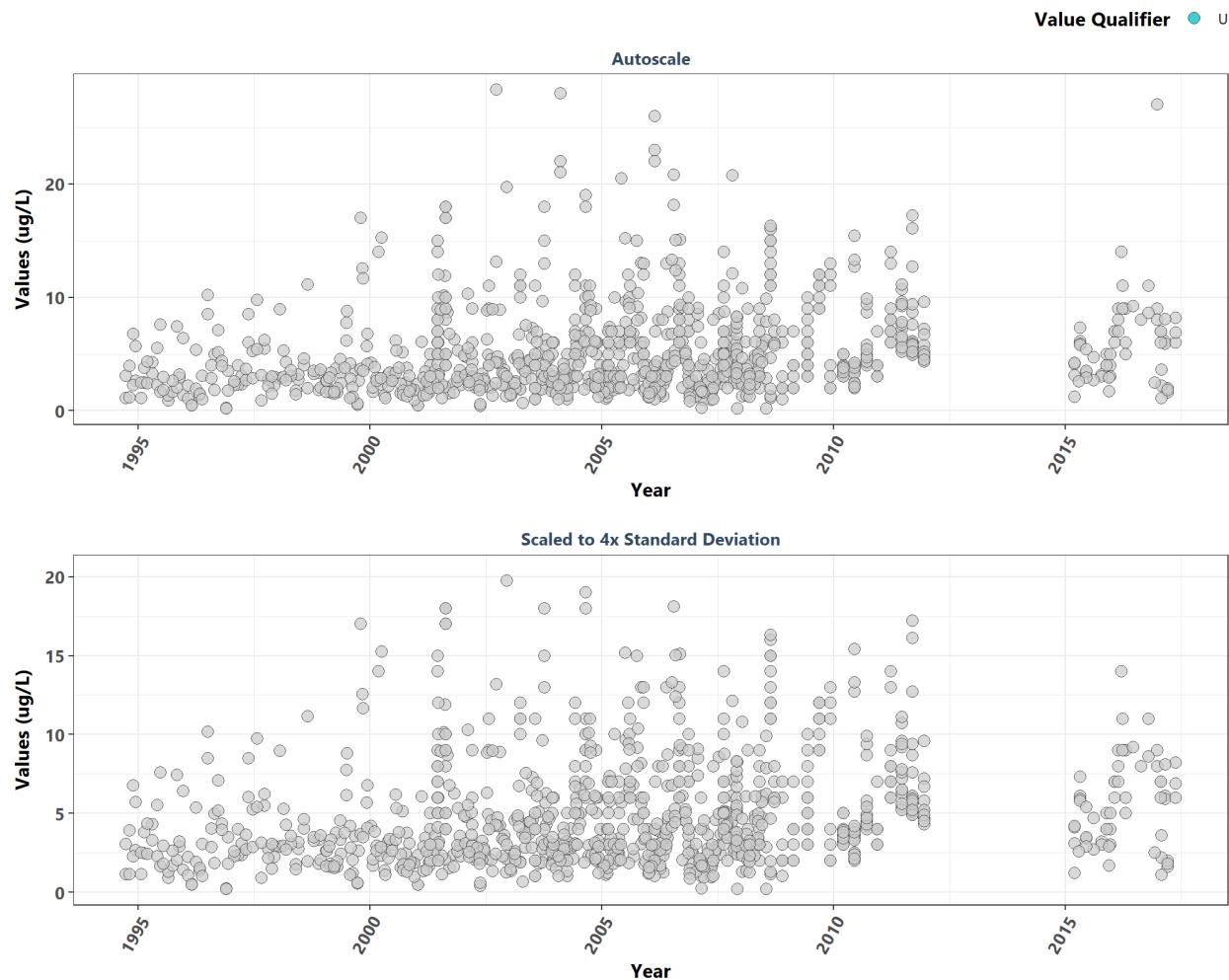
### Rookery Bay Aquatic Preserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
604	4.8343	0.023	0.01585	-18.92333	0.2752	0

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

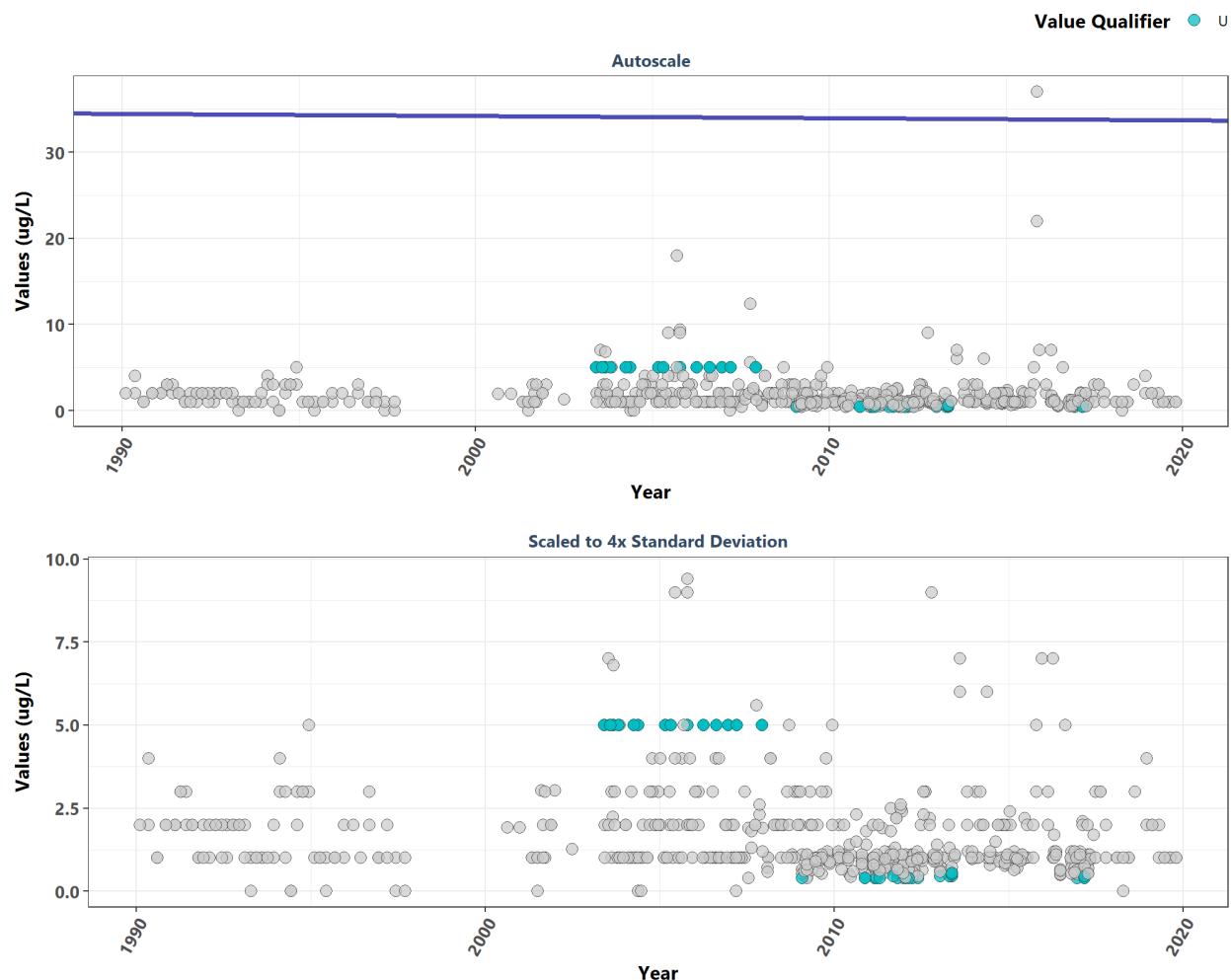
## Rookery Bay National Estuarine Research Reserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
1021	4	0.1601	0.1241333	-217.0174	0	1

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

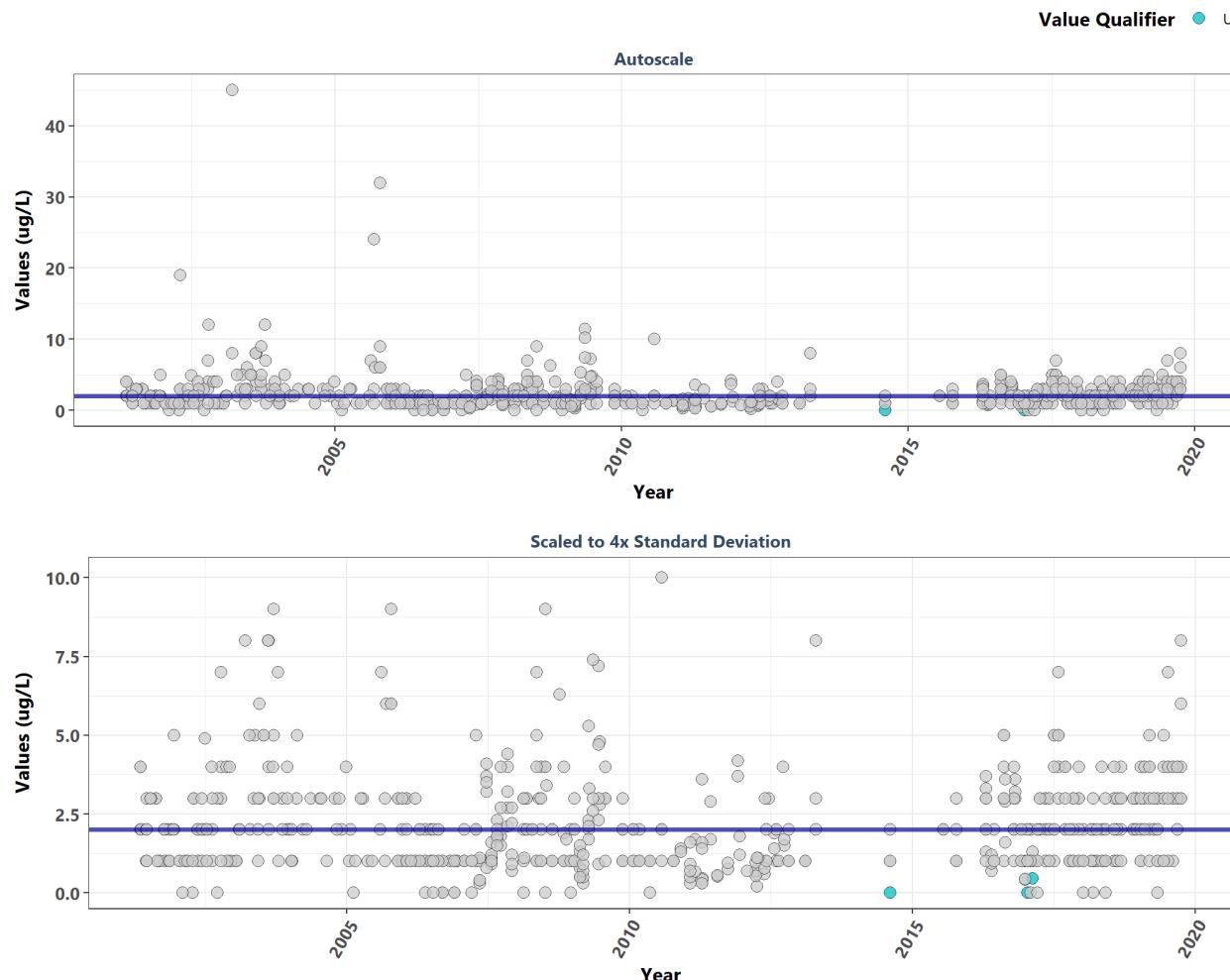
### St. Andrews State Park Aquatic Preserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
670	1	-0.1862	-0.025	84.18718	0	-1

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

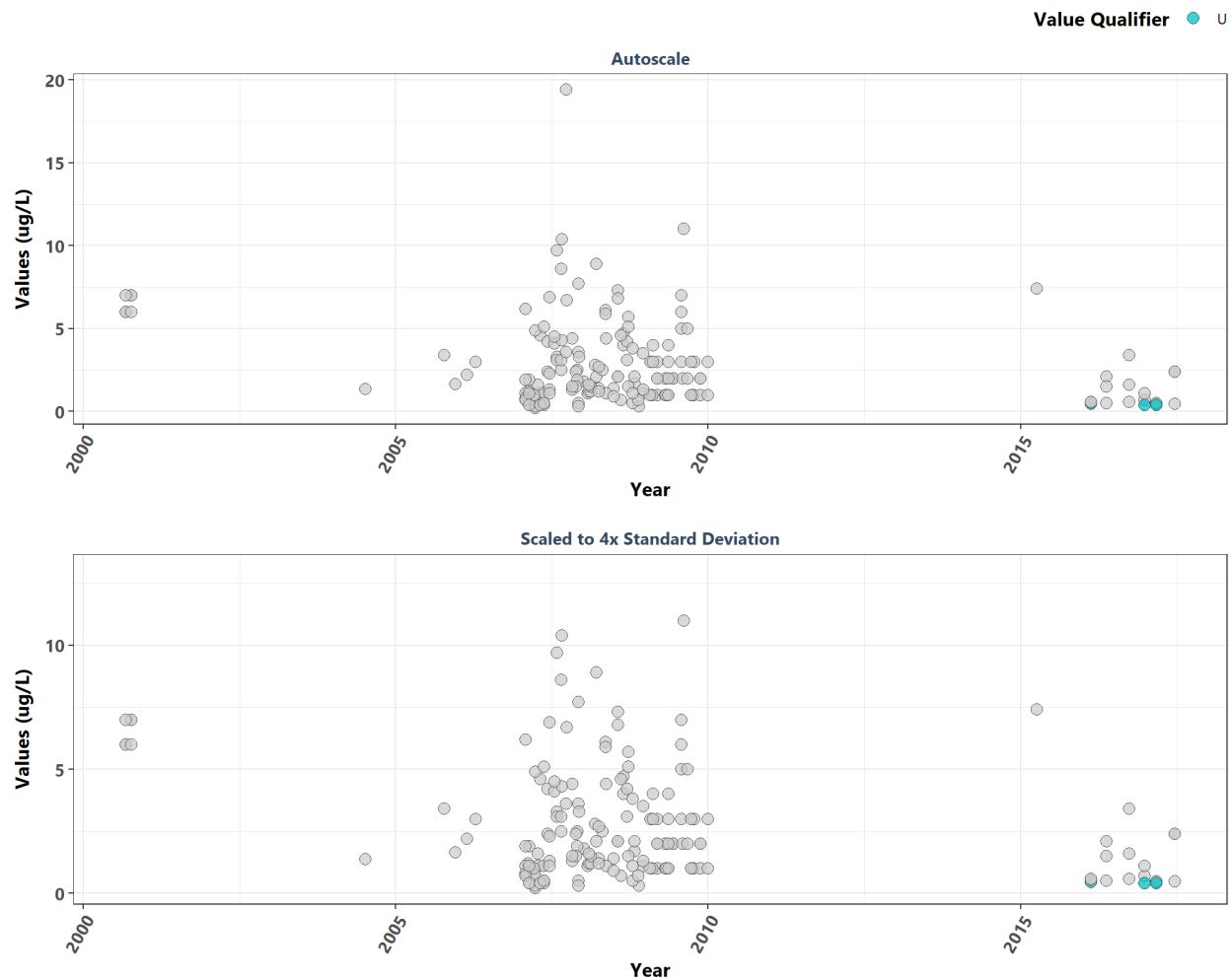
### St. Joseph Bay Aquatic Preserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
698	2	0.0321	0	2	0.2778	0

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

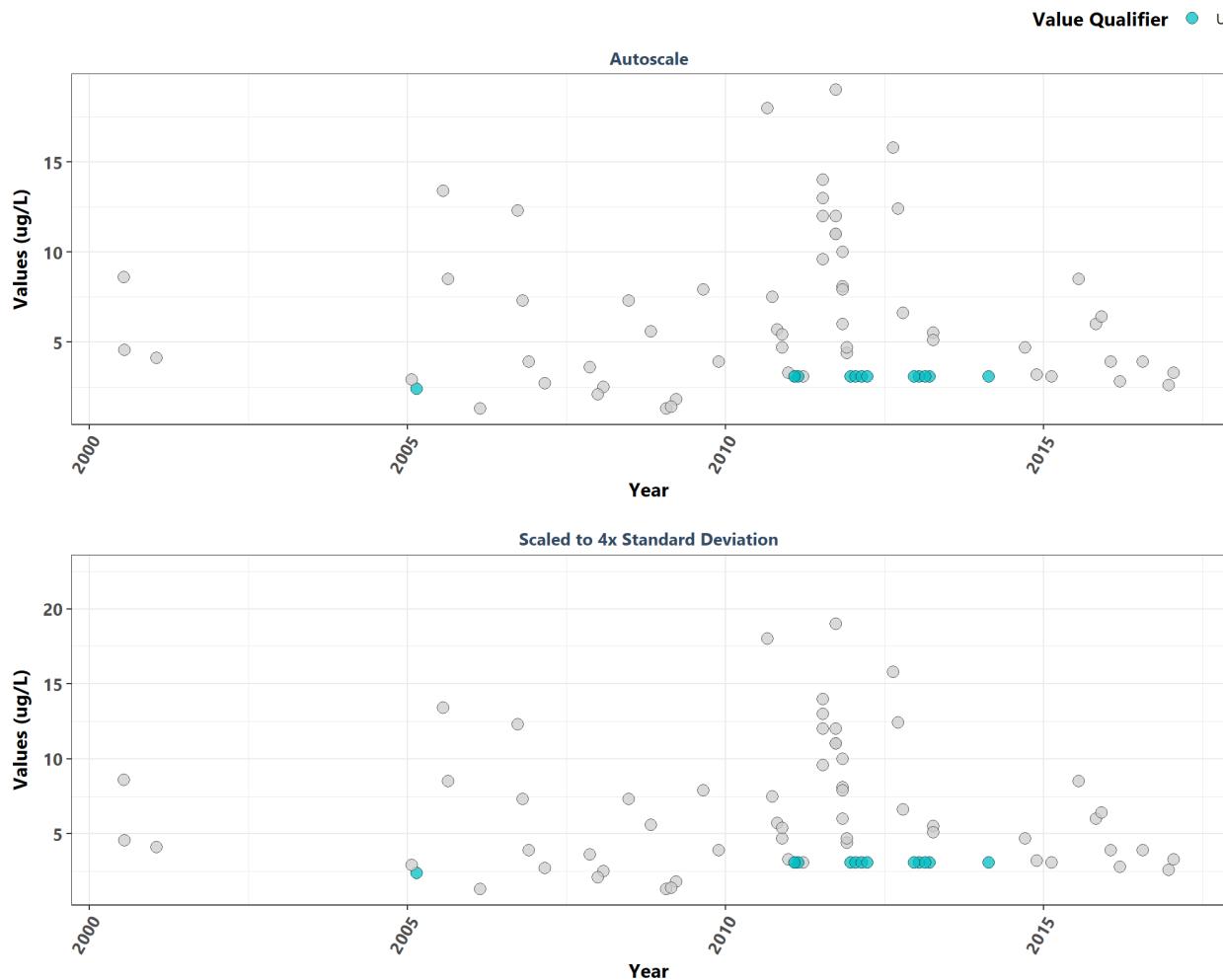
### St. Martins Marsh Aquatic Preserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
175	2	-0.1281	-0.1118056	134.128	0.0086	-1

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

### Terra Ceia Aquatic Preserve



N_Data	Median	tau	SennSlope	SennIntercept	p	Trend
71	4.563	0.1677	0.0571429	3.1	0.0407	1

<sup>a</sup> p < 0.00005 appear as 0 due to rounding

## 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$SufficientData==TRUE &
                      data$ManagedAreaName==MA_Include[i],]
    year_lower <- min(plot_data$Year)
    year_upper <- max(plot_data$Year)
    min_RV <- min(plot_data$ResultValue)
    mn_RV <- mean(plot_data$ResultValue[plot_data$ResultValue <
                                             quantile(data$ResultValue, 0.98)])
    sd_RV <- sd(plot_data$ResultValue[plot_data$ResultValue <
                                             quantile(data$ResultValue, 0.98)])
    x_scale <- ifelse(year_upper - year_lower > 30, 10, 5)
    y_scale <- mn_RV + 4 * sd_RV

    ##Year plots
    p1 <- ggplot(data=plot_data,
                  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,
                  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)

p0 <- ggplot() + labs(title=paste0(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)
YMset <- ggarrange(leg1, p4, p5 + theme(legend.position="none"), p6,
                    ncol=1, heights=c(0.1, 1, 1, 1))

p00 <- ggplot() + labs(title=paste0(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)
Mset <- ggarrange(leg2, p7, p8 + theme(legend.position="none"), p9,
                  ncol=1, heights=c(0.1, 1, 1, 1))

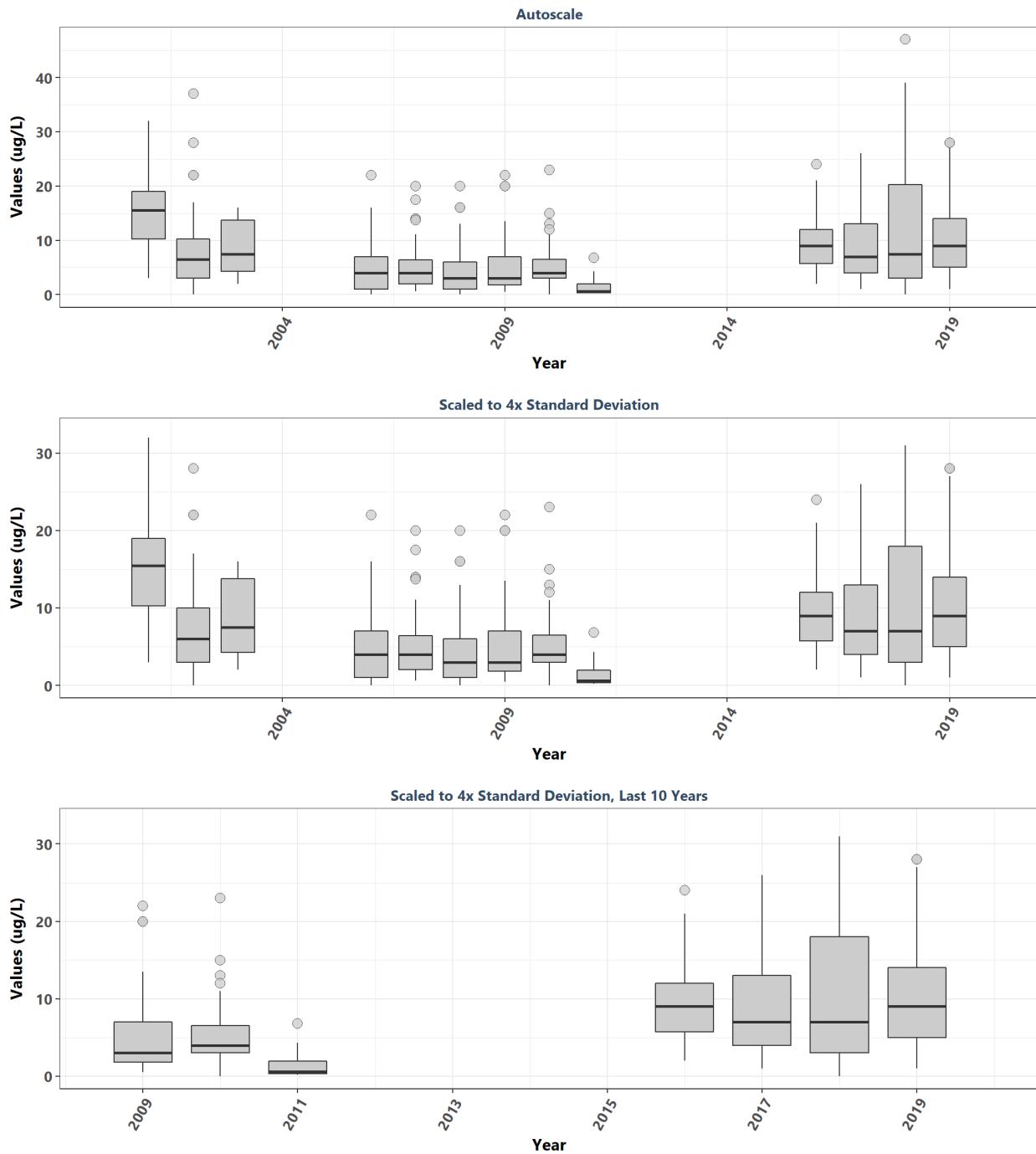
p000 <- ggplot() + labs(title=paste0(MA_Include[i]),
                         subtitle="By Month") + plot_theme +
  theme(panel.border=element_blank(),
        panel.grid.major=element_blank(),
        panel.grid.minor=element_blank(), axis.line=element_blank())

print(ggarrange(p0, Yset, ncol=1, heights=c(0.07, 1)))
print(ggarrange(p00, YMset, ncol=1, heights=c(0.07, 1)))
print(ggarrange(p000, Mset, ncol=1, heights=c(0.07, 1, 0.7)))

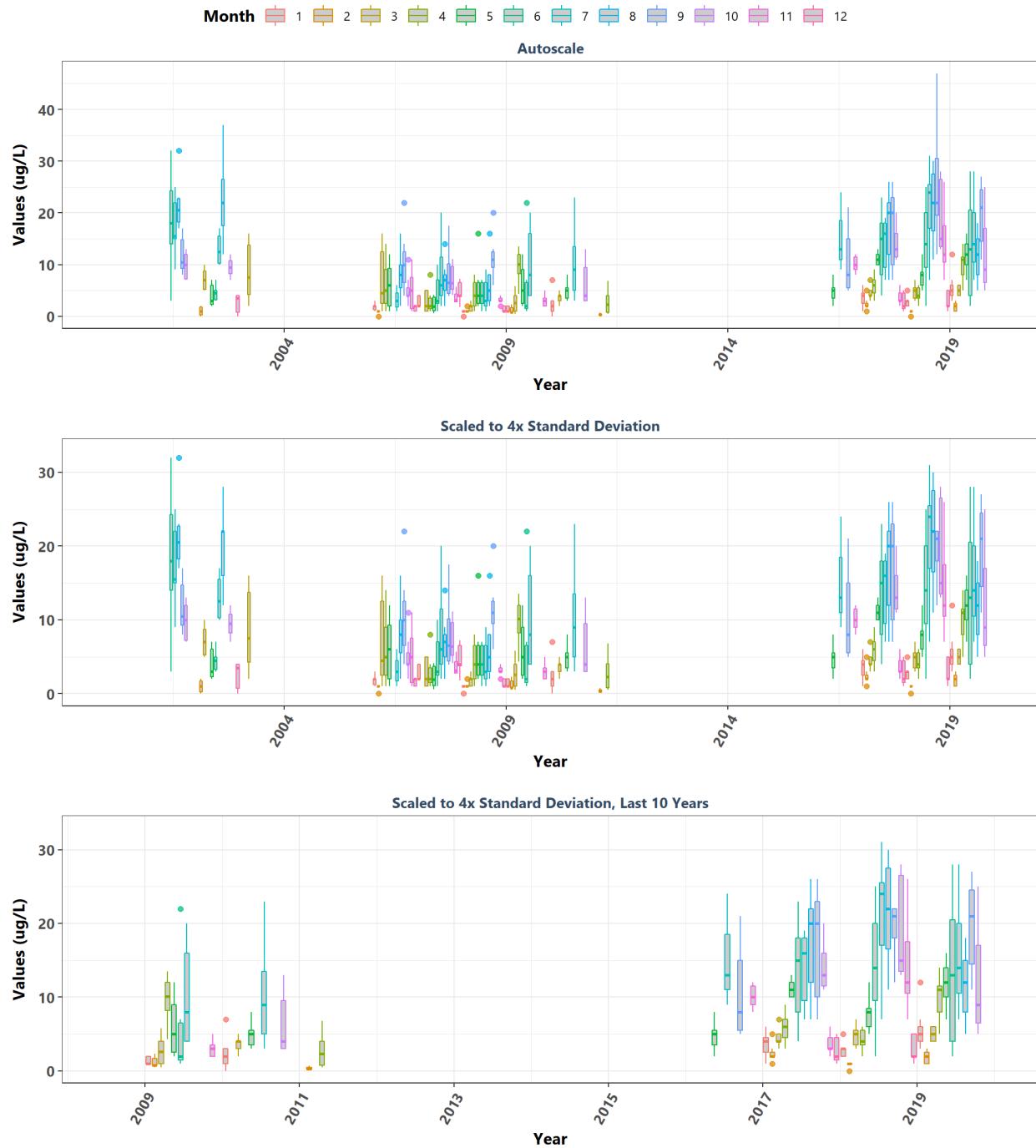
rm(plot_data)
rm(p1, p2, p3, p4, p5, p6, p7, p8, p9, p0, p00, p000, leg1, leg2,
  Yset, YMset, Mset)
}
}

```

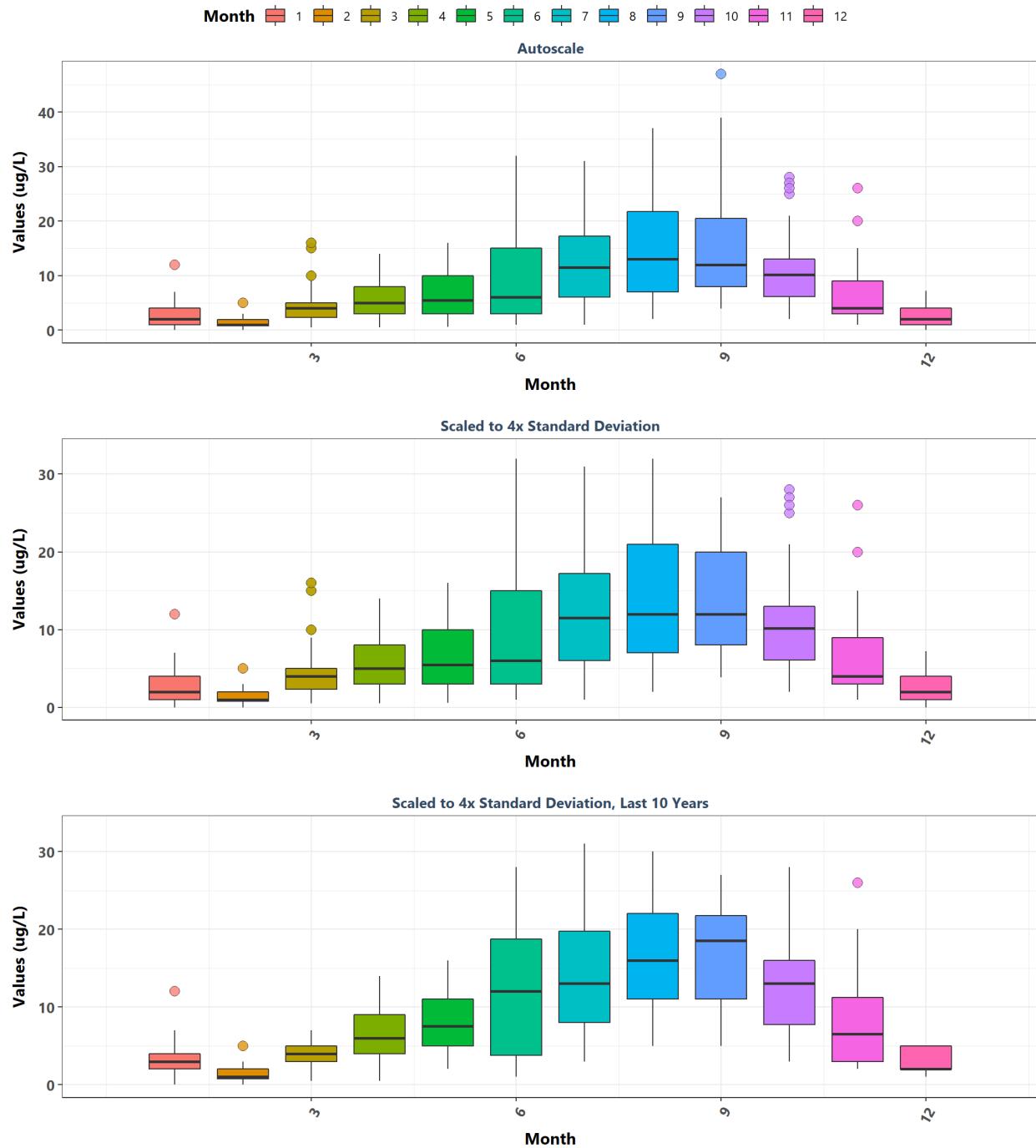
**Alligator Harbor Aquatic Preserve**  
By Year



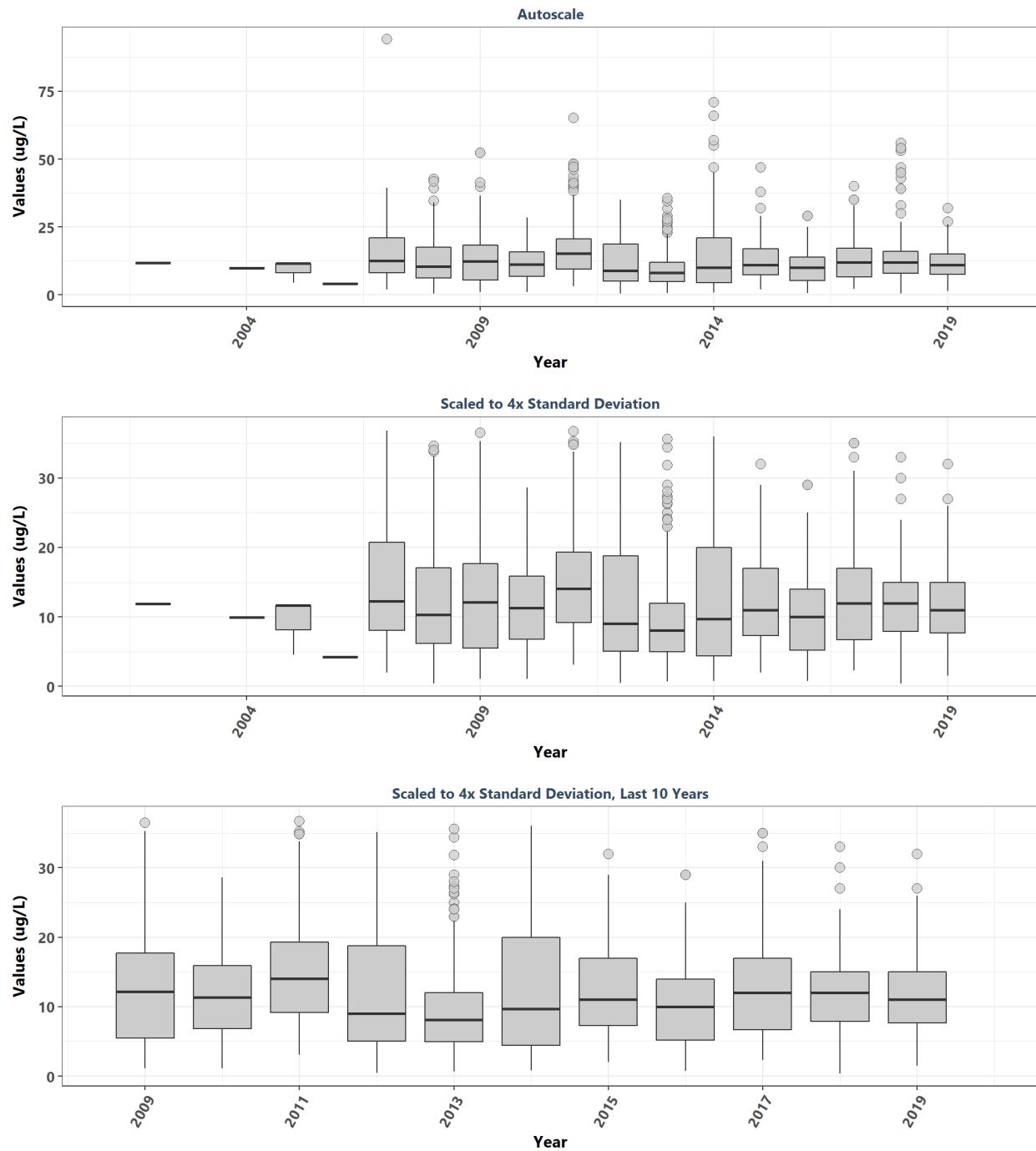
**Alligator Harbor Aquatic Preserve**  
By Year & Month



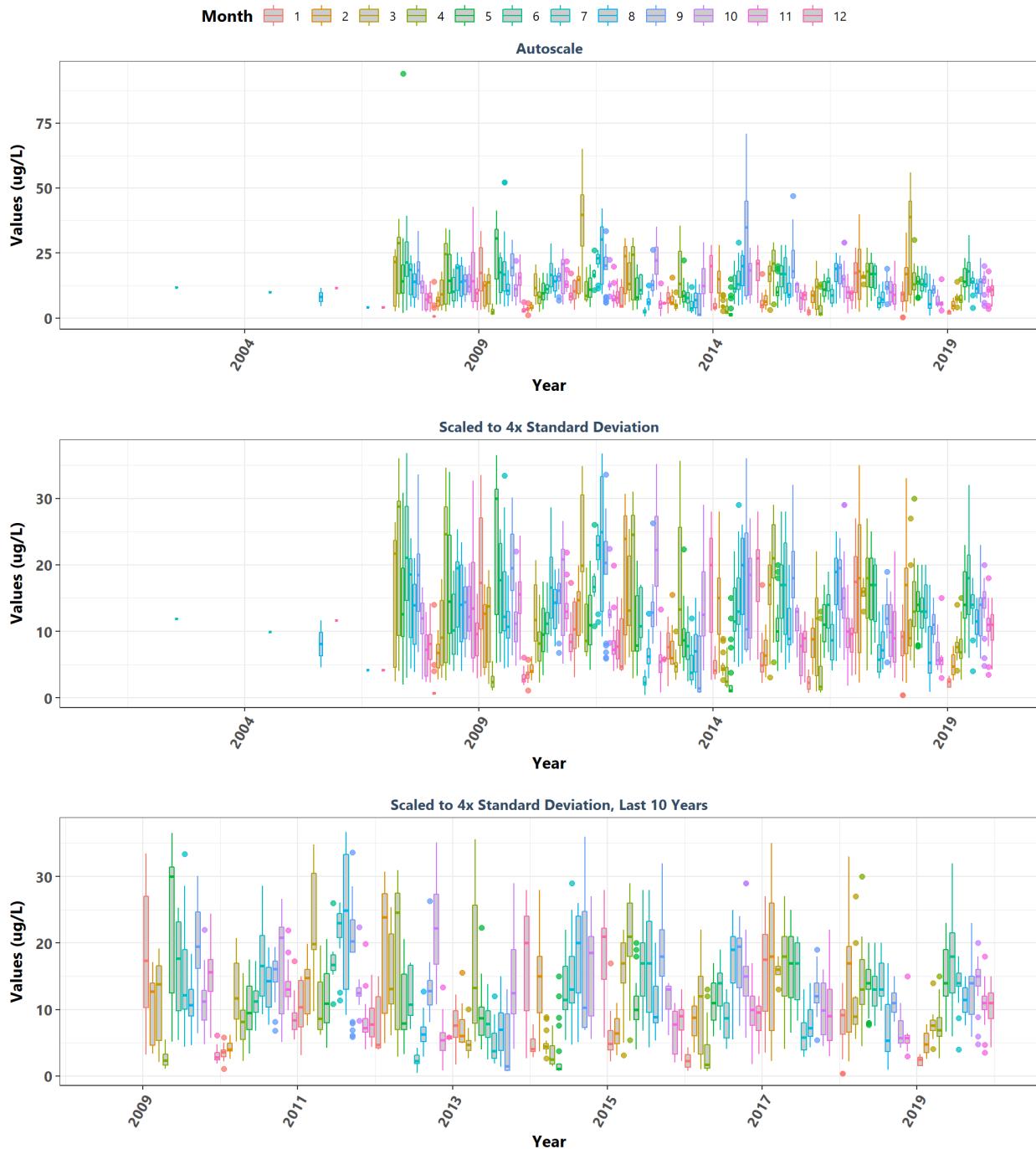
**Alligator Harbor Aquatic Preserve**  
By Month



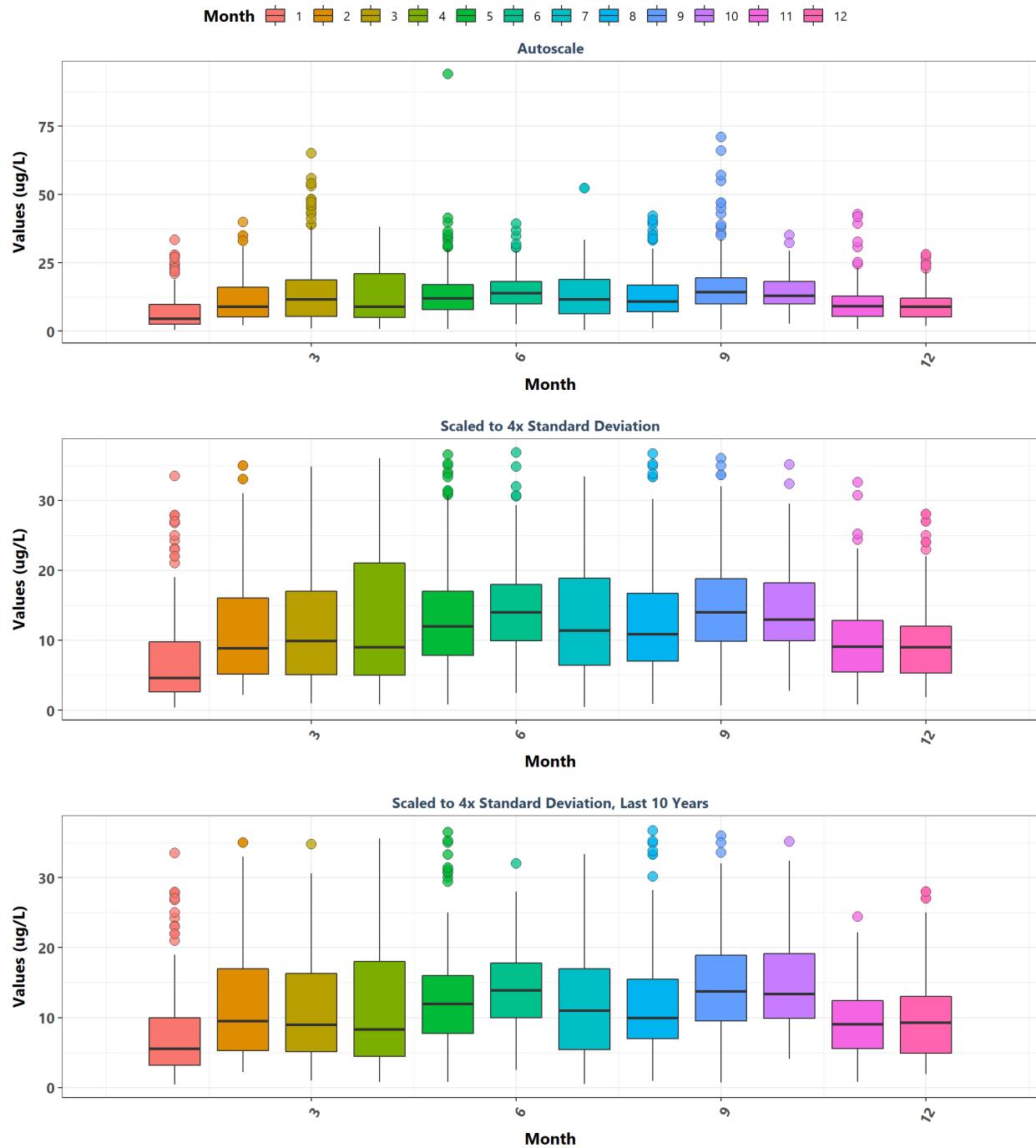
**Apalachicola Bay Aquatic Preserve**  
By Year



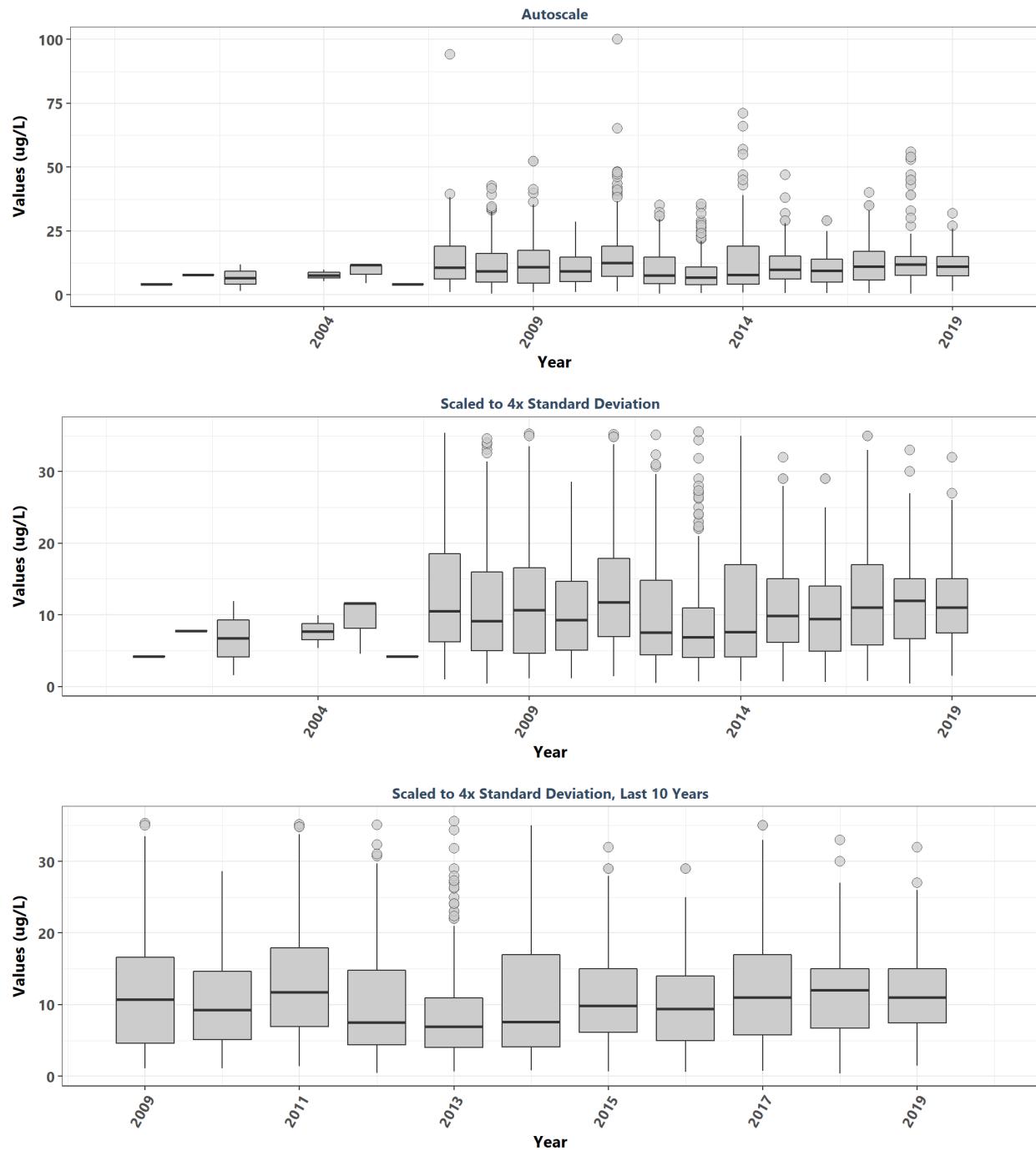
**Apalachicola Bay Aquatic Preserve**  
By Year & Month



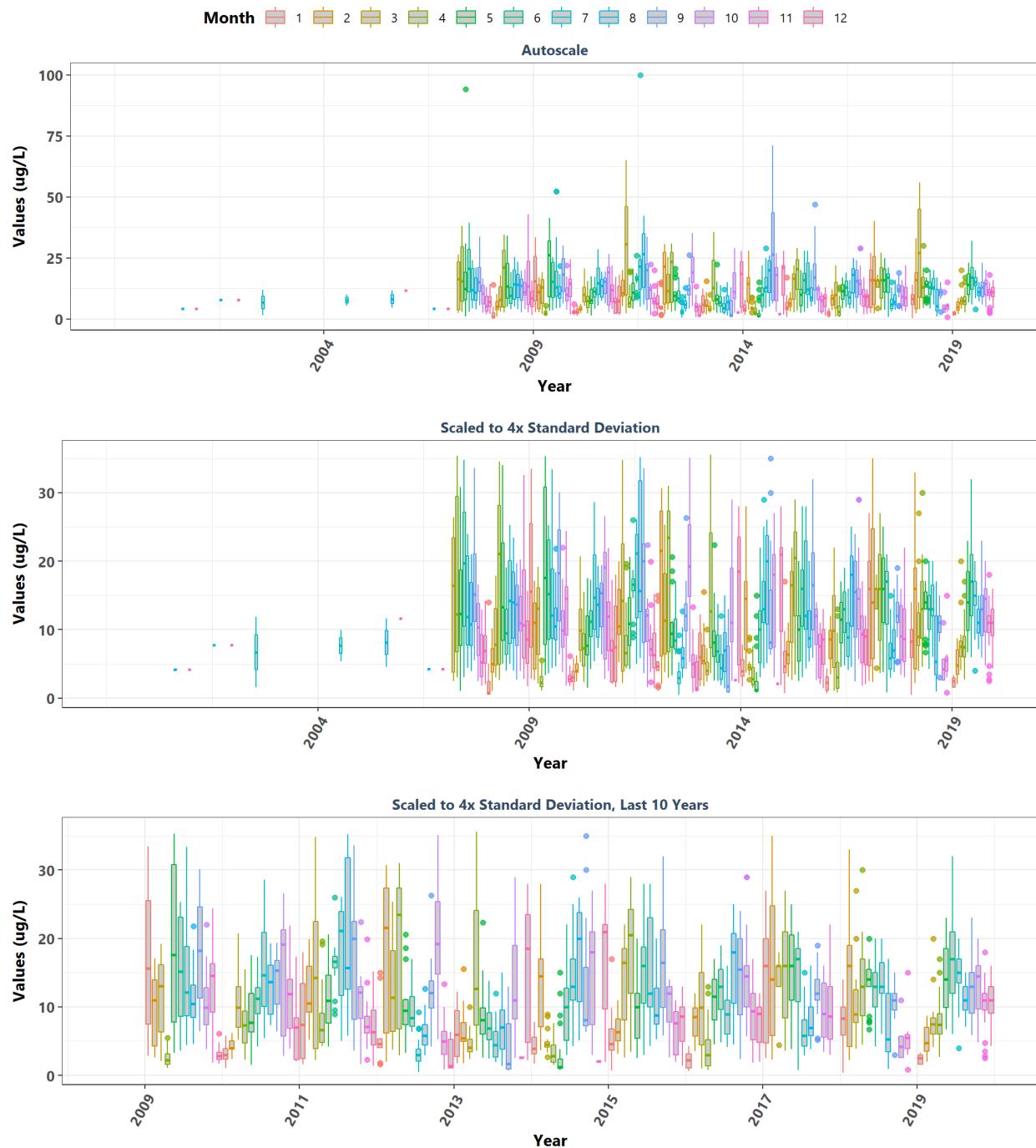
**Apalachicola Bay Aquatic Preserve**  
By Month



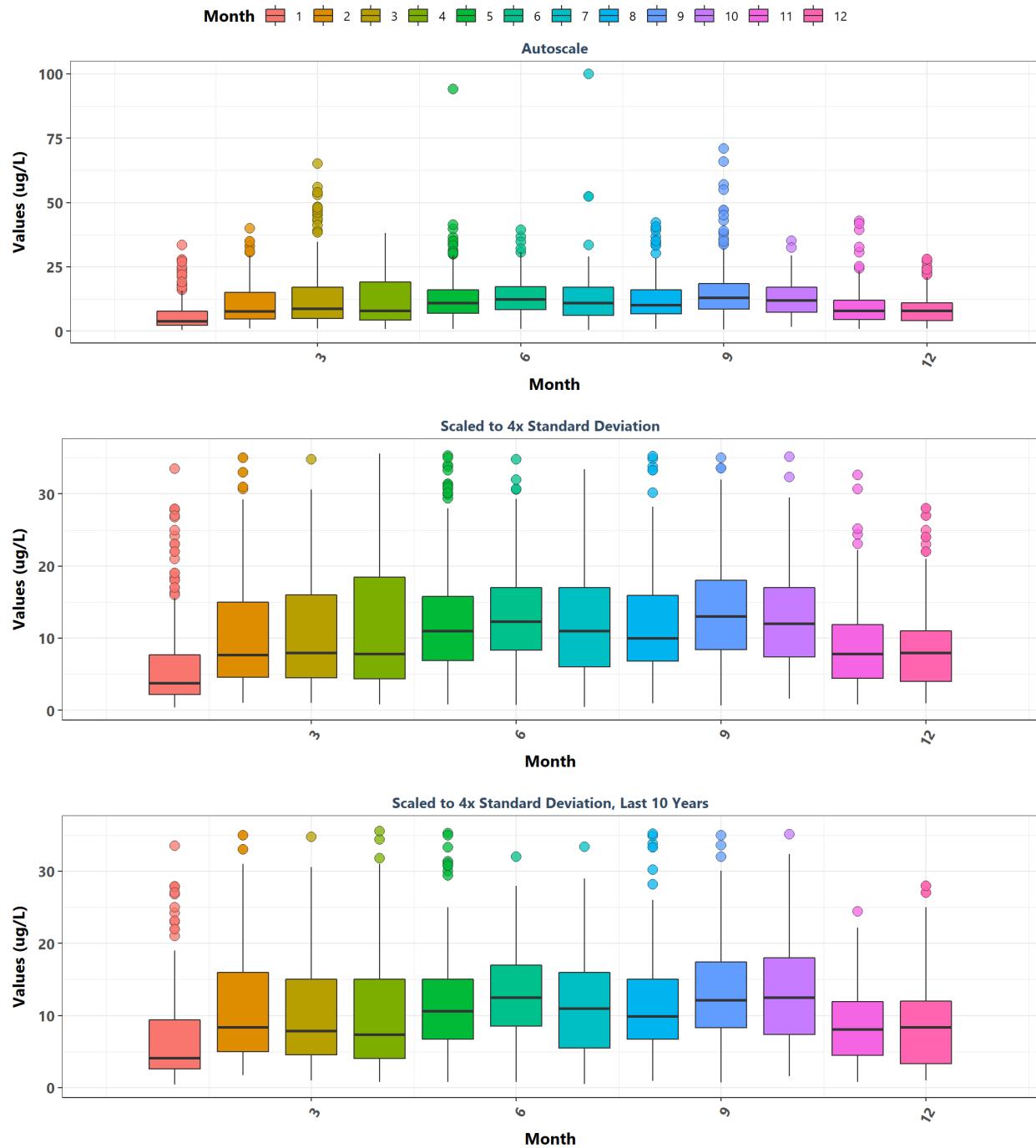
**Apalachicola National Estuarine Research Reserve**  
By Year



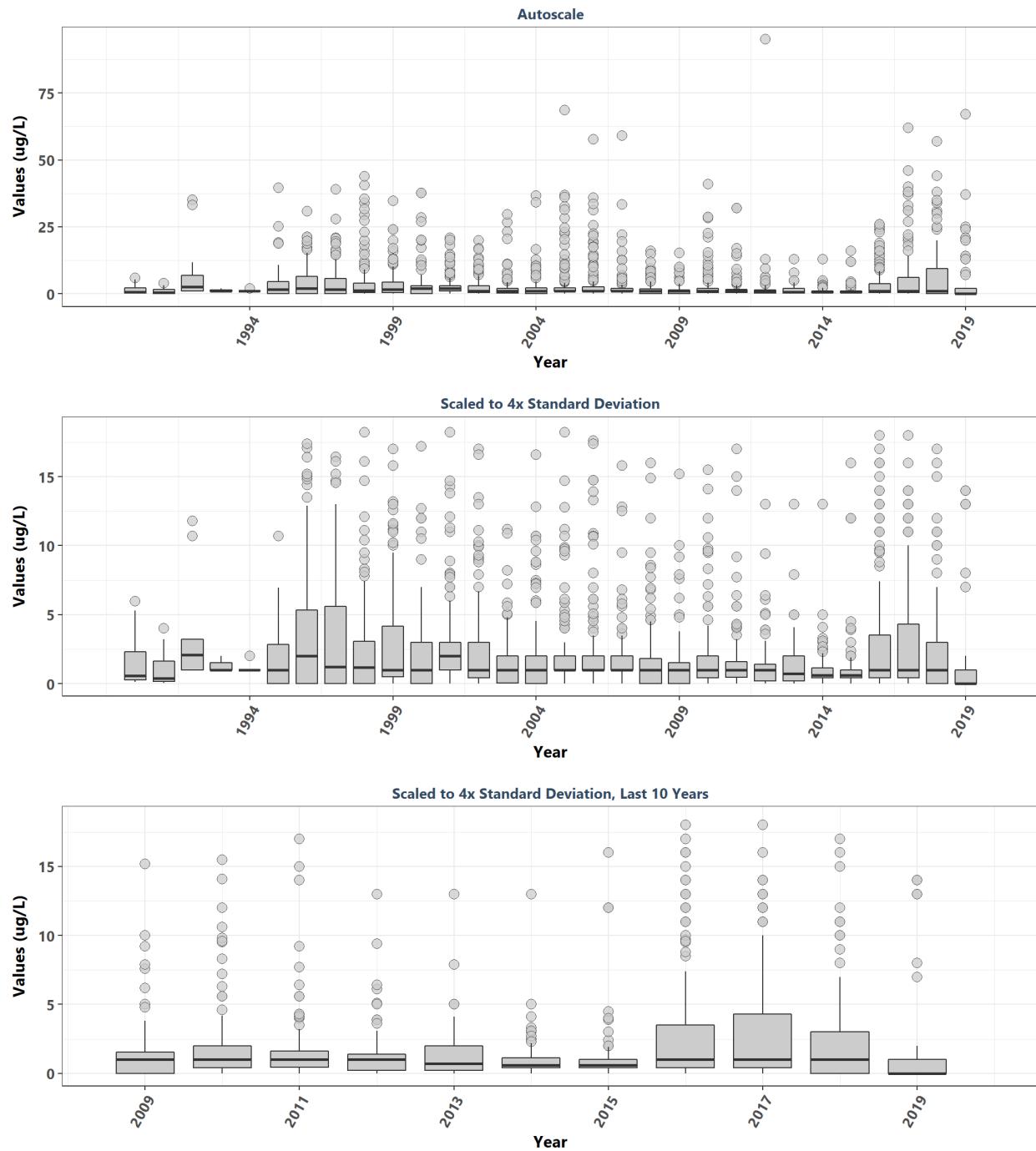
**Apalachicola National Estuarine Research Reserve**  
By Year & Month



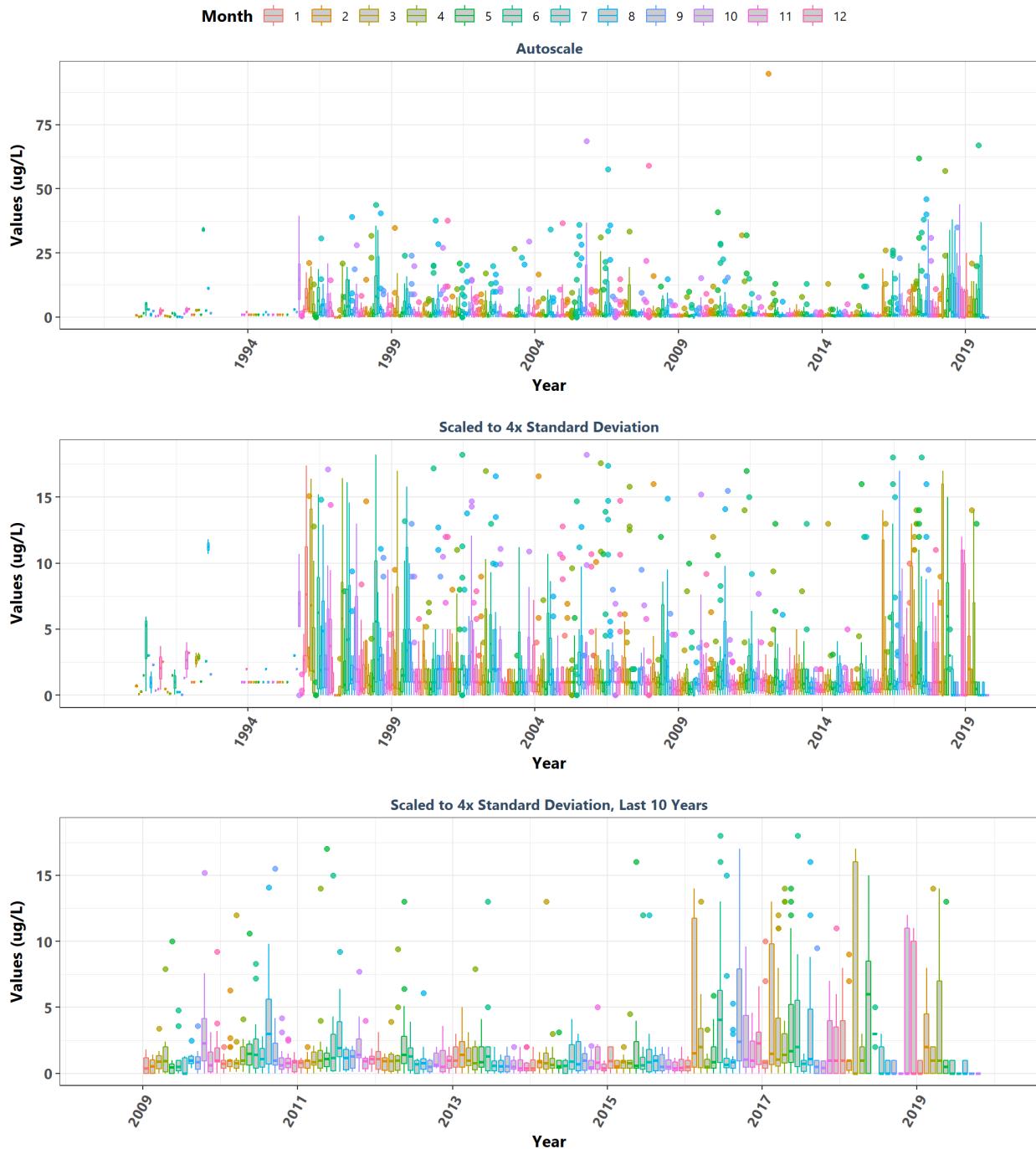
**Apalachicola National Estuarine Research Reserve**  
By Month



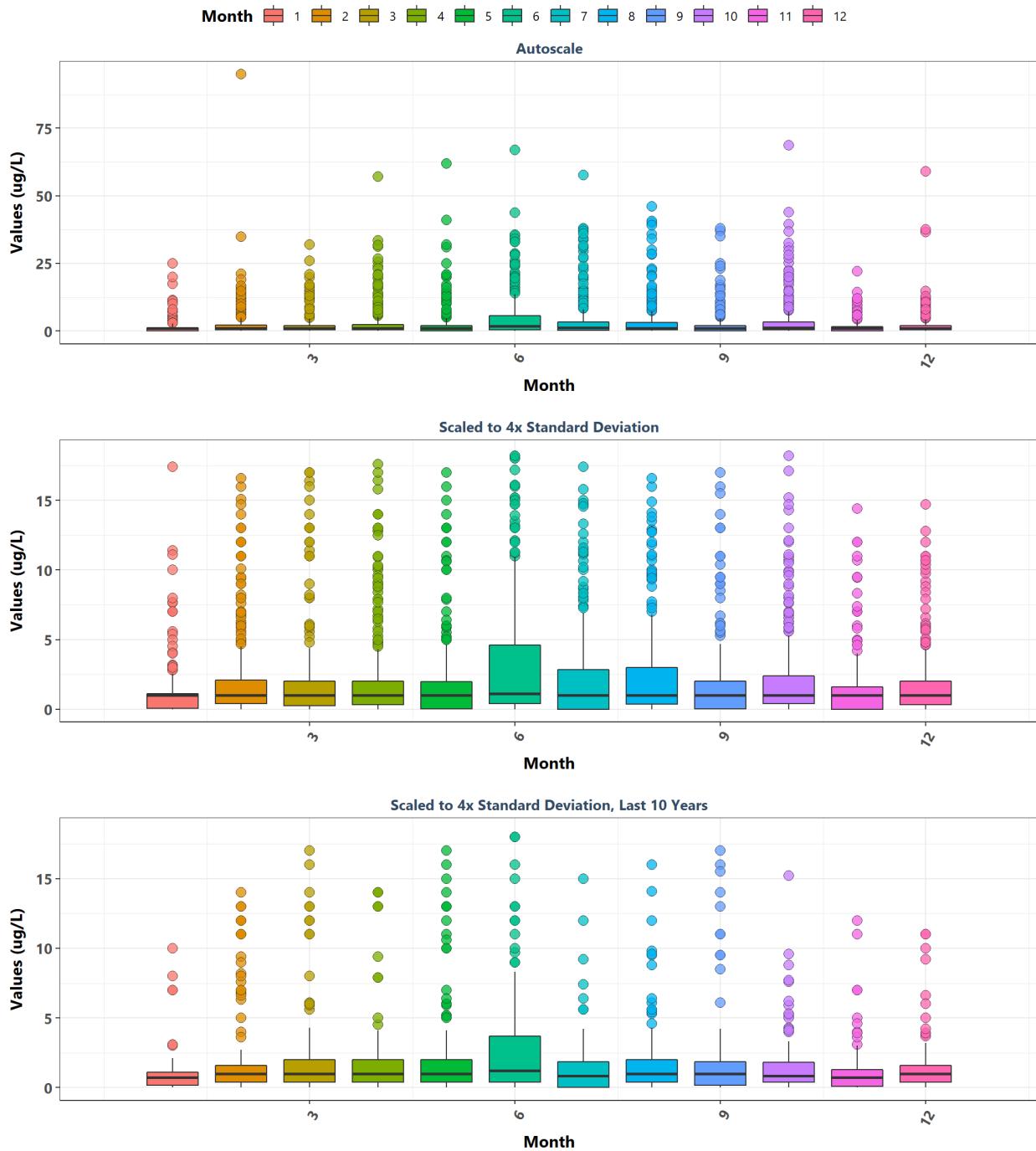
**Big Bend Seagrasses Aquatic Preserve**  
By Year



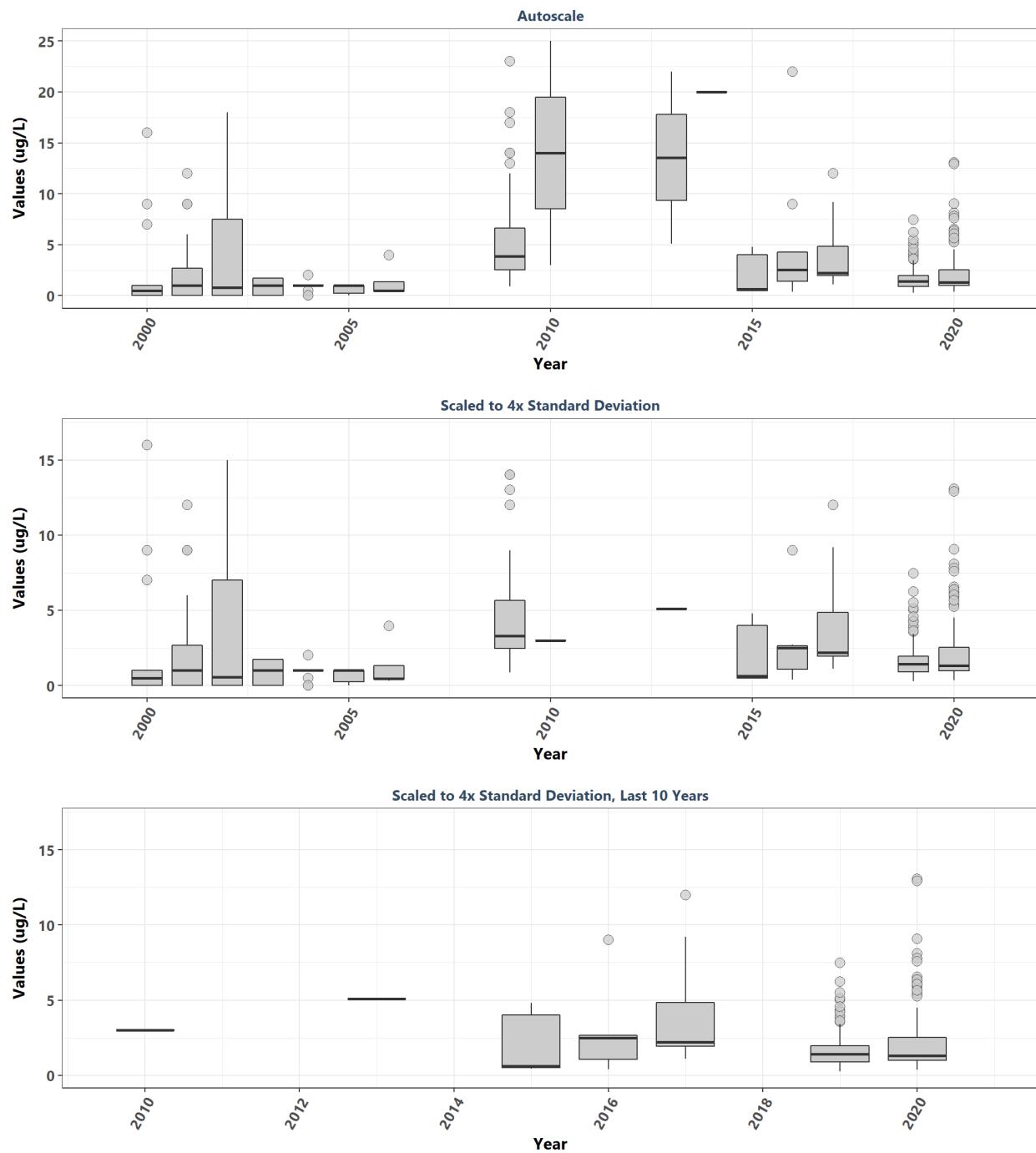
**Big Bend Seagrasses Aquatic Preserve**  
By Year & Month



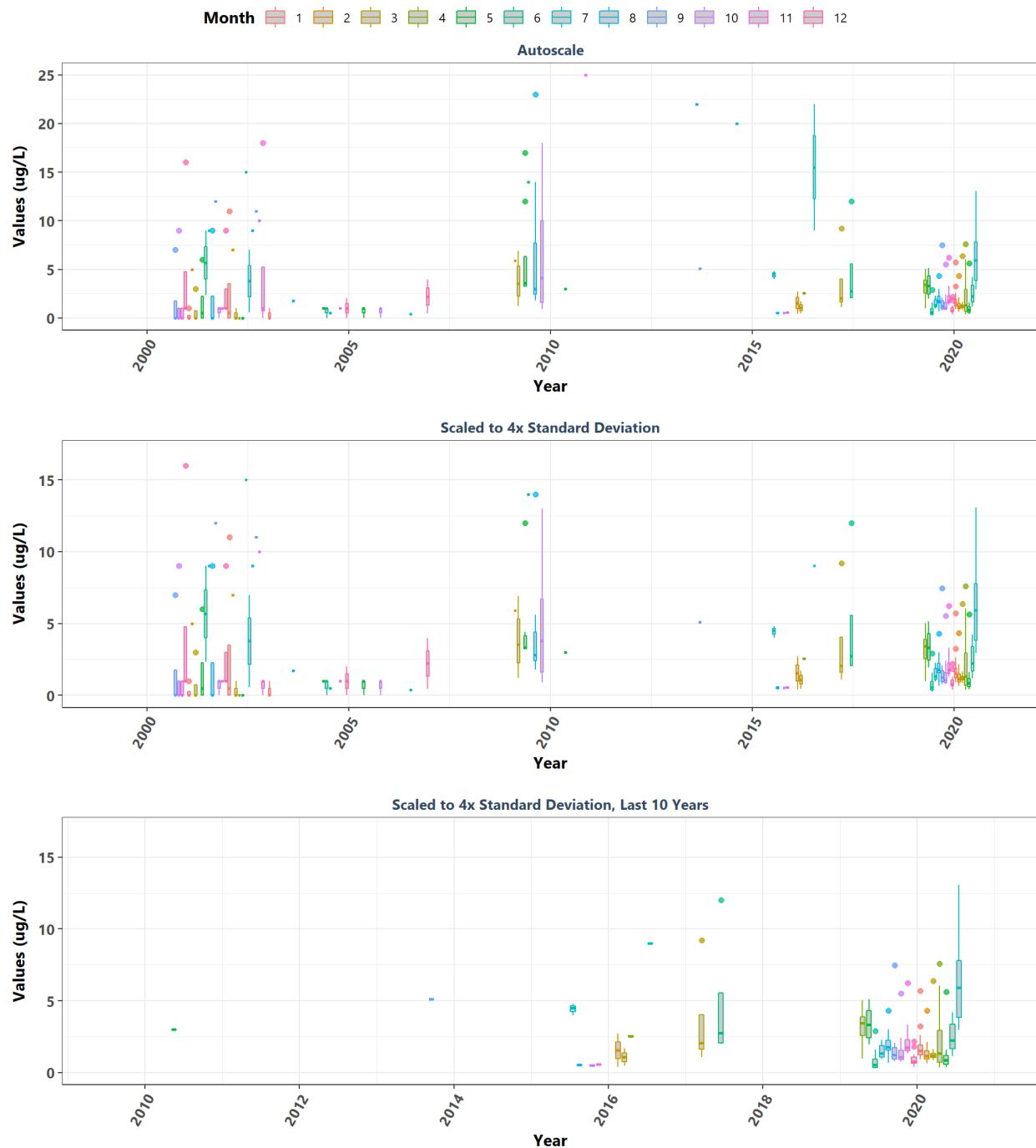
**Big Bend Seagrasses Aquatic Preserve**  
By Month



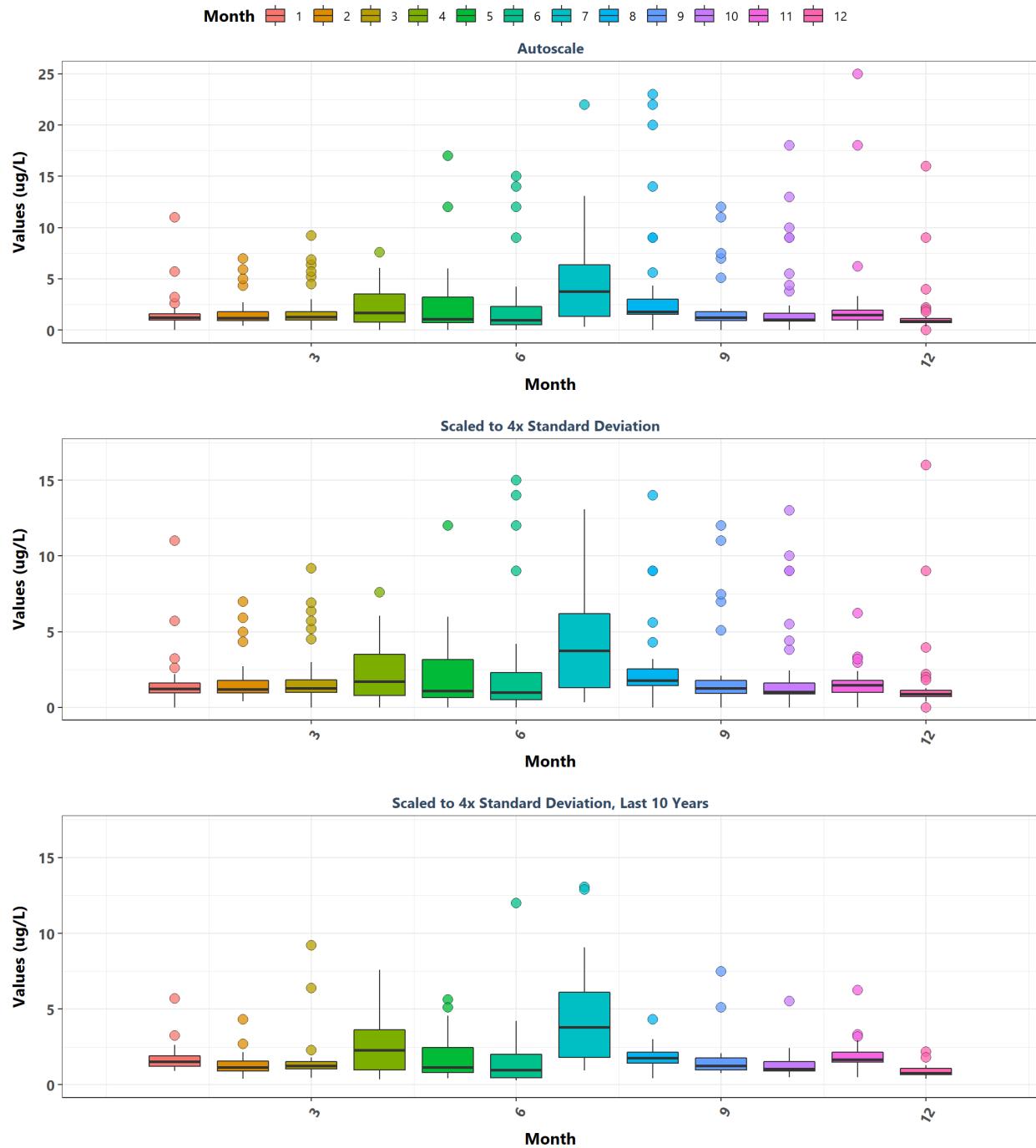
**Biscayne Bay Aquatic Preserve**  
By Year



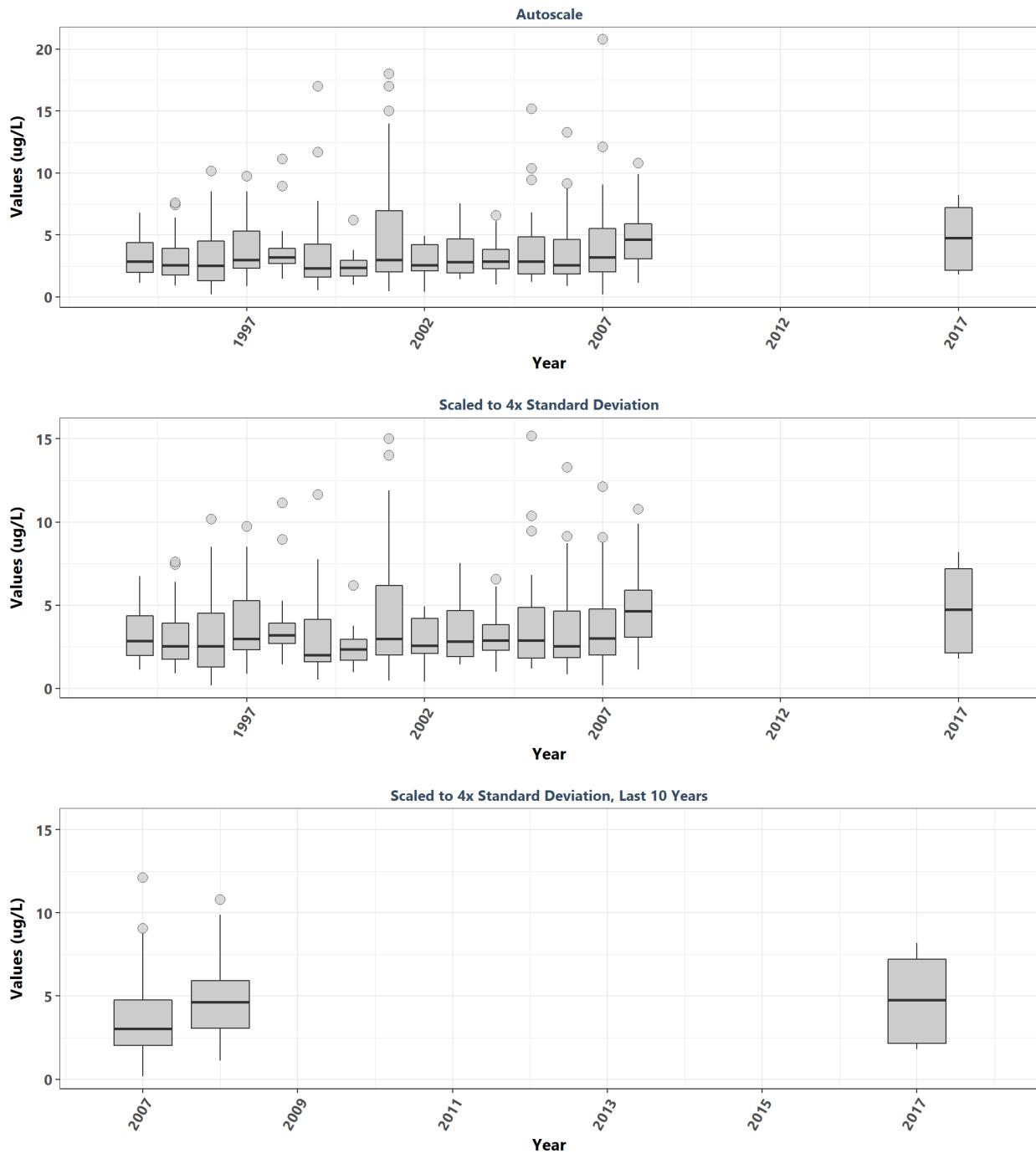
**Biscayne Bay Aquatic Preserve**  
By Year & Month



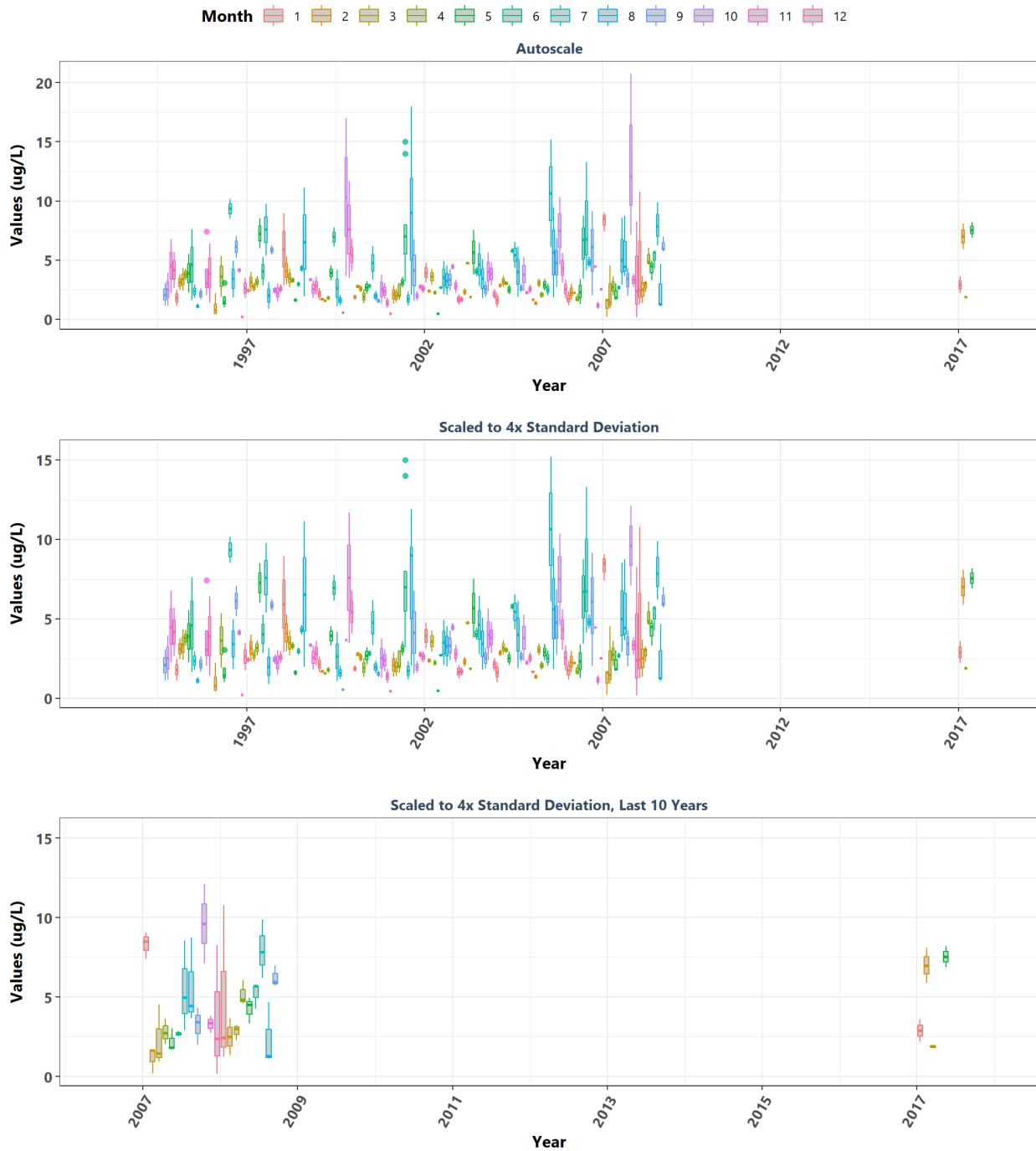
**Biscayne Bay Aquatic Preserve**  
By Month



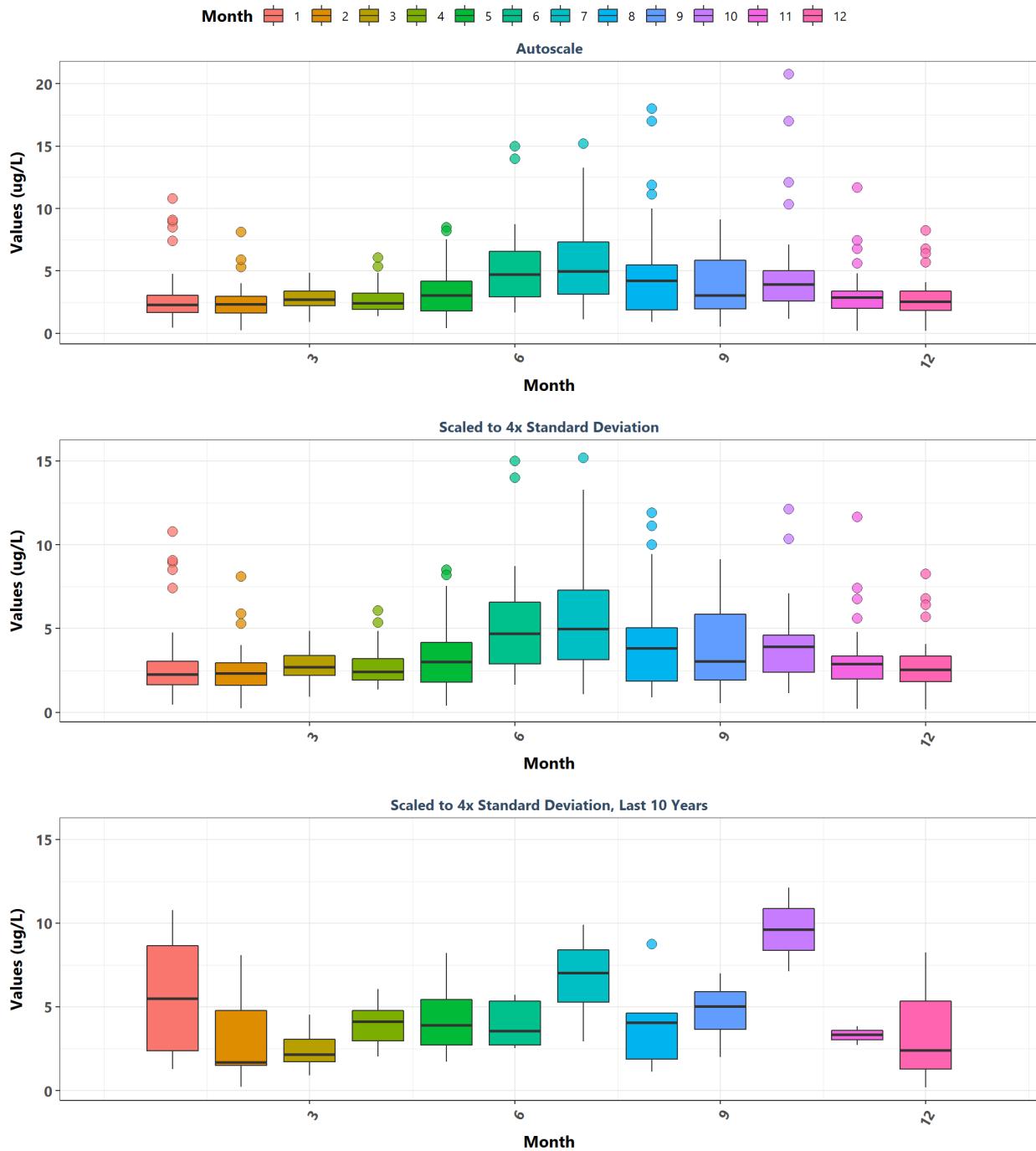
**Cape Romano-Ten Thousand Islands Aquatic Preserve**  
By Year



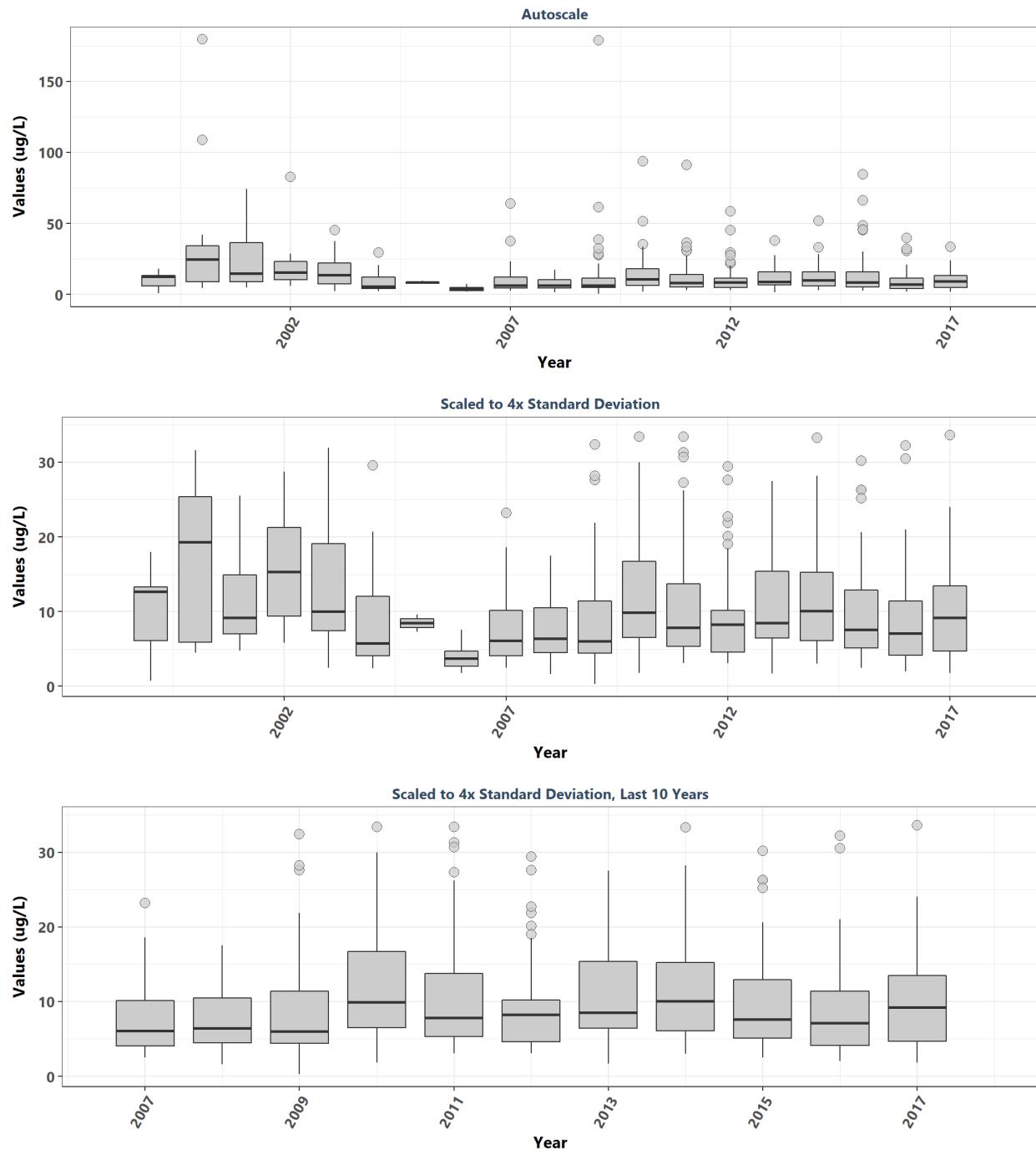
**Cape Romano-Ten Thousand Islands Aquatic Preserve**  
By Year & Month



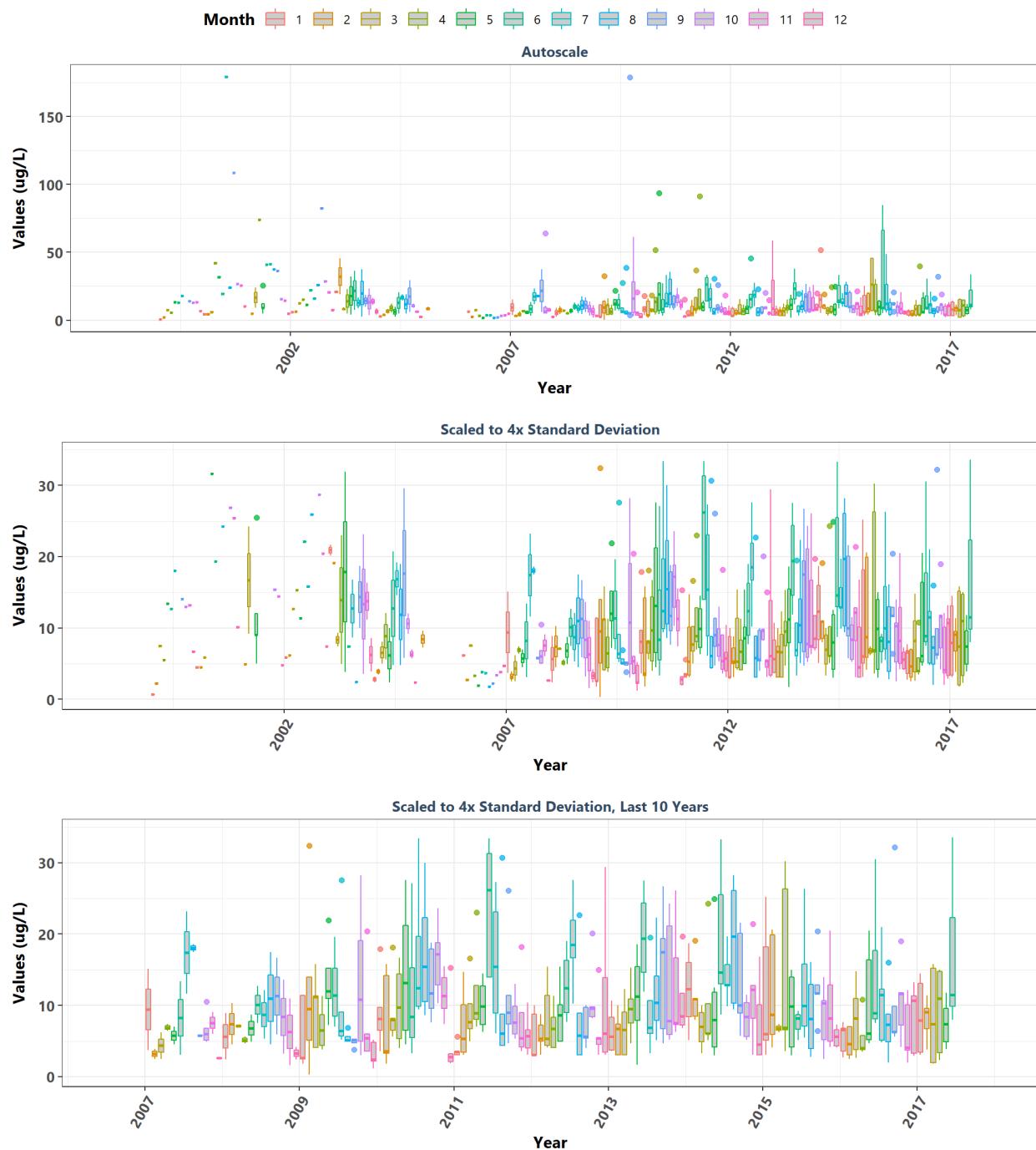
**Cape Romano-Ten Thousand Islands Aquatic Preserve**  
By Month



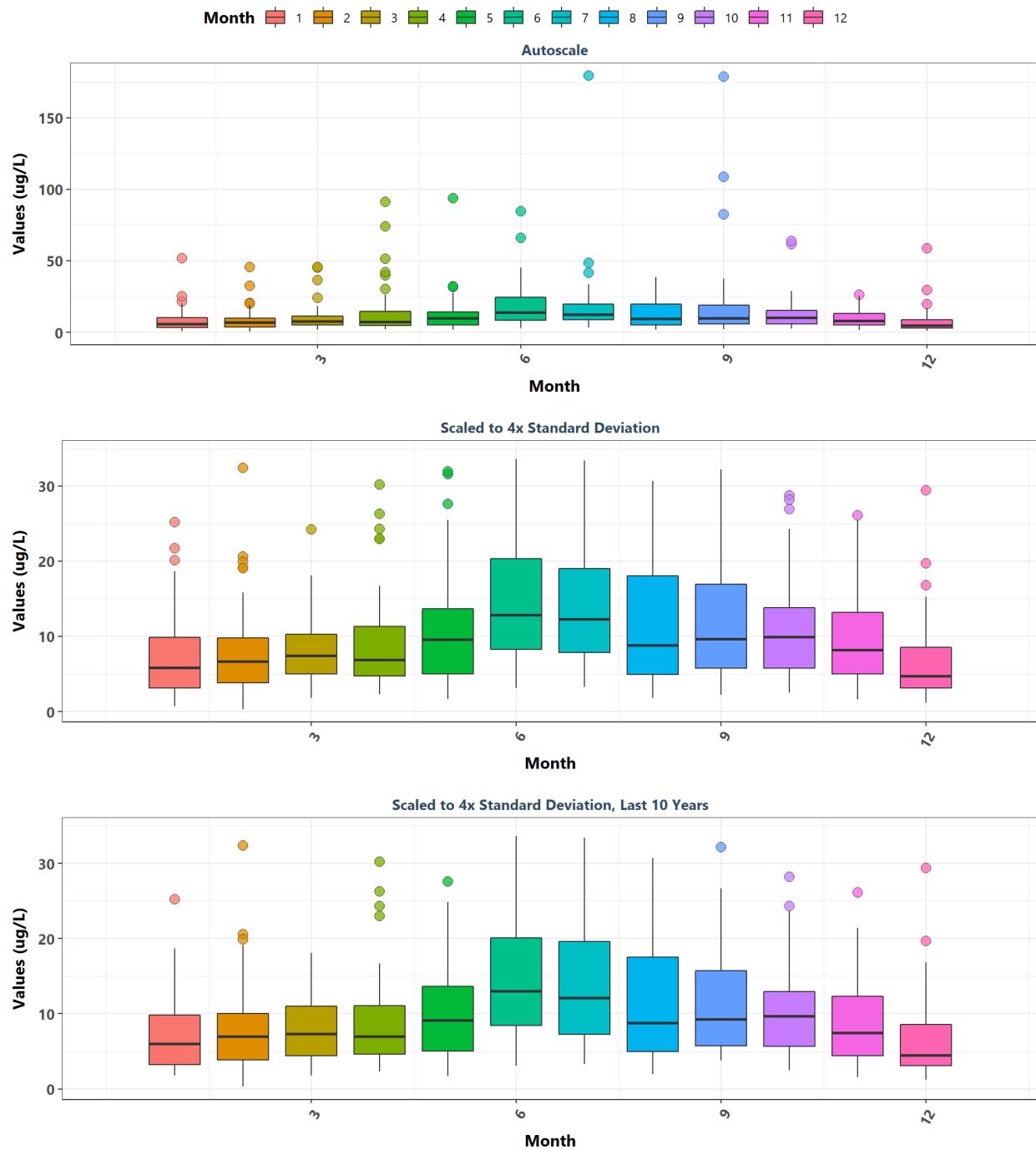
**Cockroach Bay Aquatic Preserve**  
By Year



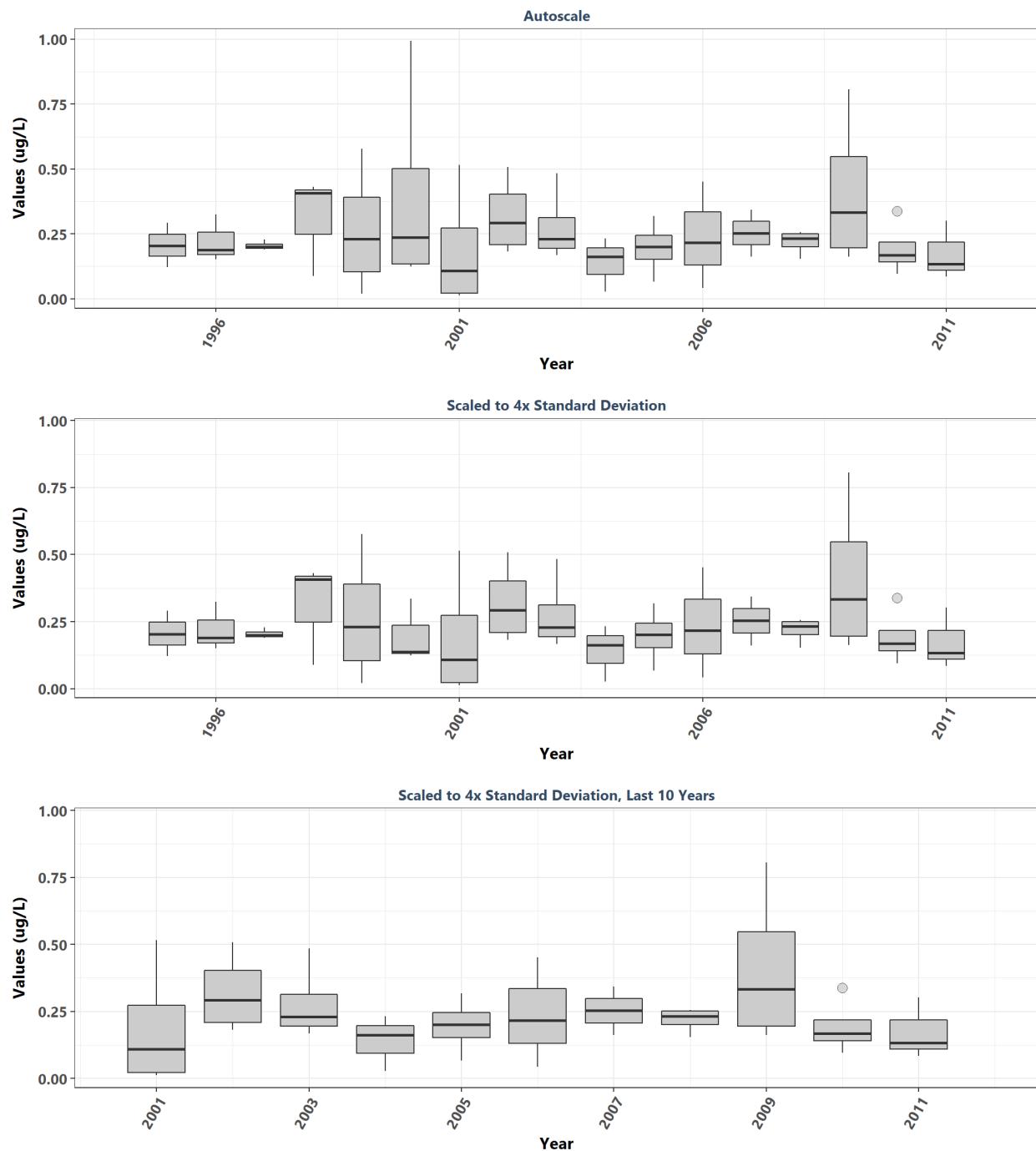
**Cockroach Bay Aquatic Preserve**  
By Year & Month



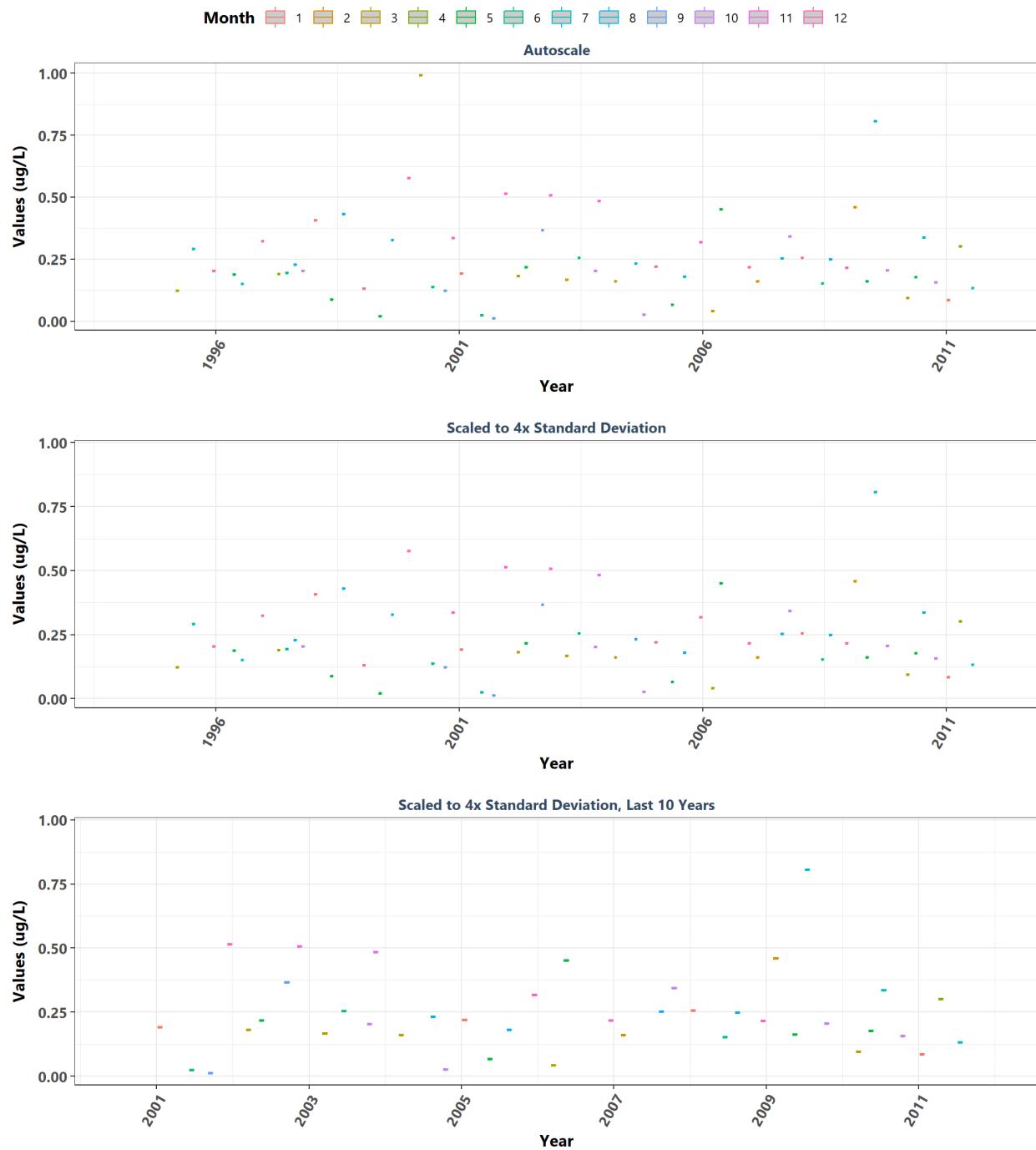
**Cockroach Bay Aquatic Preserve**  
By Month



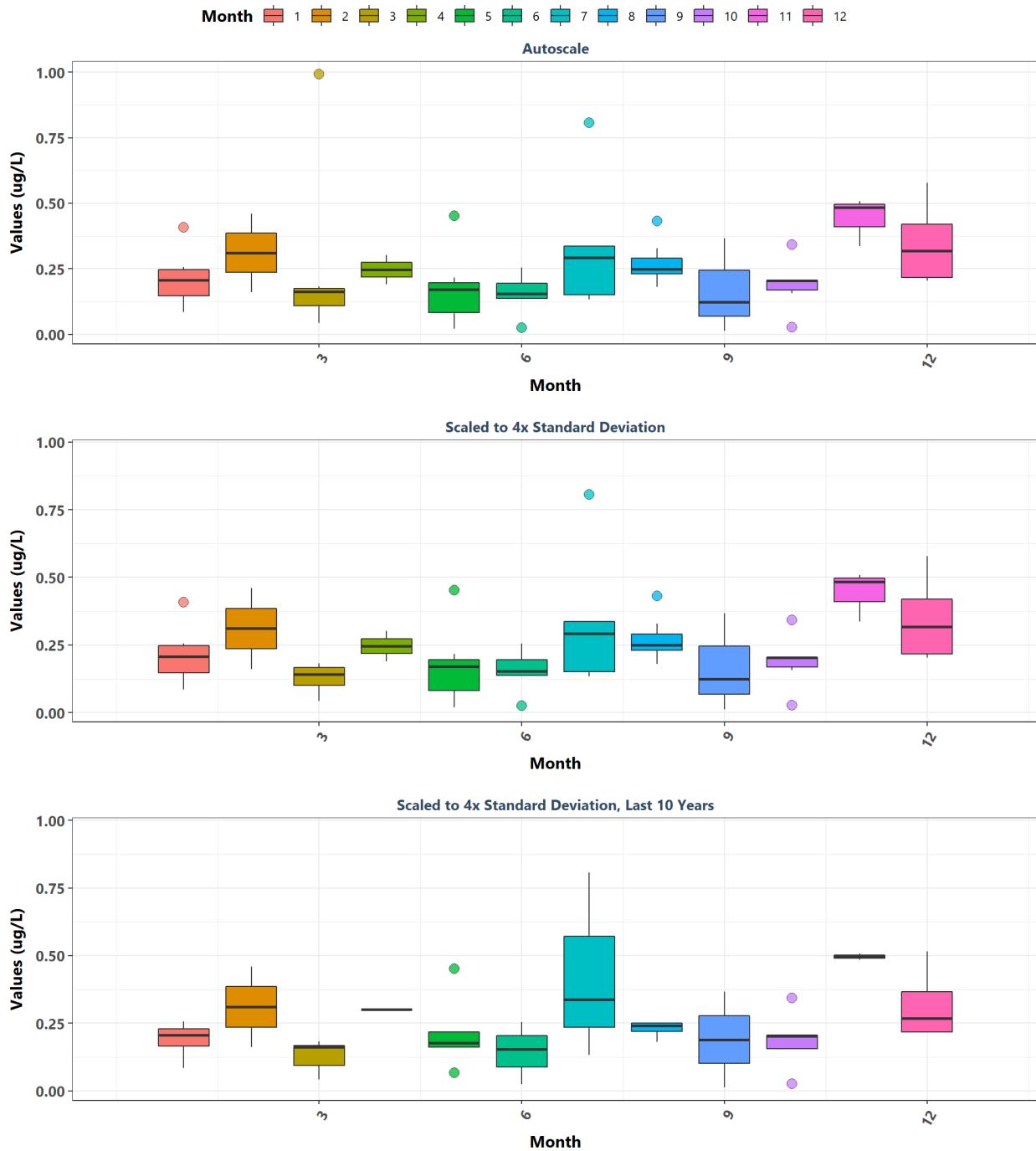
**Coupon Bight Aquatic Preserve**  
By Year



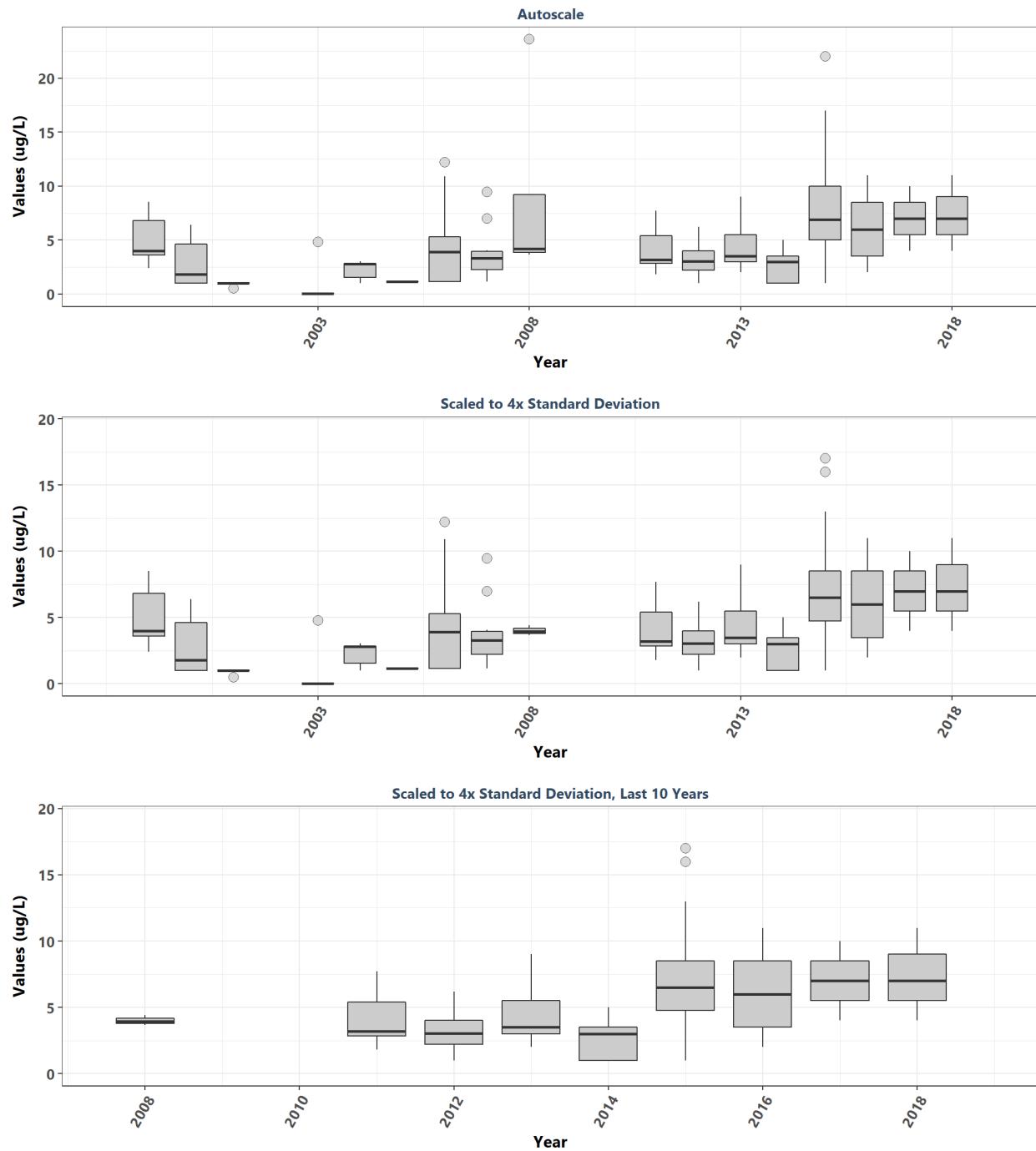
**Coupon Eight Aquatic Preserve**  
By Year & Month



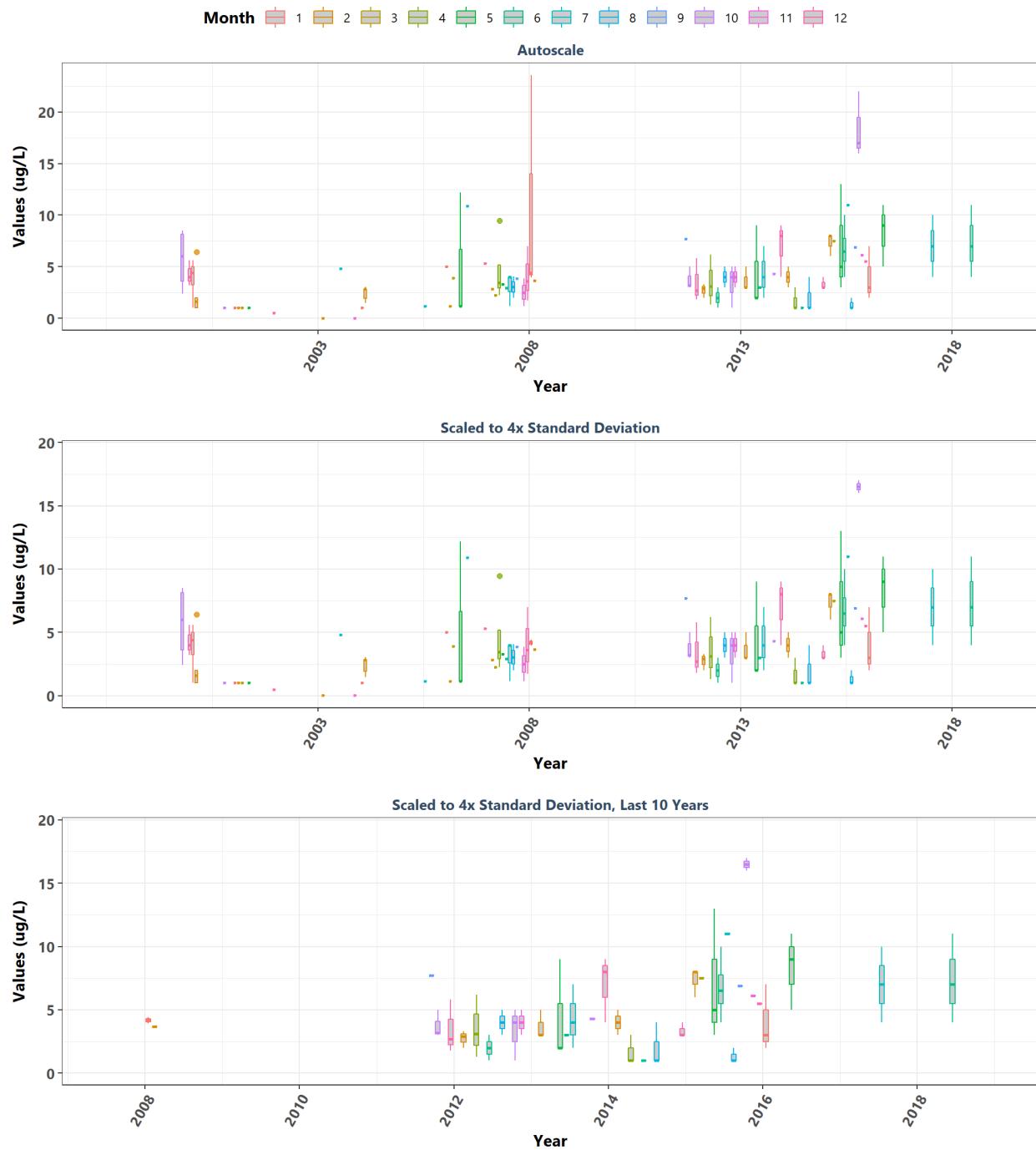
**Coupon Eight Aquatic Preserve  
By Month**



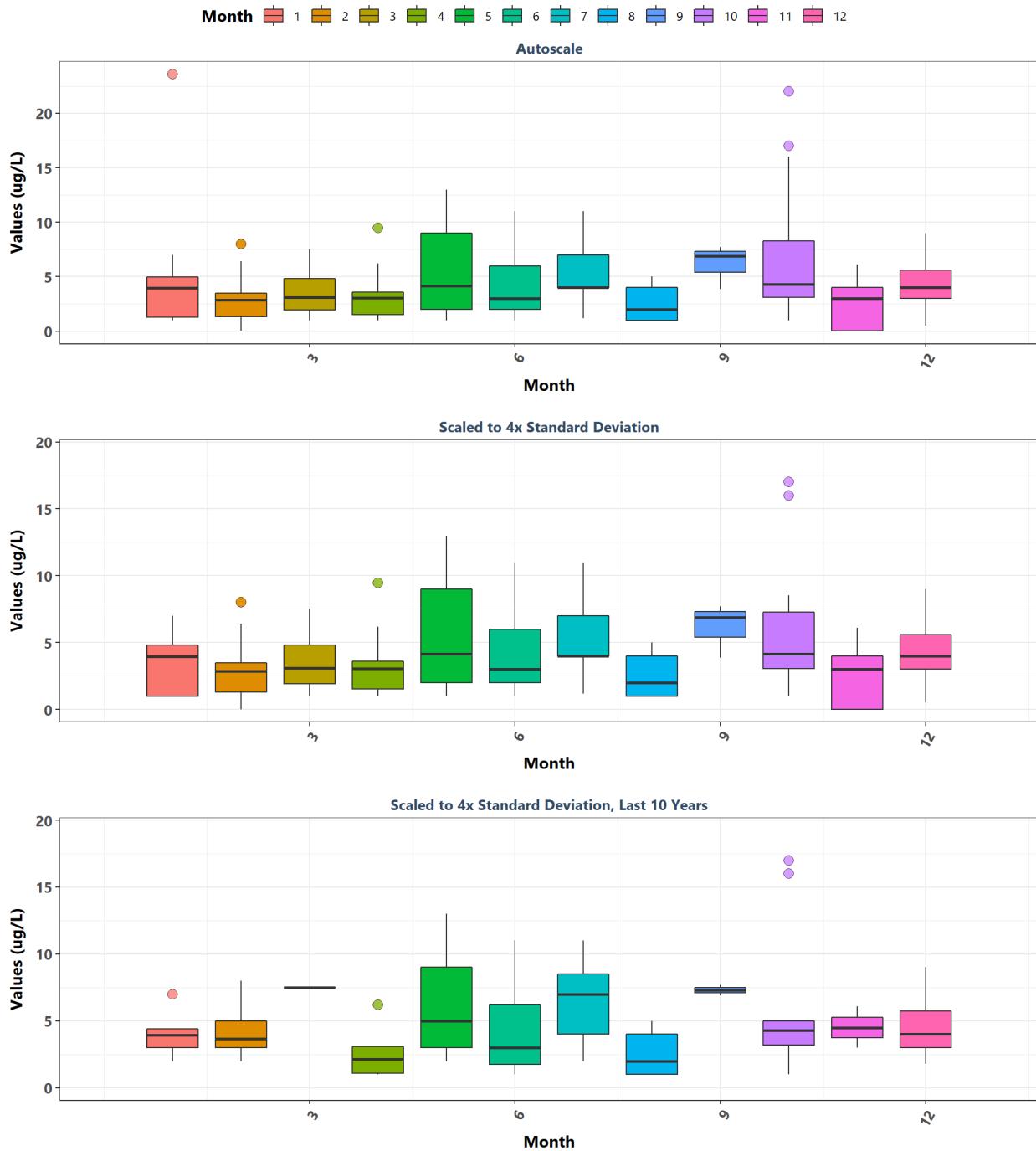
**Estero Bay Aquatic Preserve**  
By Year



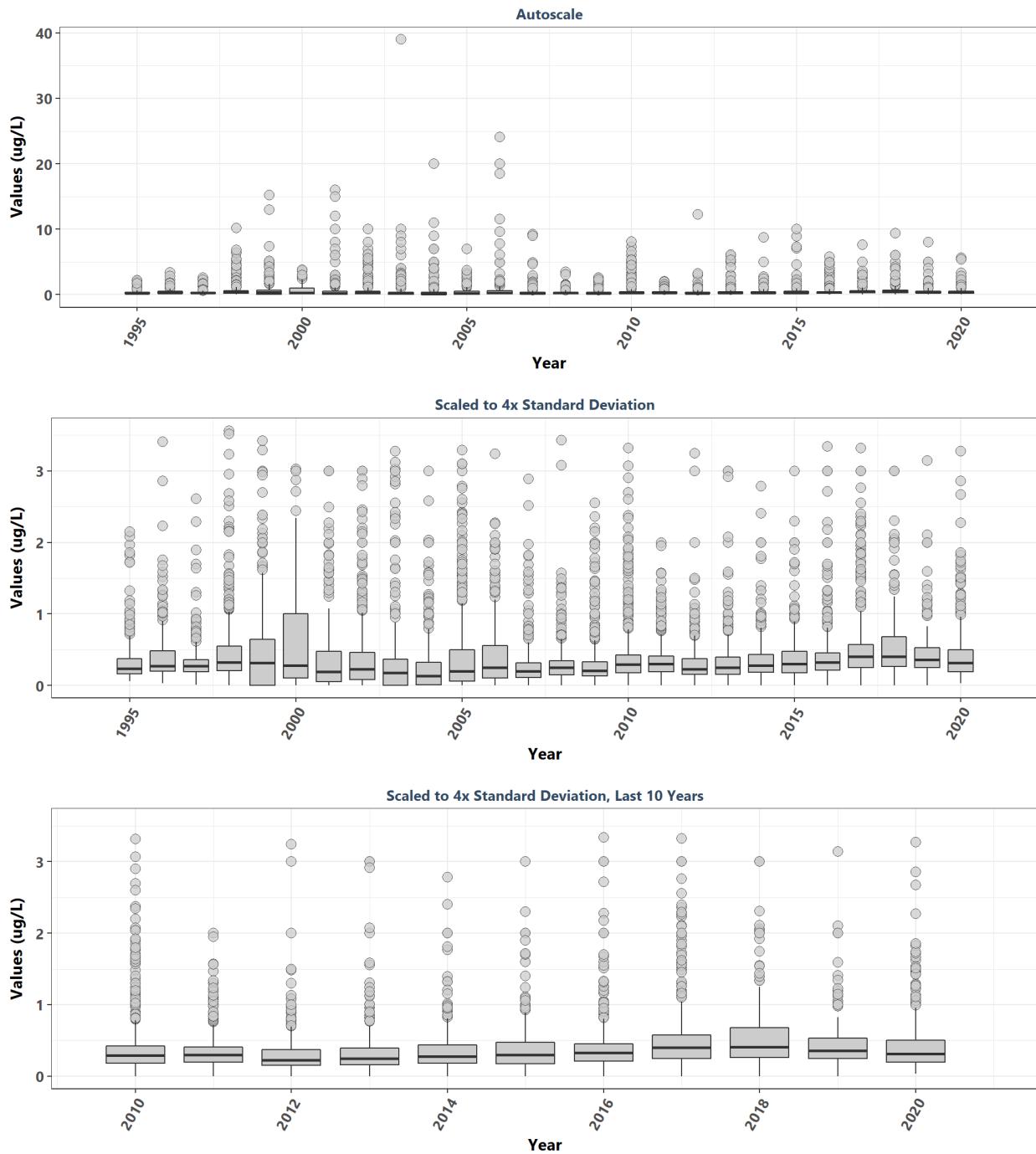
**Estero Bay Aquatic Preserve**  
By Year & Month



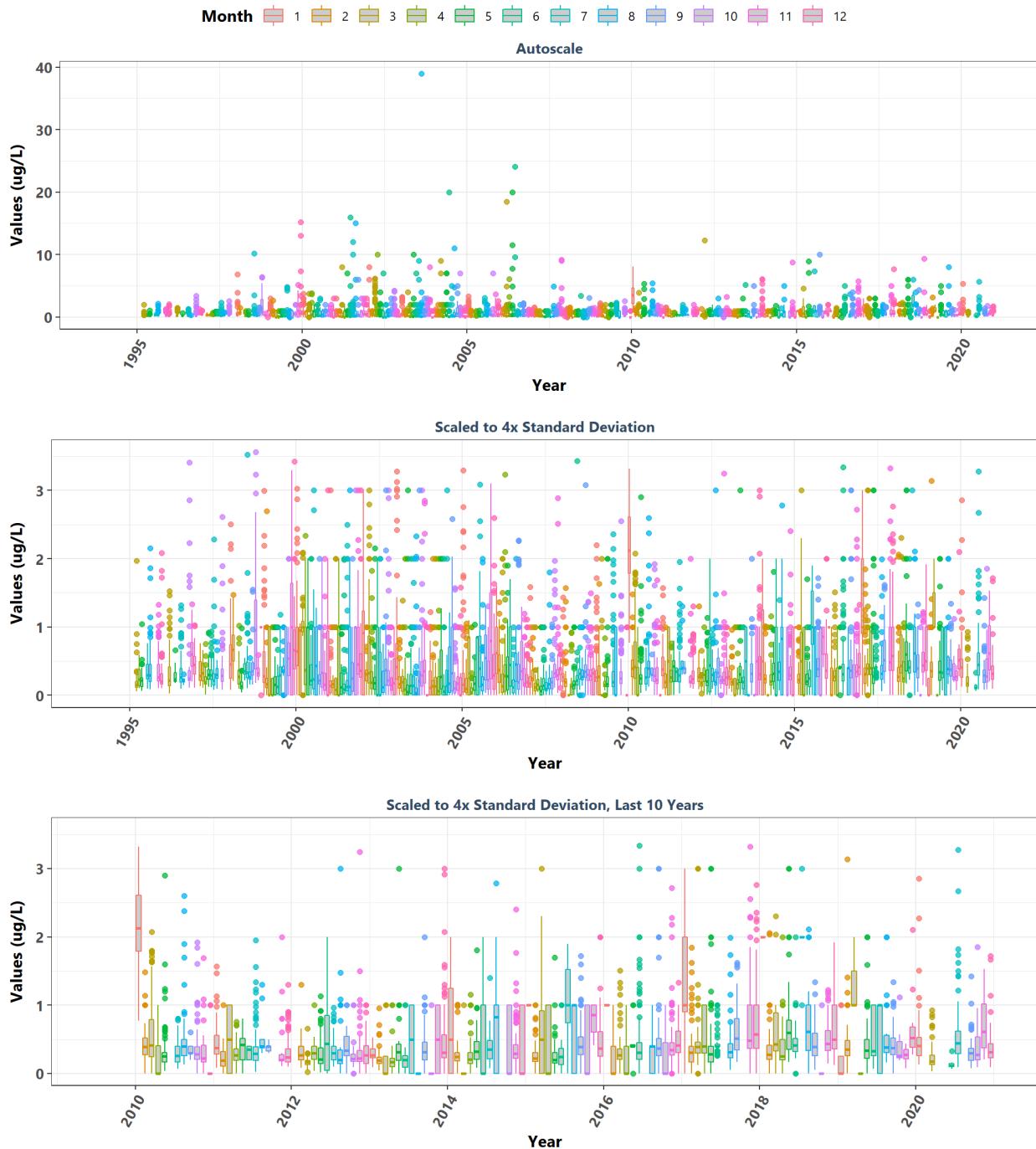
**Estero Bay Aquatic Preserve**  
By Month



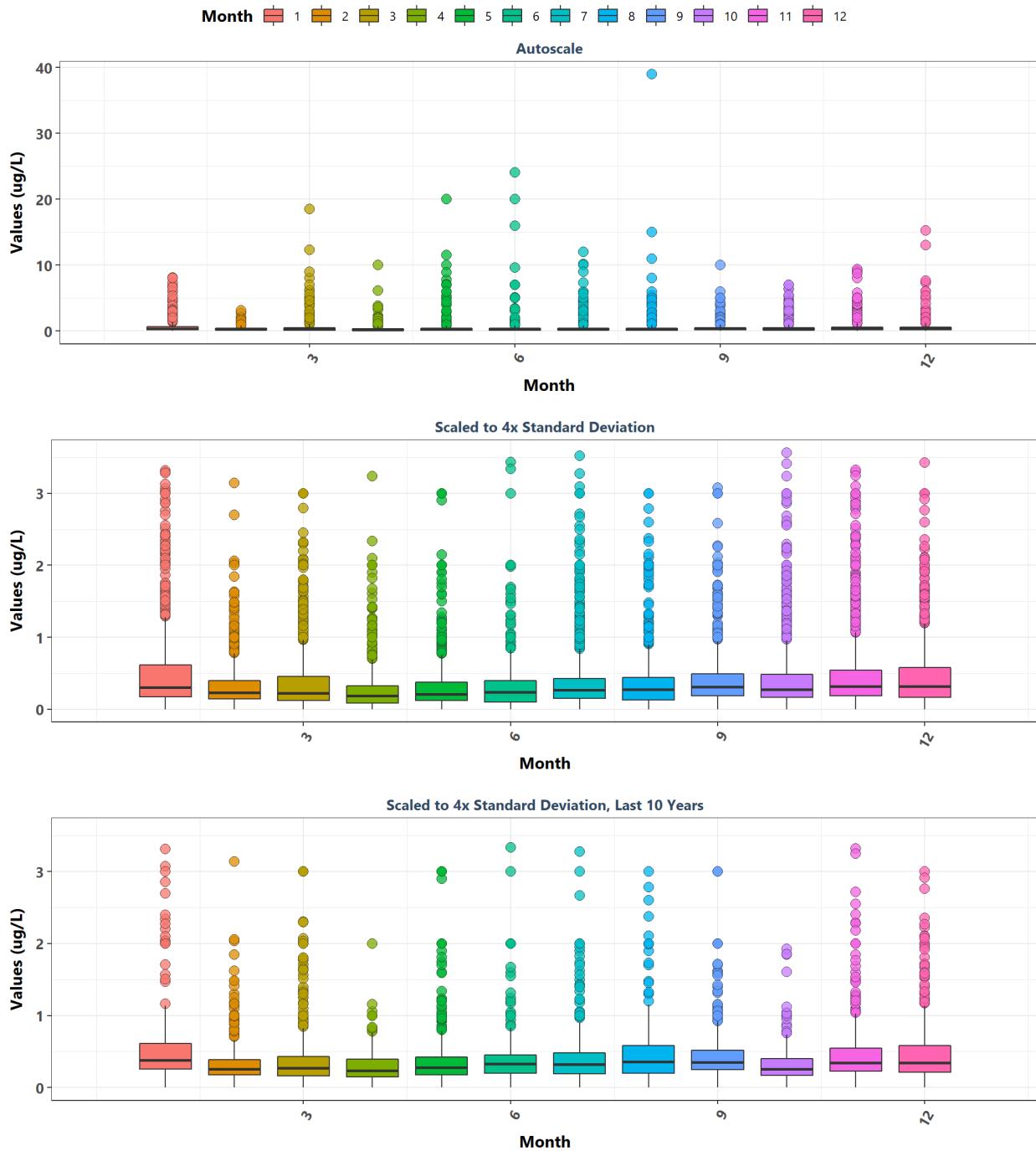
**Florida Keys National Marine Sanctuary**  
By Year



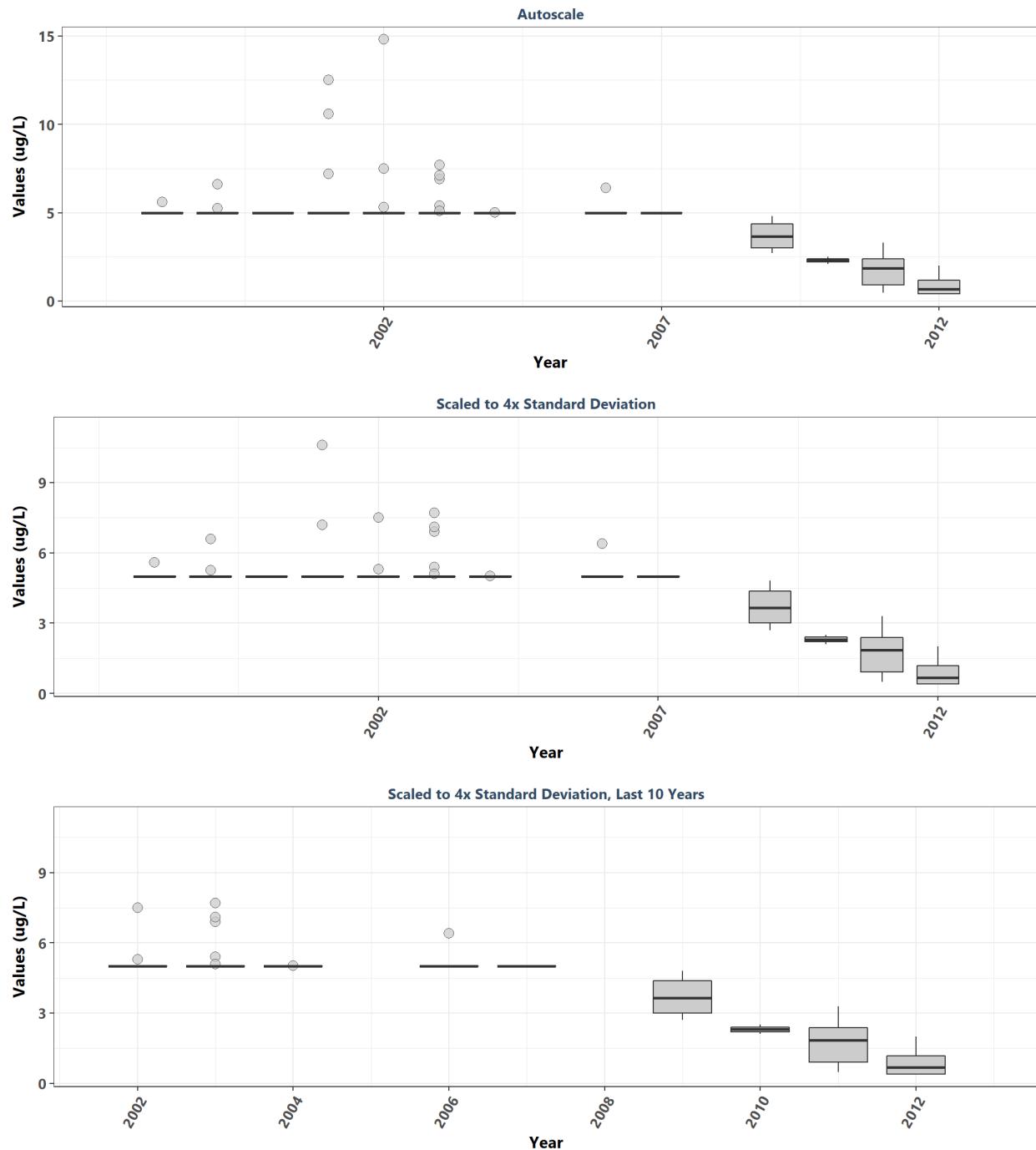
**Florida Keys National Marine Sanctuary**  
By Year & Month



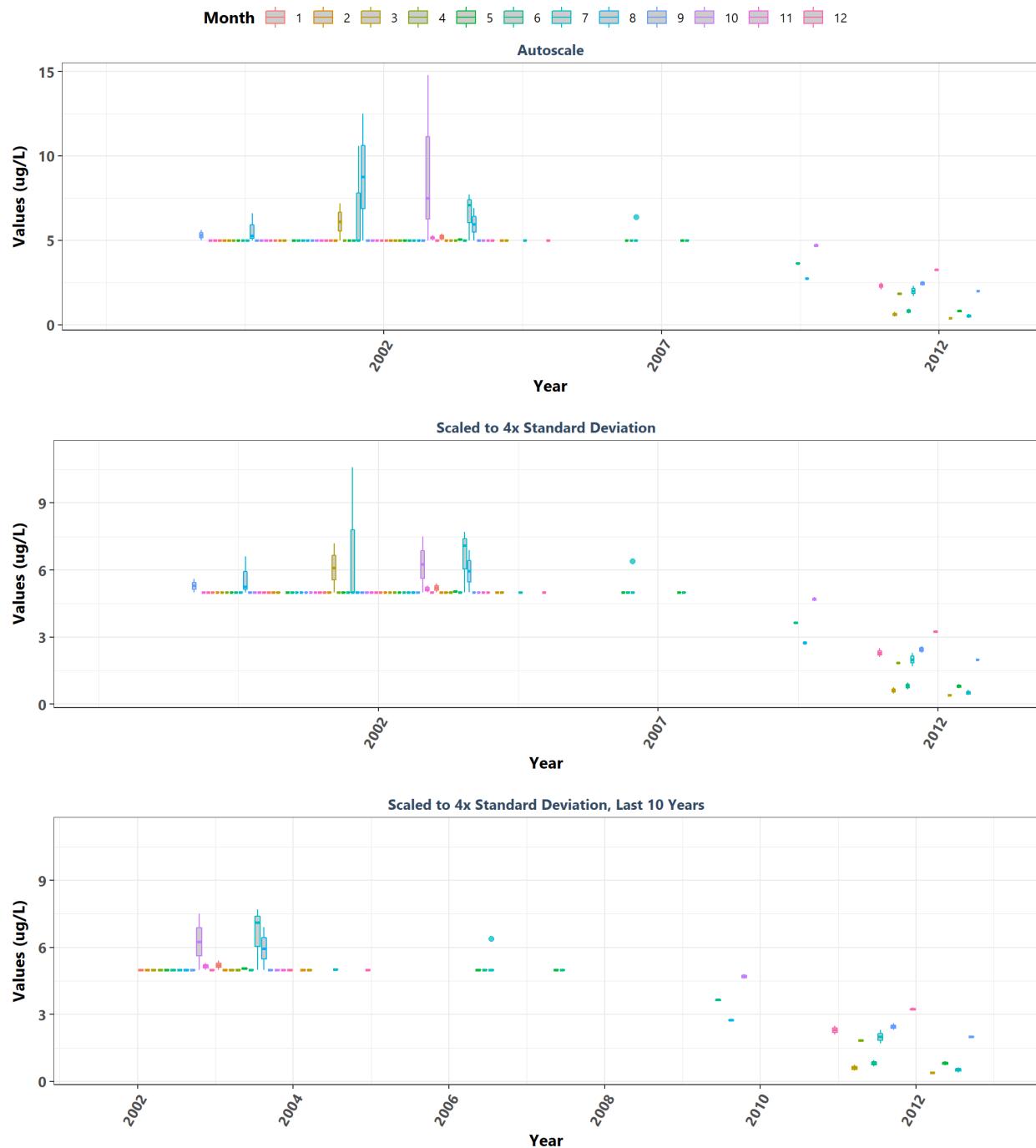
**Florida Keys National Marine Sanctuary**  
By Month



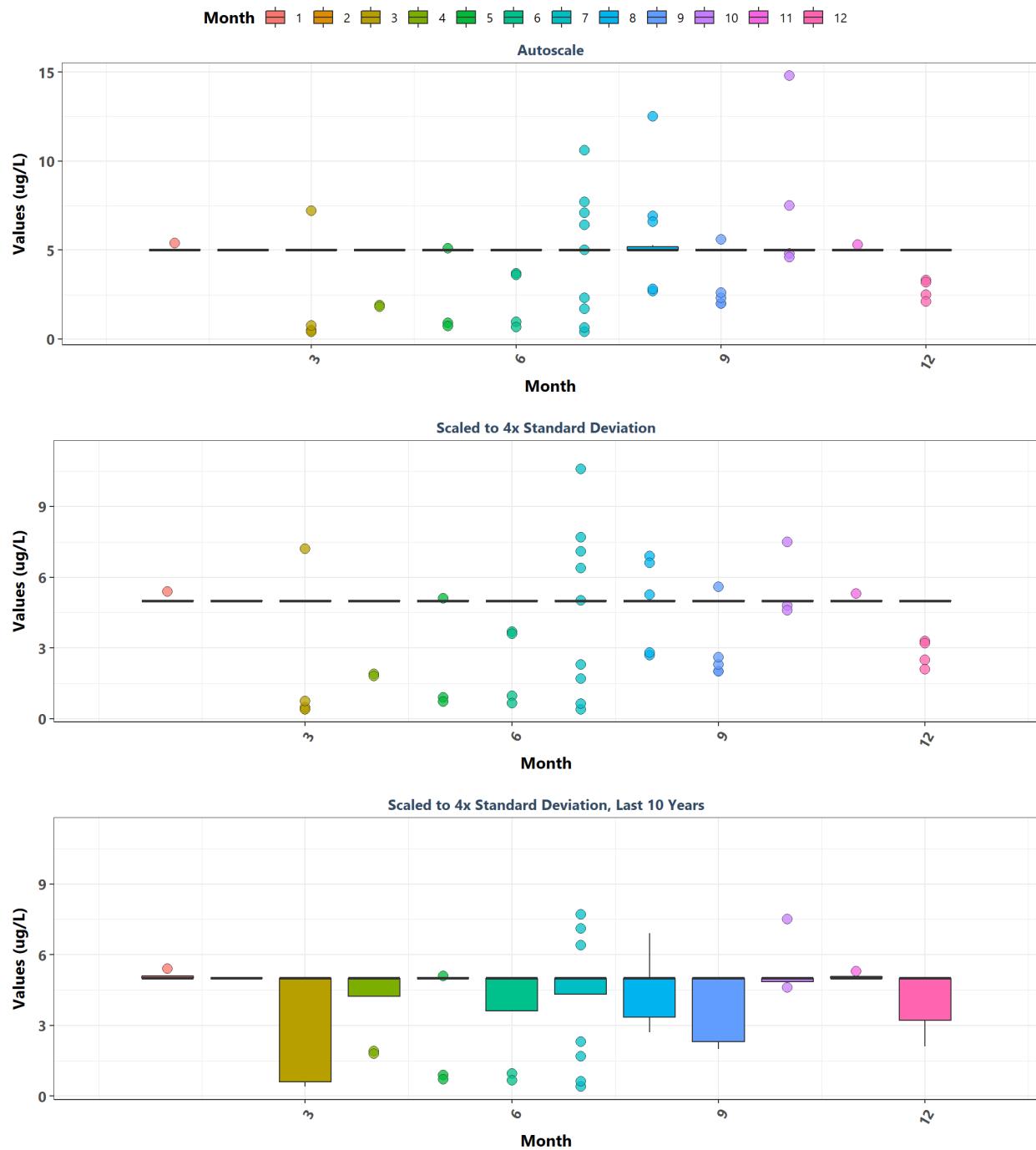
**Fort Pickens State Park Aquatic Preserve**  
By Year



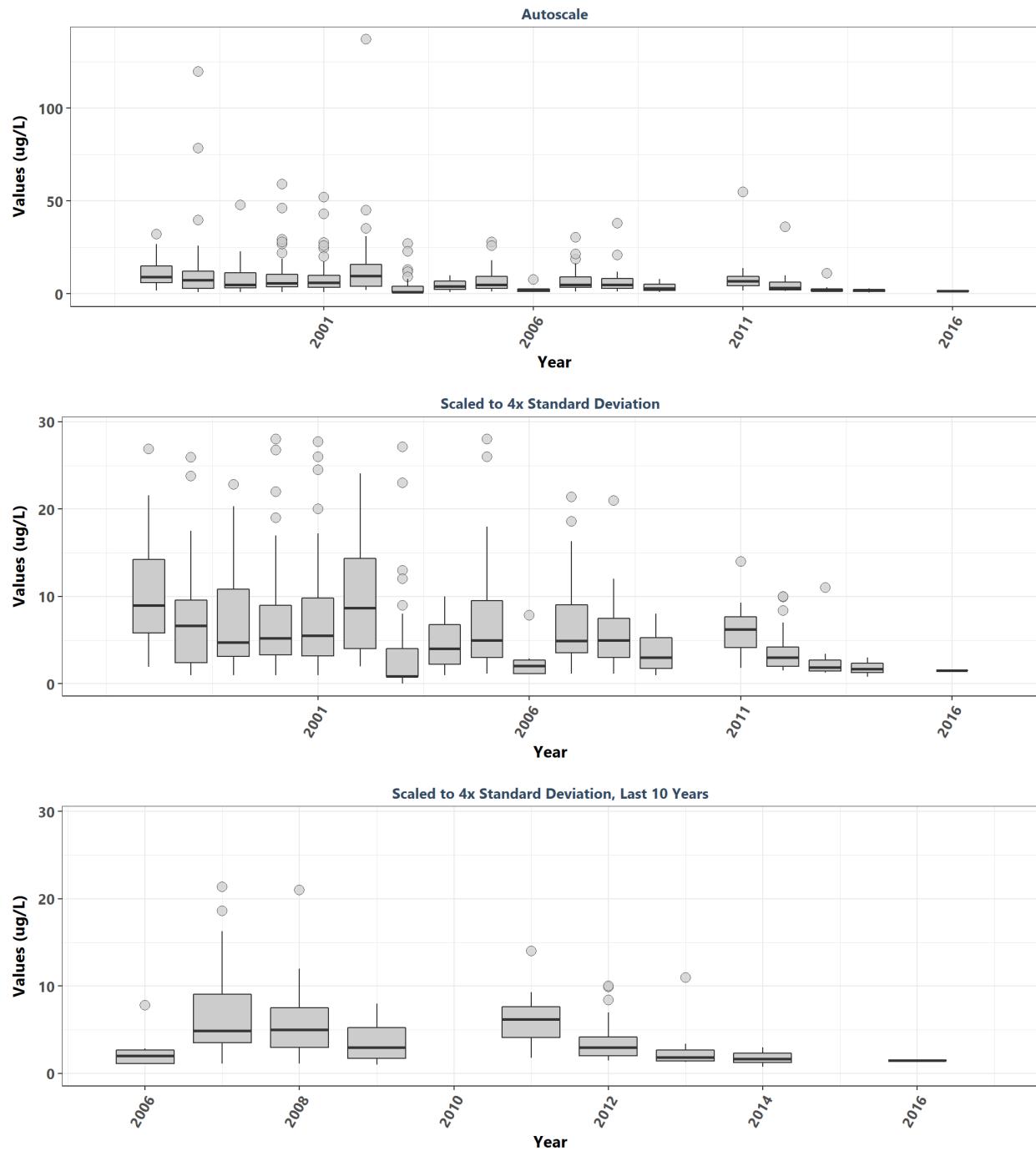
**Fort Pickens State Park Aquatic Preserve**  
By Year & Month



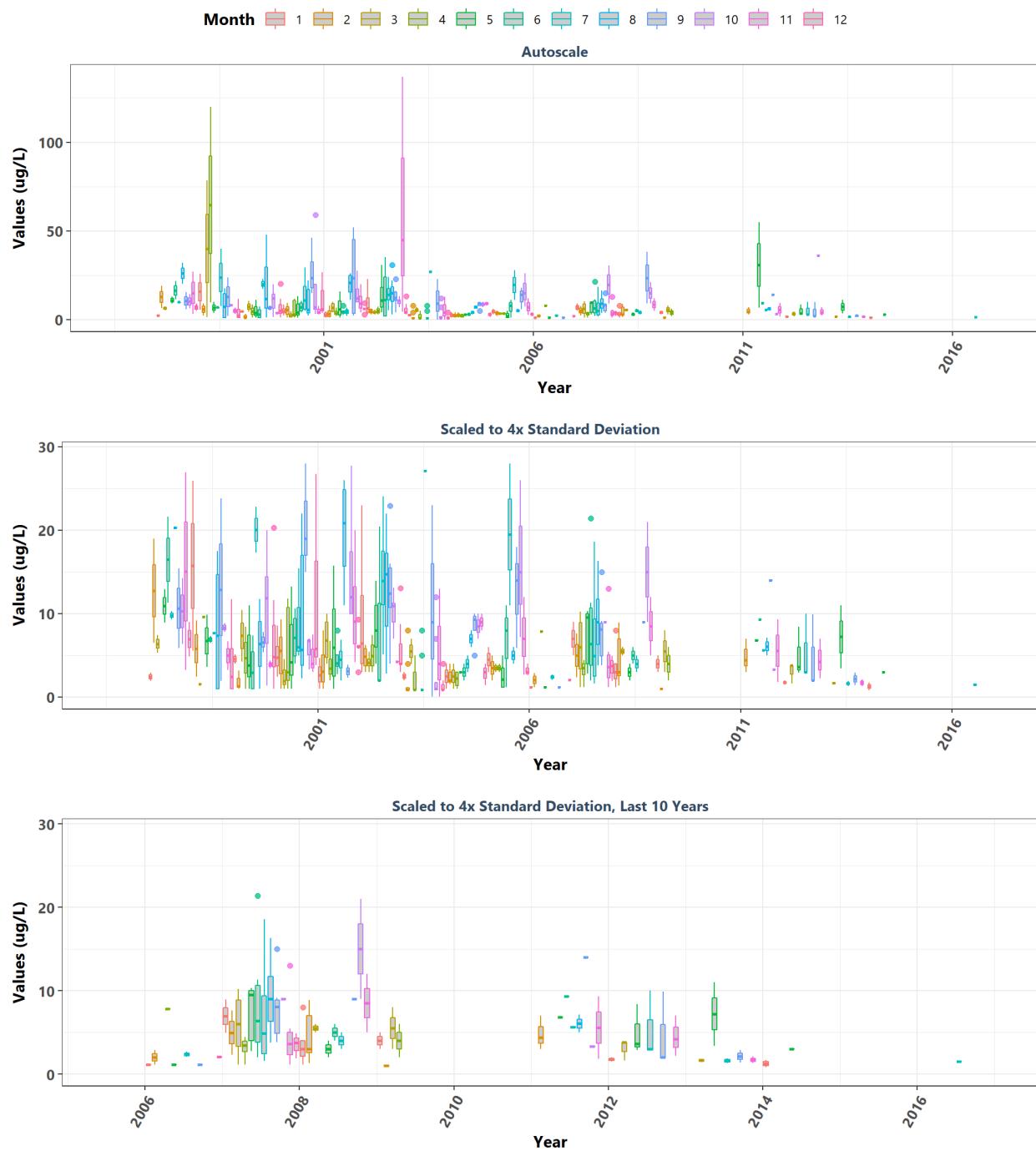
**Fort Pickens State Park Aquatic Preserve**  
By Month



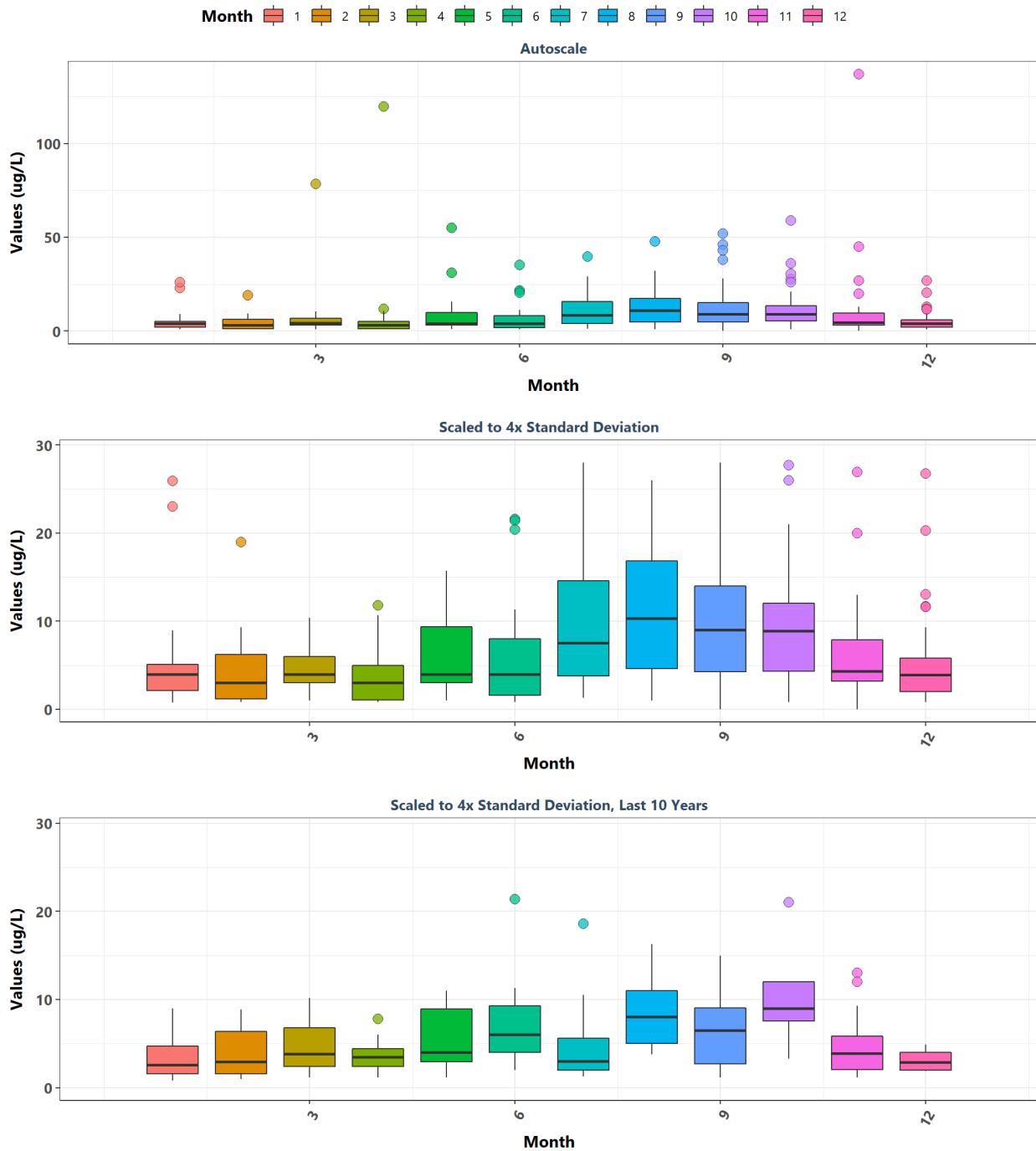
**Gasparilla Sound-Charlotte Harbor Aquatic Preserve**  
By Year



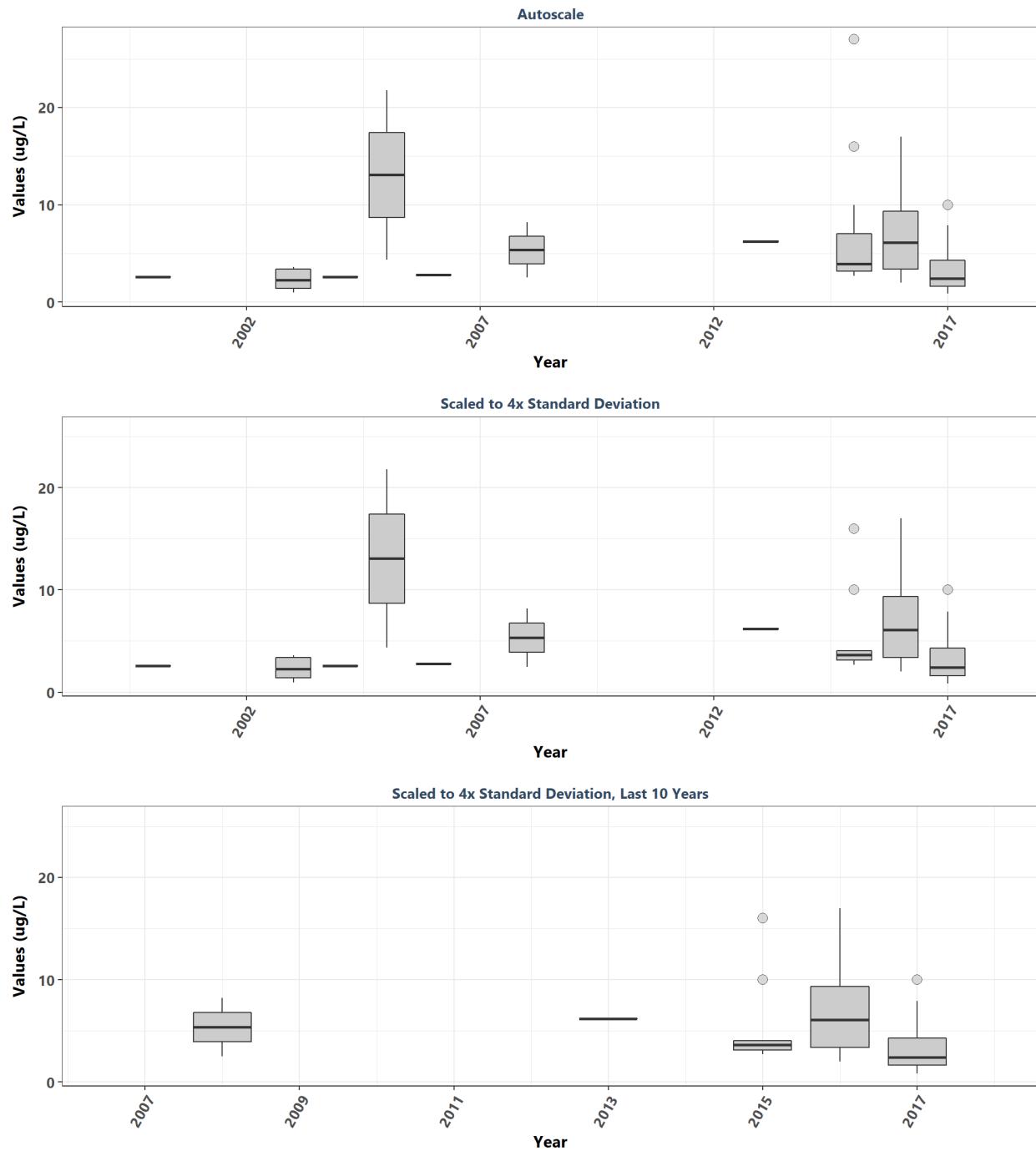
**Gasparilla Sound-Charlotte Harbor Aquatic Preserve**  
By Year & Month



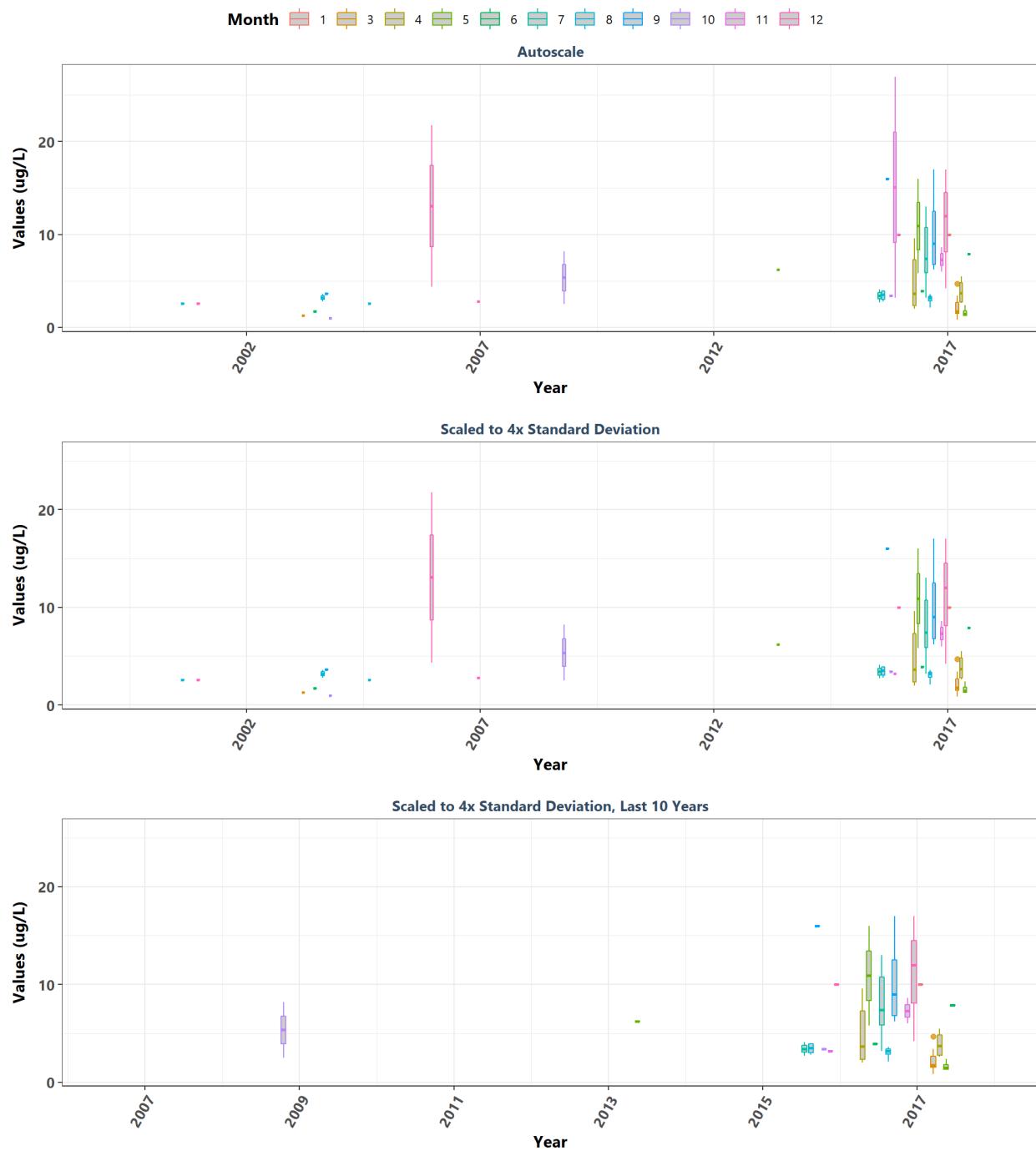
**Gasparilla Sound-Charlotte Harbor Aquatic Preserve**  
By Month



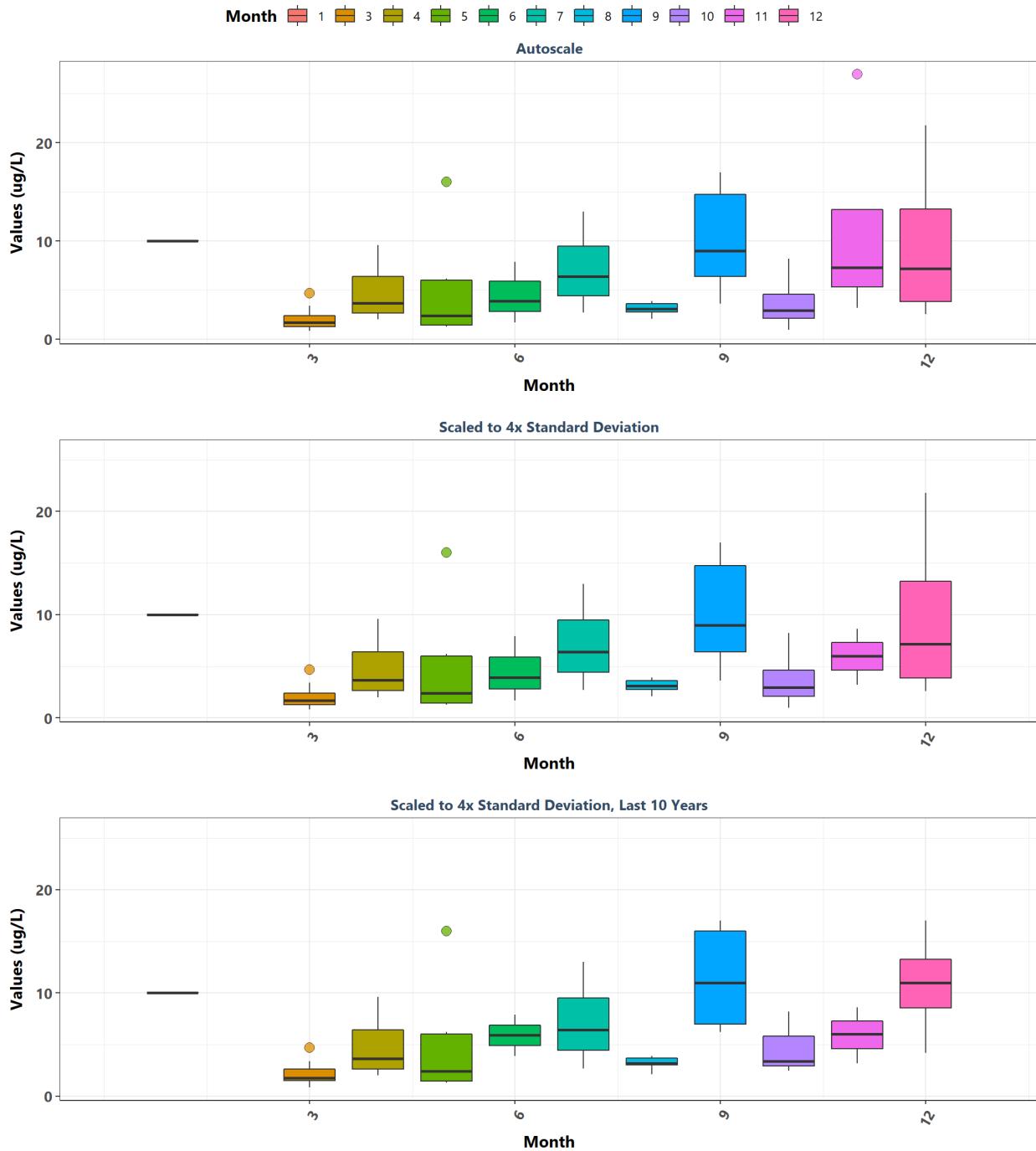
**Indian River-Malabar to Vero Beach Aquatic Preserve**  
By Year



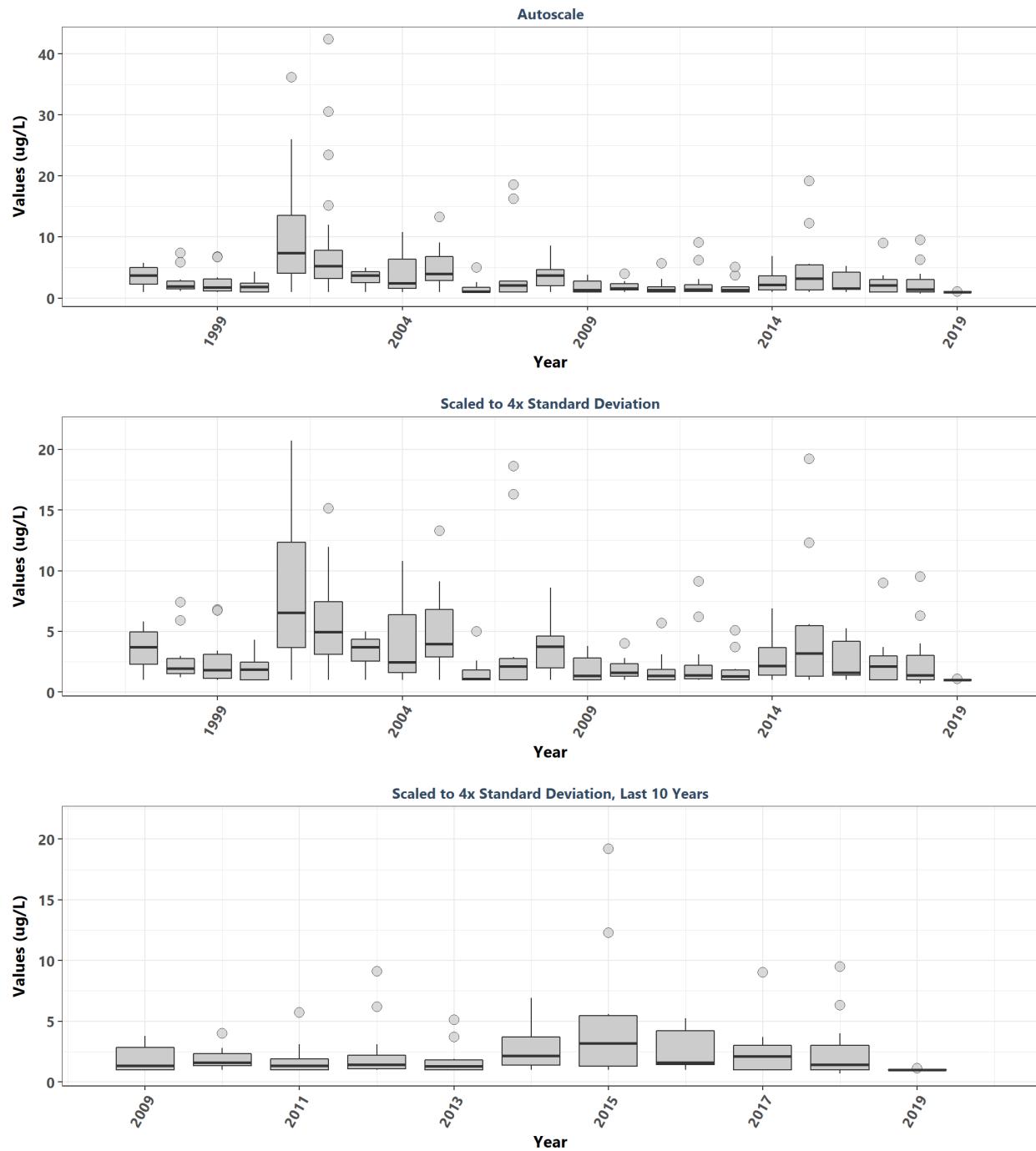
**Indian River-Malabar to Vero Beach Aquatic Preserve**  
By Year & Month



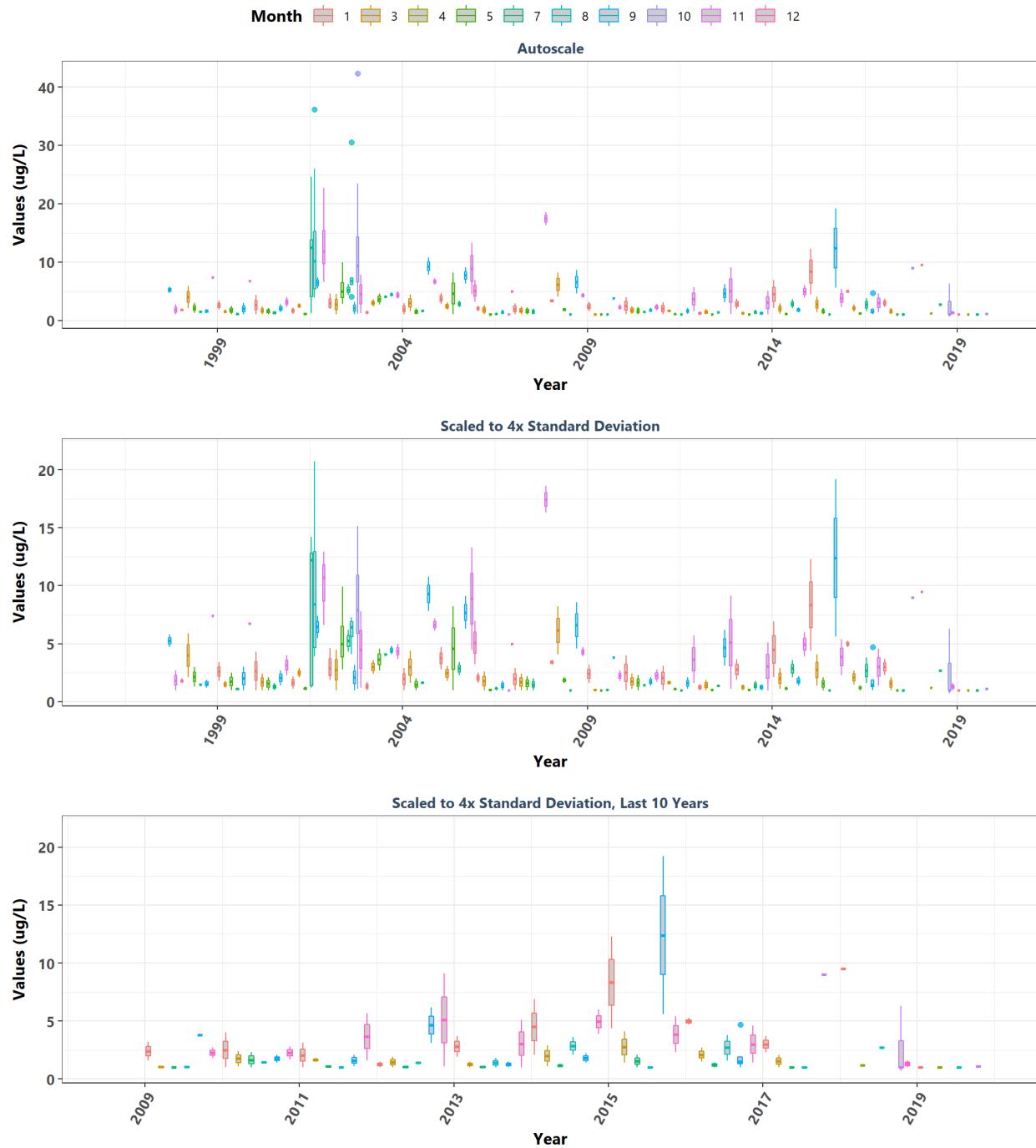
**Indian River-Malabar to Vero Beach Aquatic Preserve**  
By Month



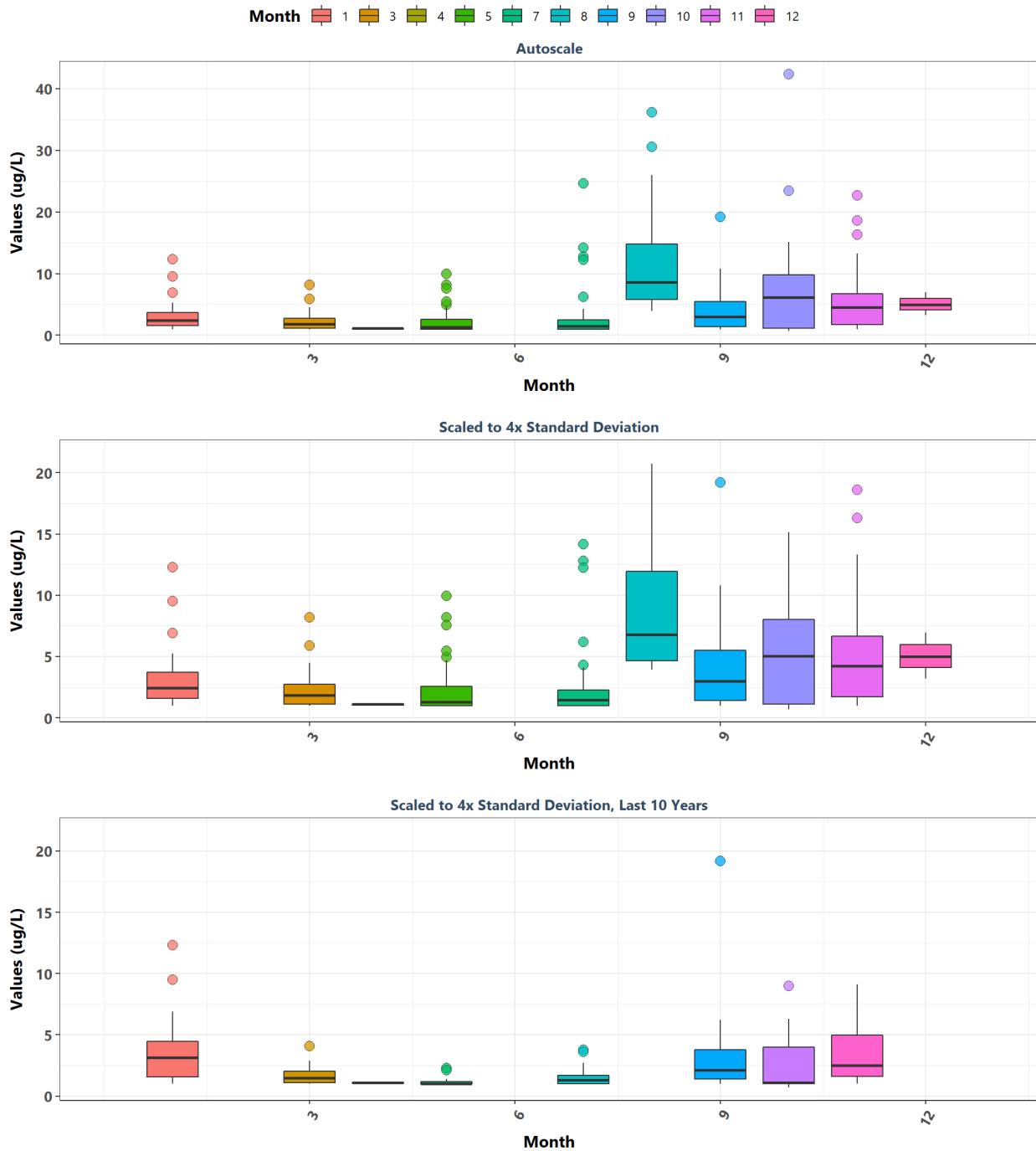
**Jensen Beach to Jupiter Inlet Aquatic Preserve**  
By Year



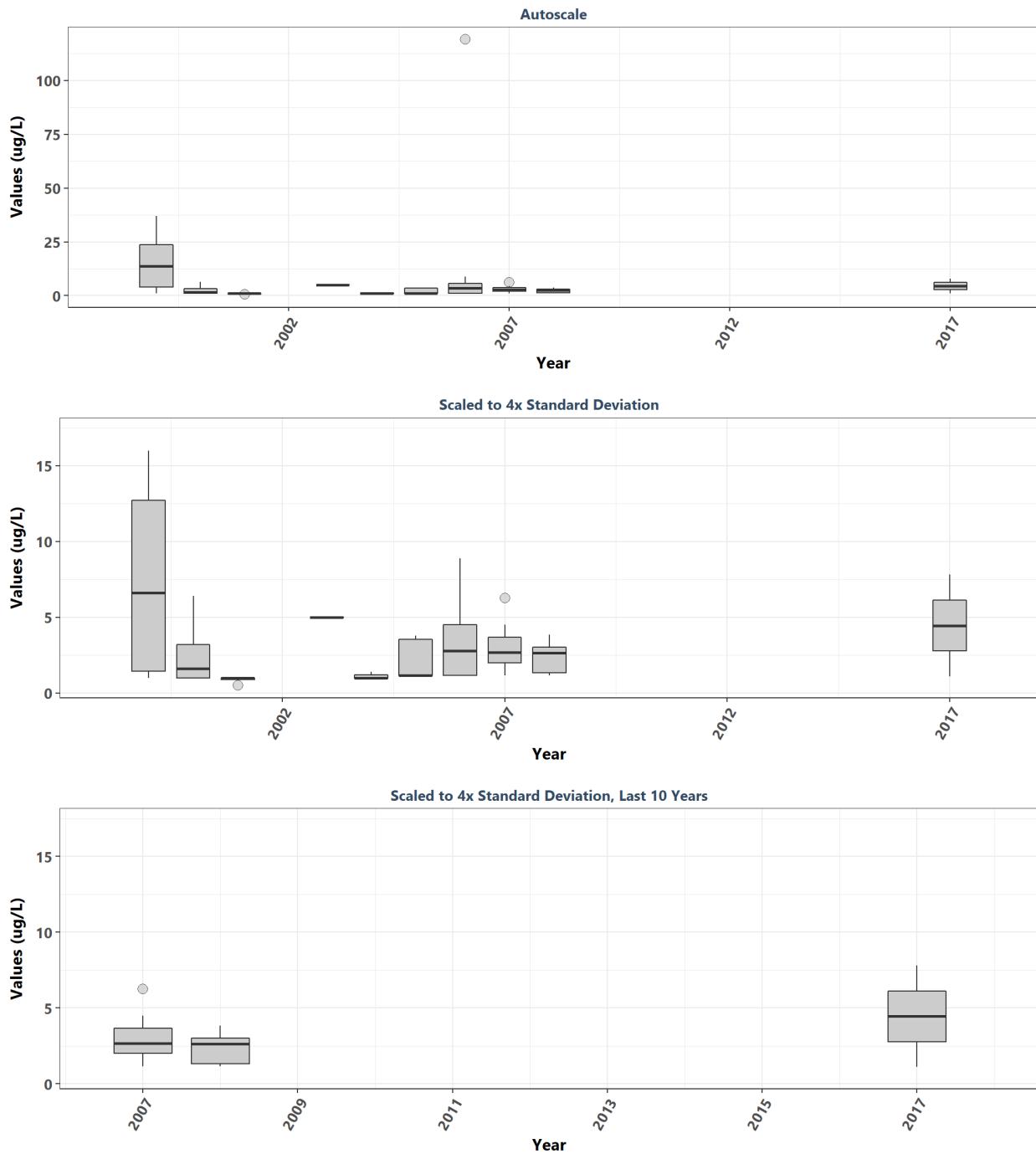
**Jensen Beach to Jupiter Inlet Aquatic Preserve**  
By Year & Month



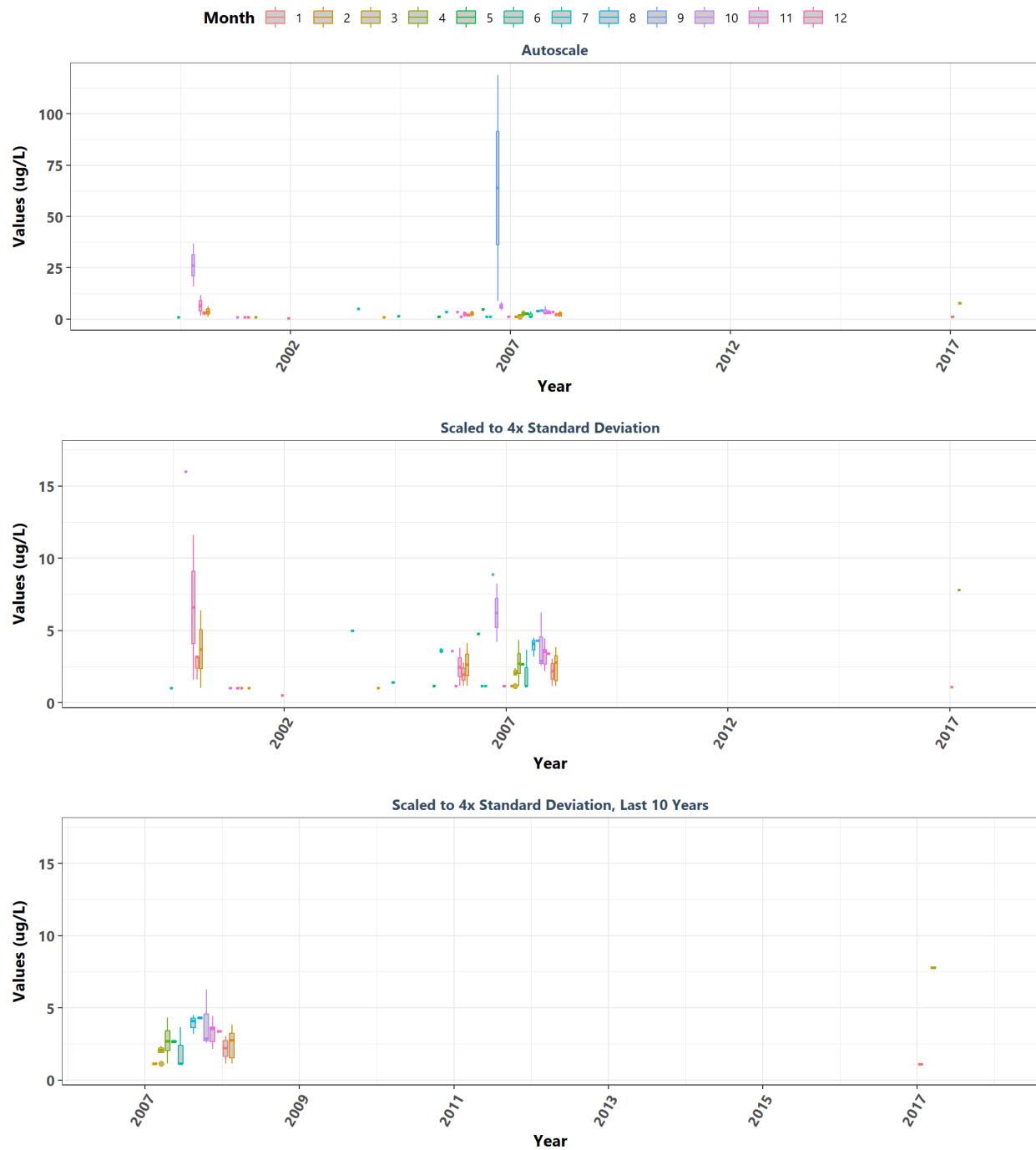
**Jensen Beach to Jupiter Inlet Aquatic Preserve**  
By Month



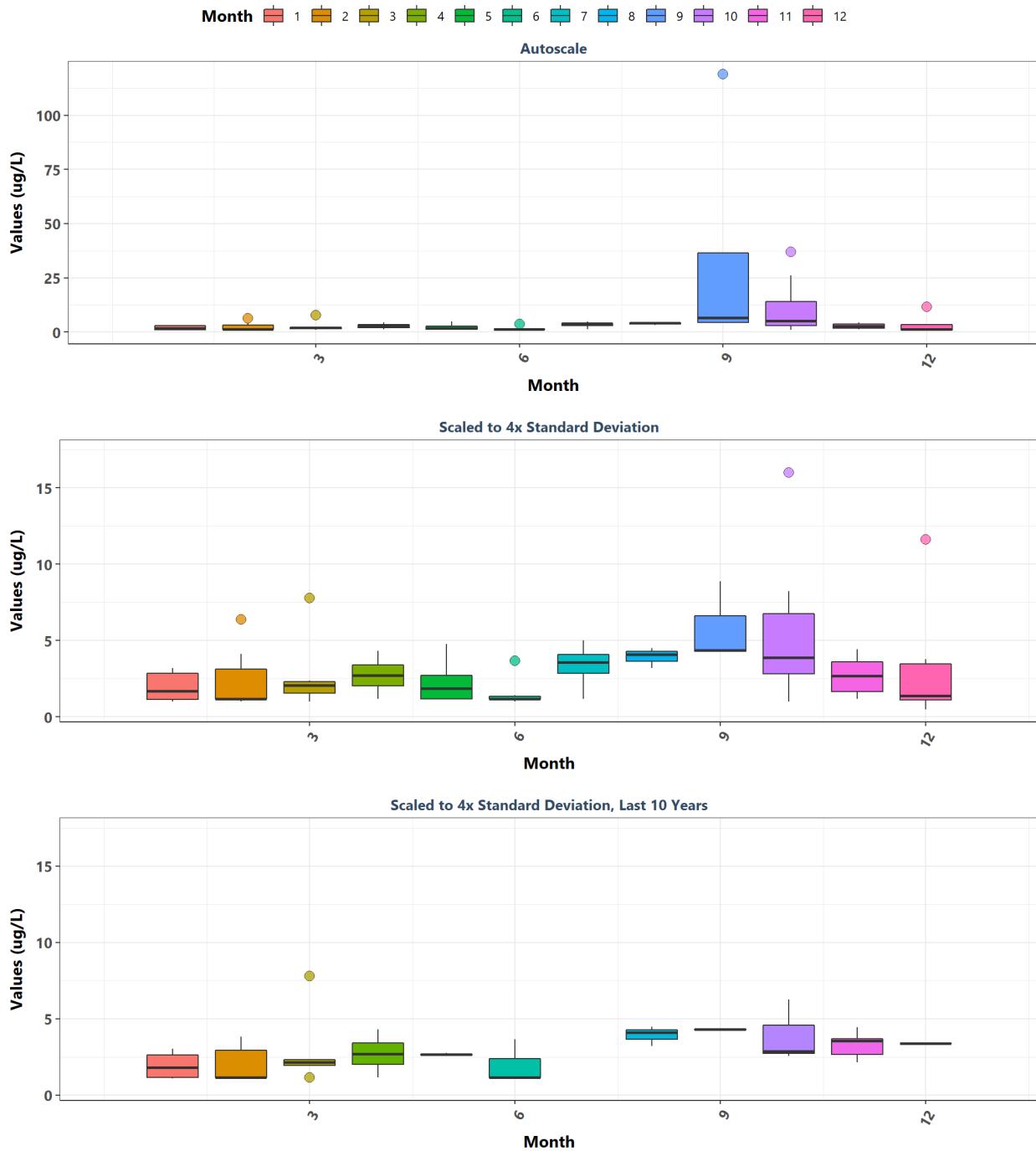
**Lemon Bay Aquatic Preserve**  
By Year



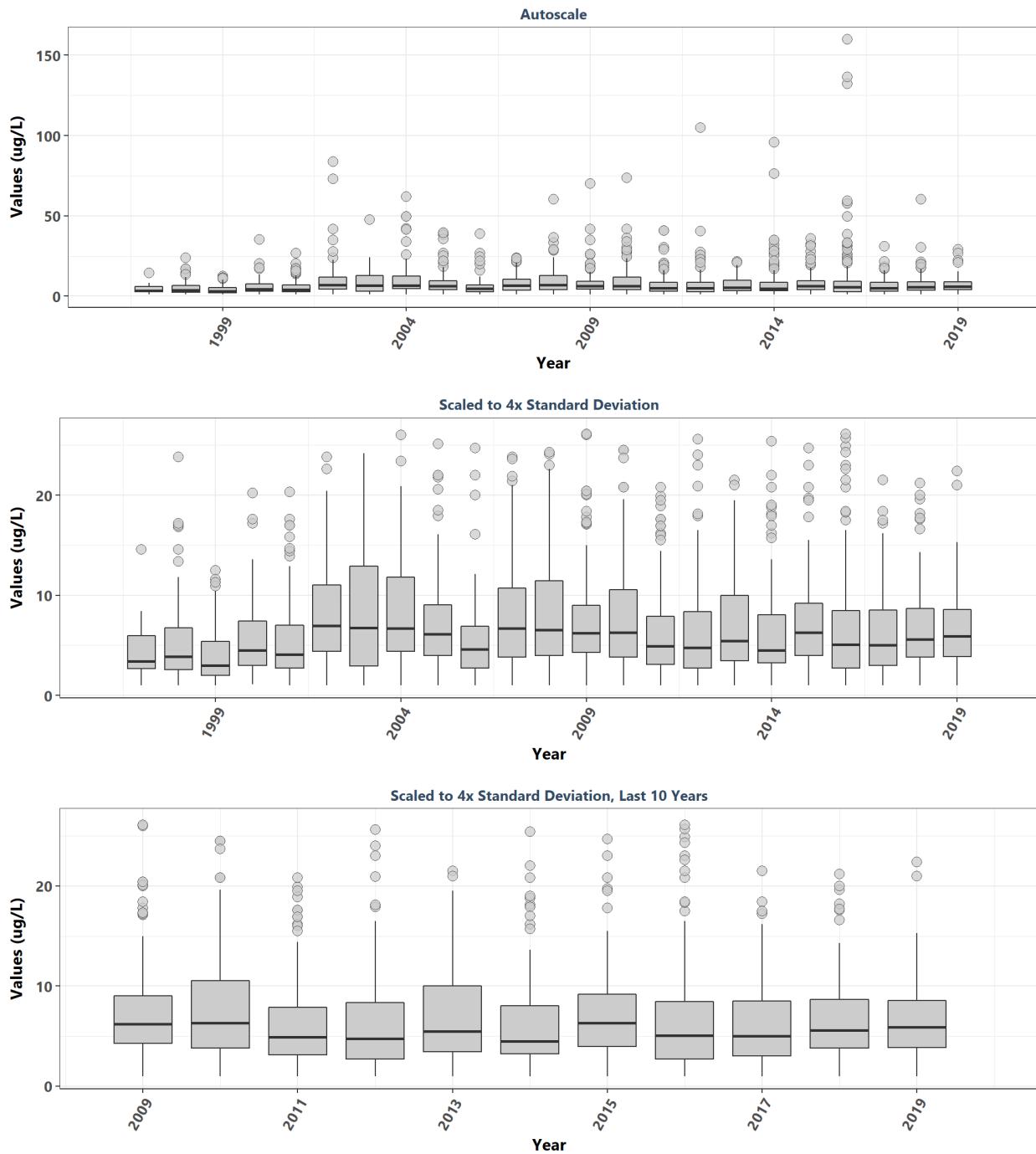
**Lemon Bay Aquatic Preserve**  
By Year & Month



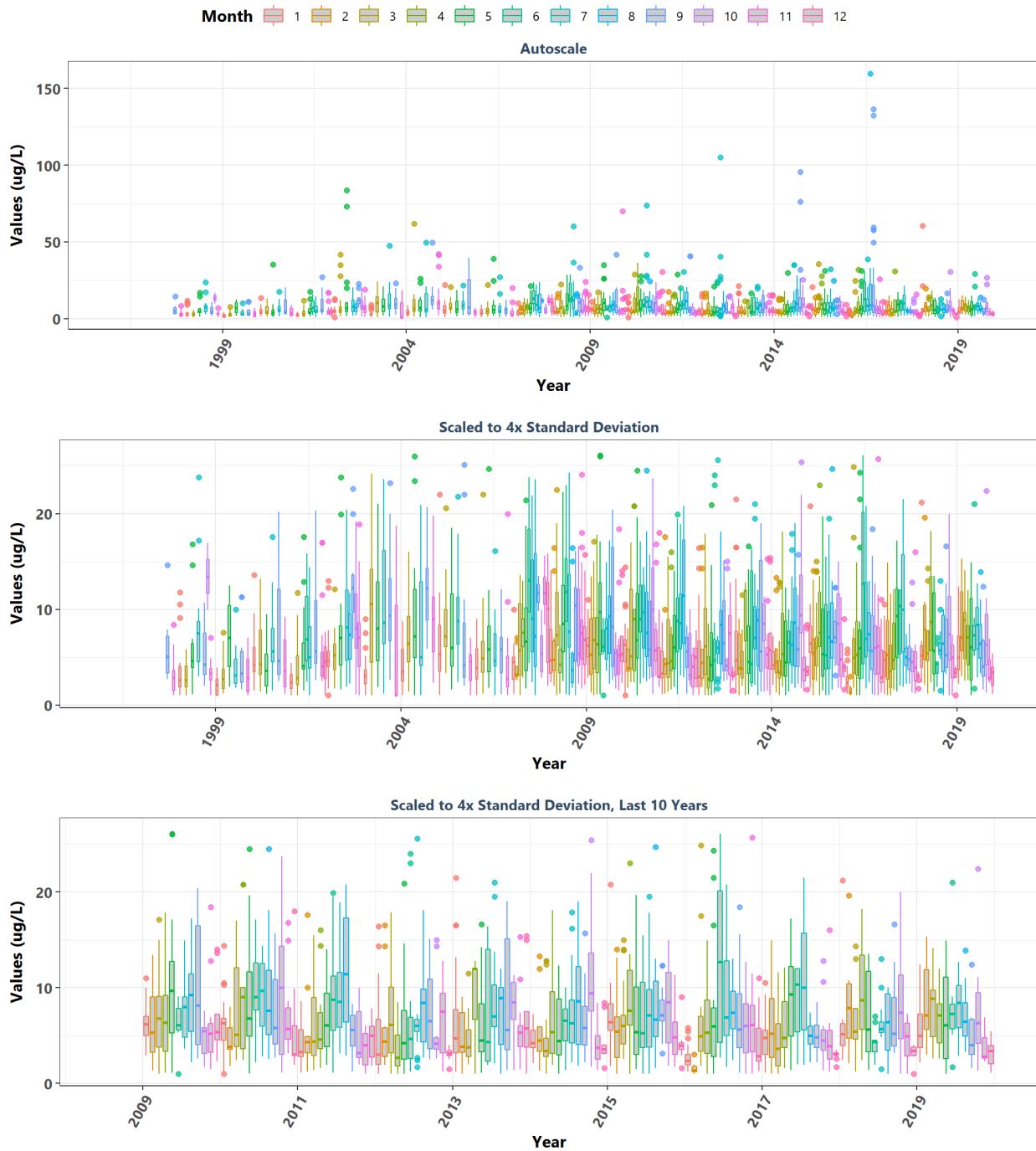
**Lemon Bay Aquatic Preserve**  
By Month



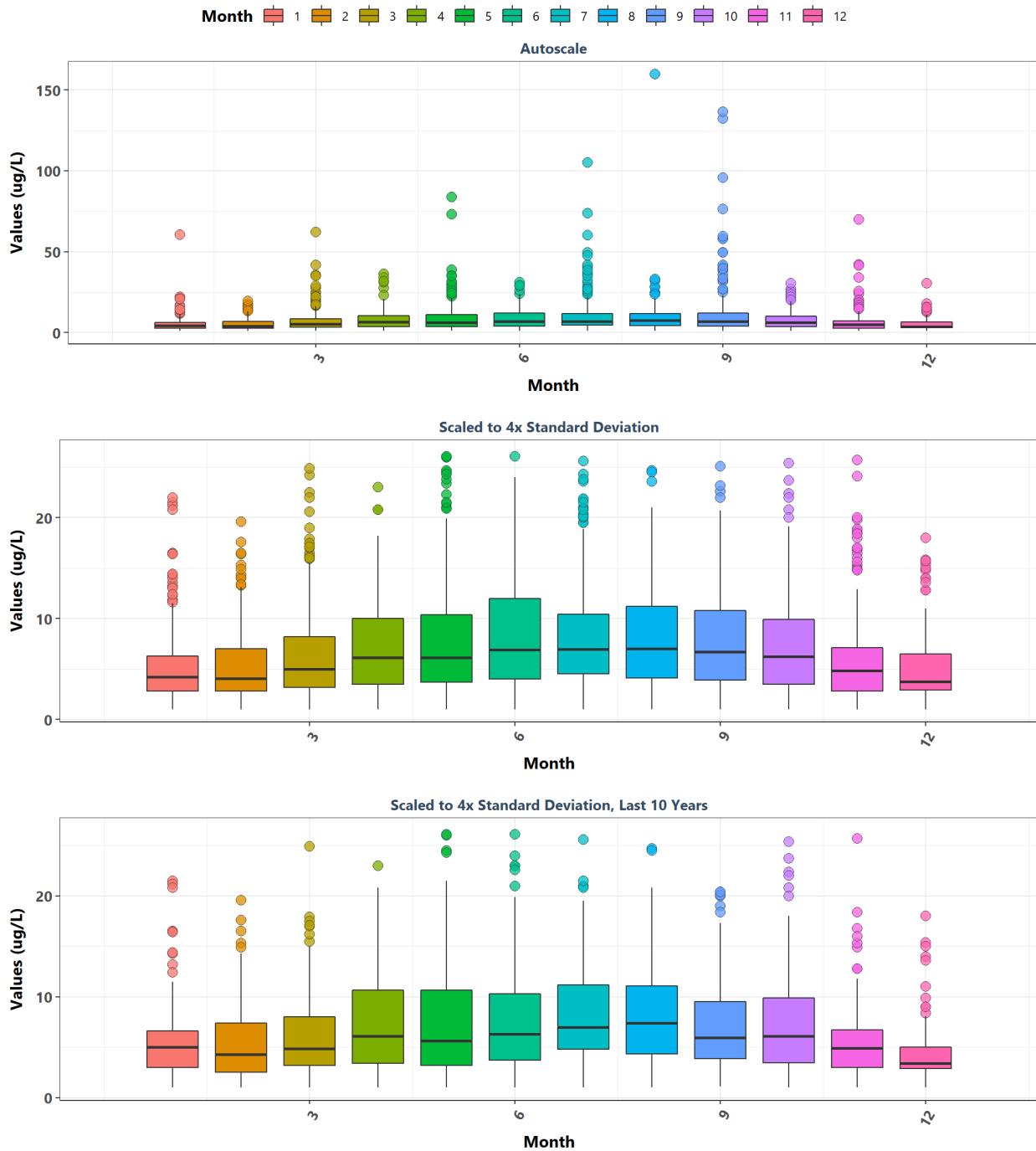
**Loxahatchee River-Lake Worth Creek Aquatic Preserve**  
By Year



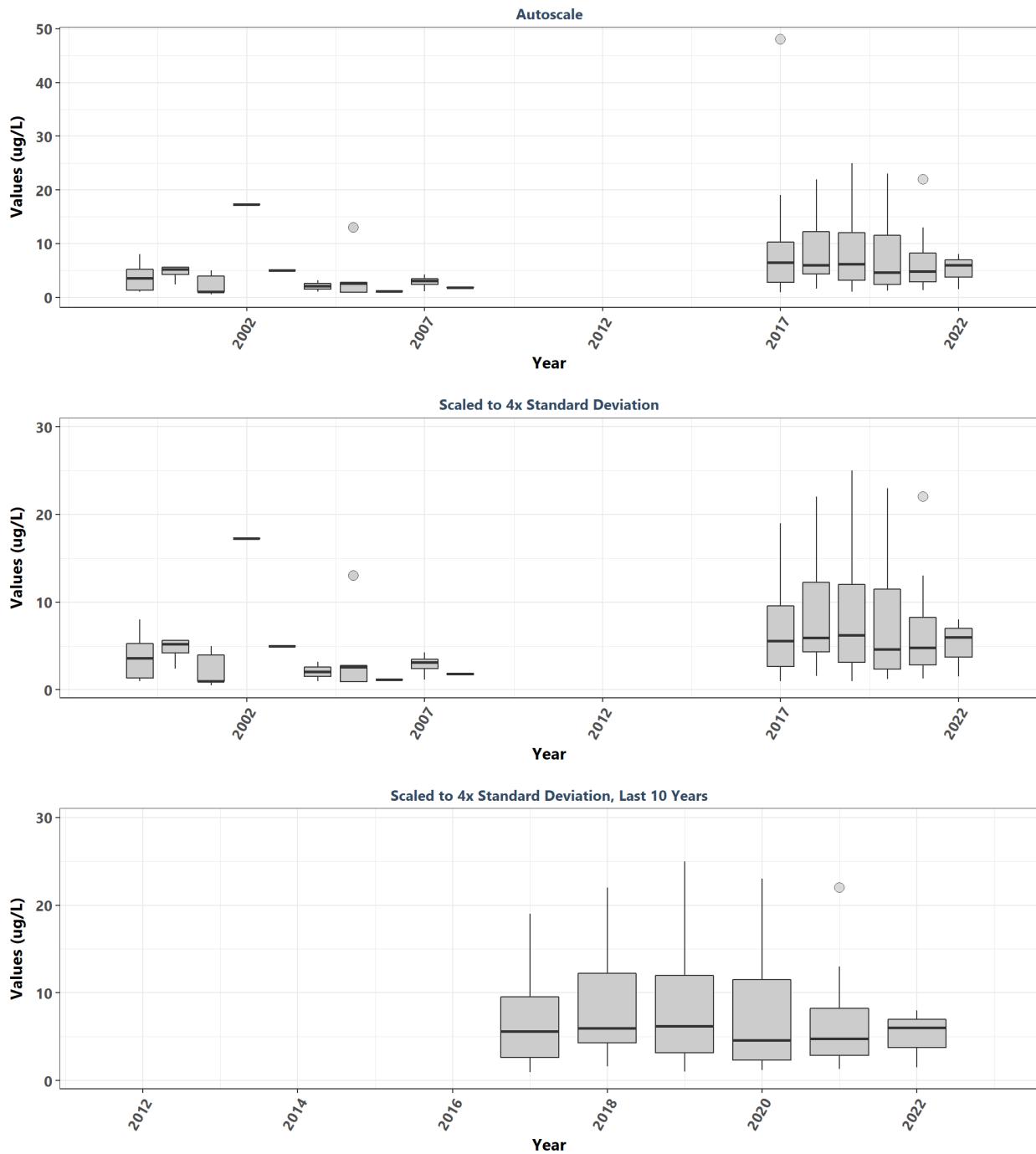
**Loxahatchee River-Lake Worth Creek Aquatic Preserve**  
By Year & Month



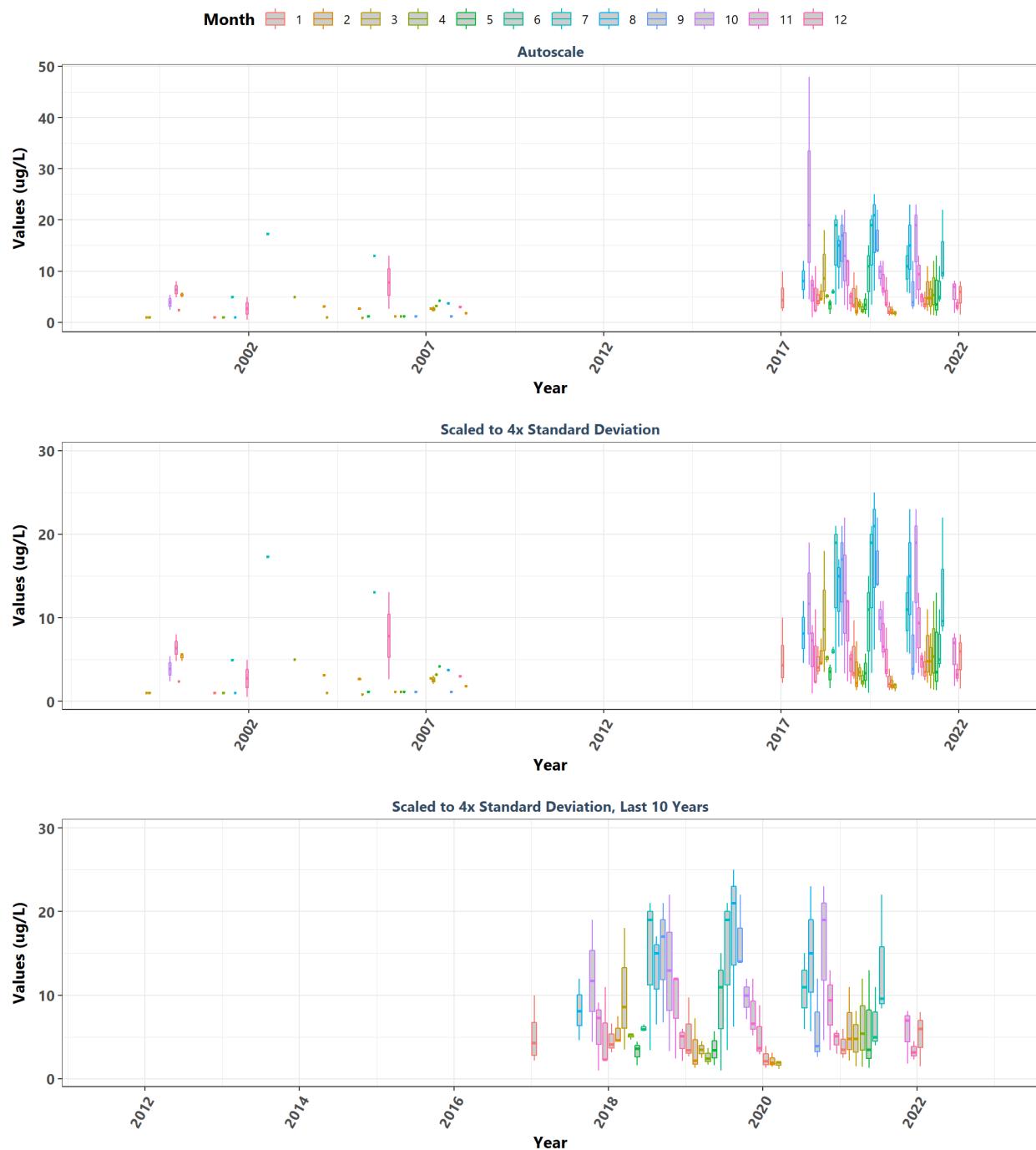
**Loxahatchee River-Lake Worth Creek Aquatic Preserve**  
By Month



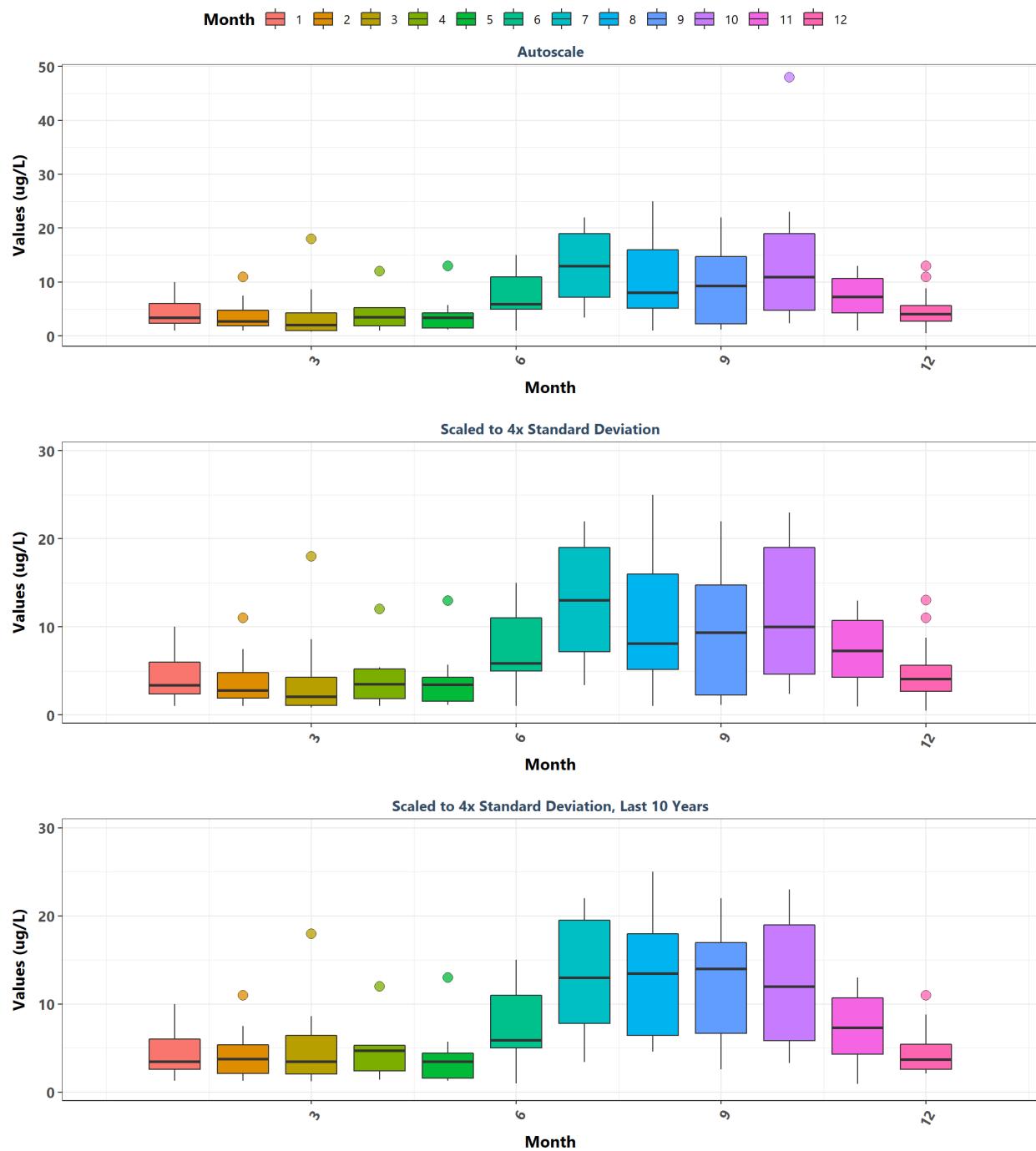
**Matlacha Pass Aquatic Preserve**  
By Year



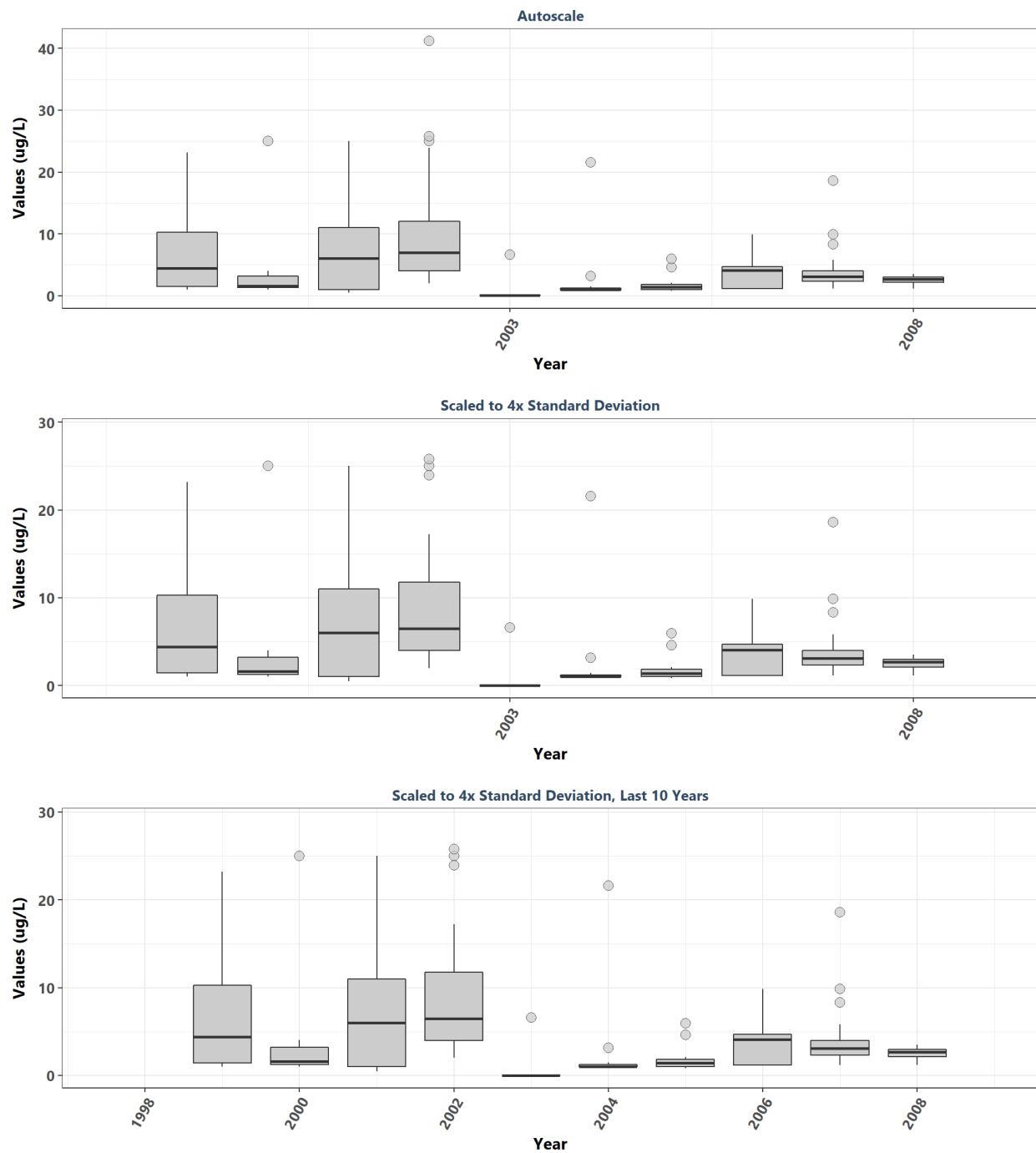
**Matlacha Pass Aquatic Preserve**  
By Year & Month



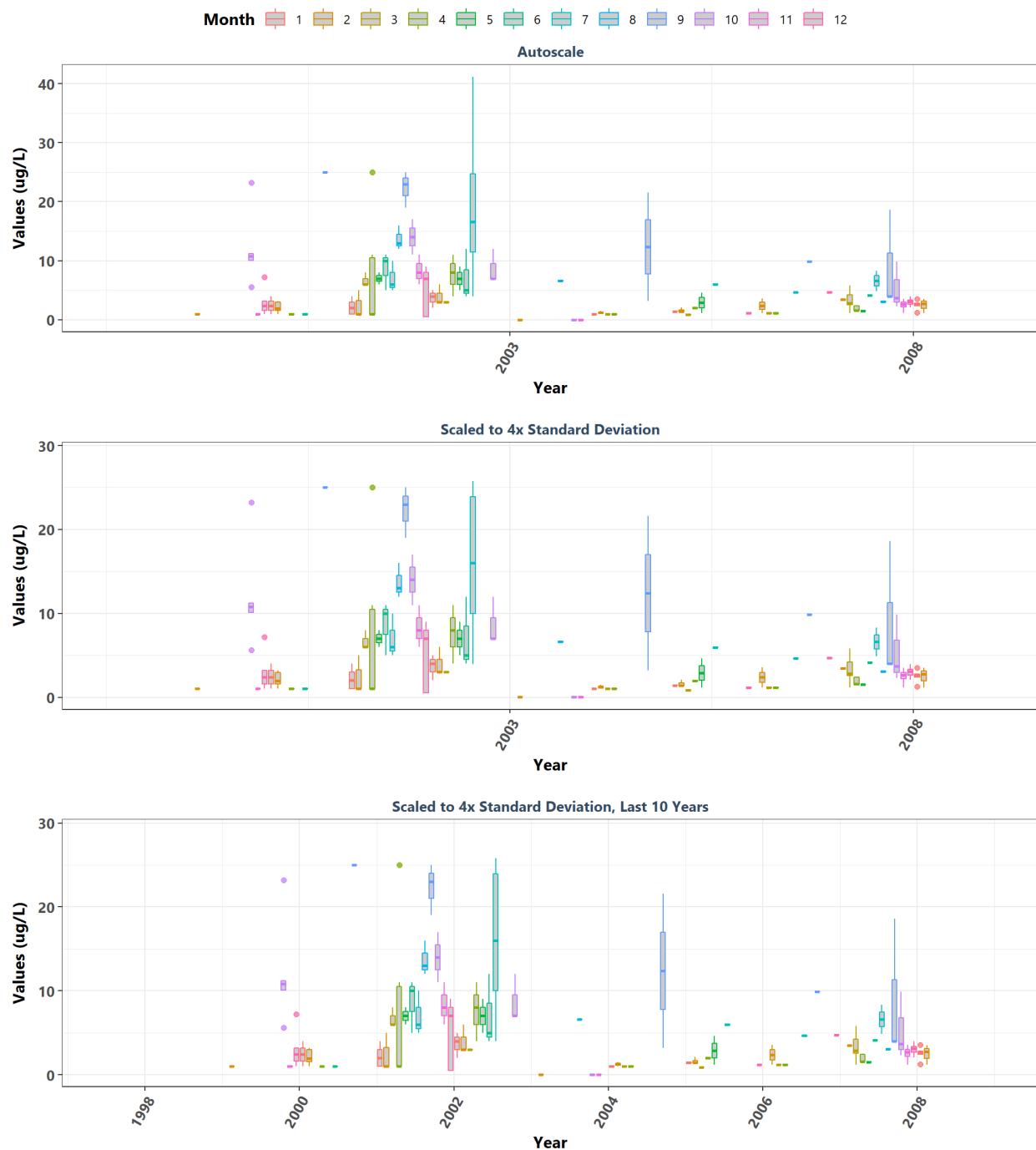
**Matlacha Pass Aquatic Preserve**  
By Month



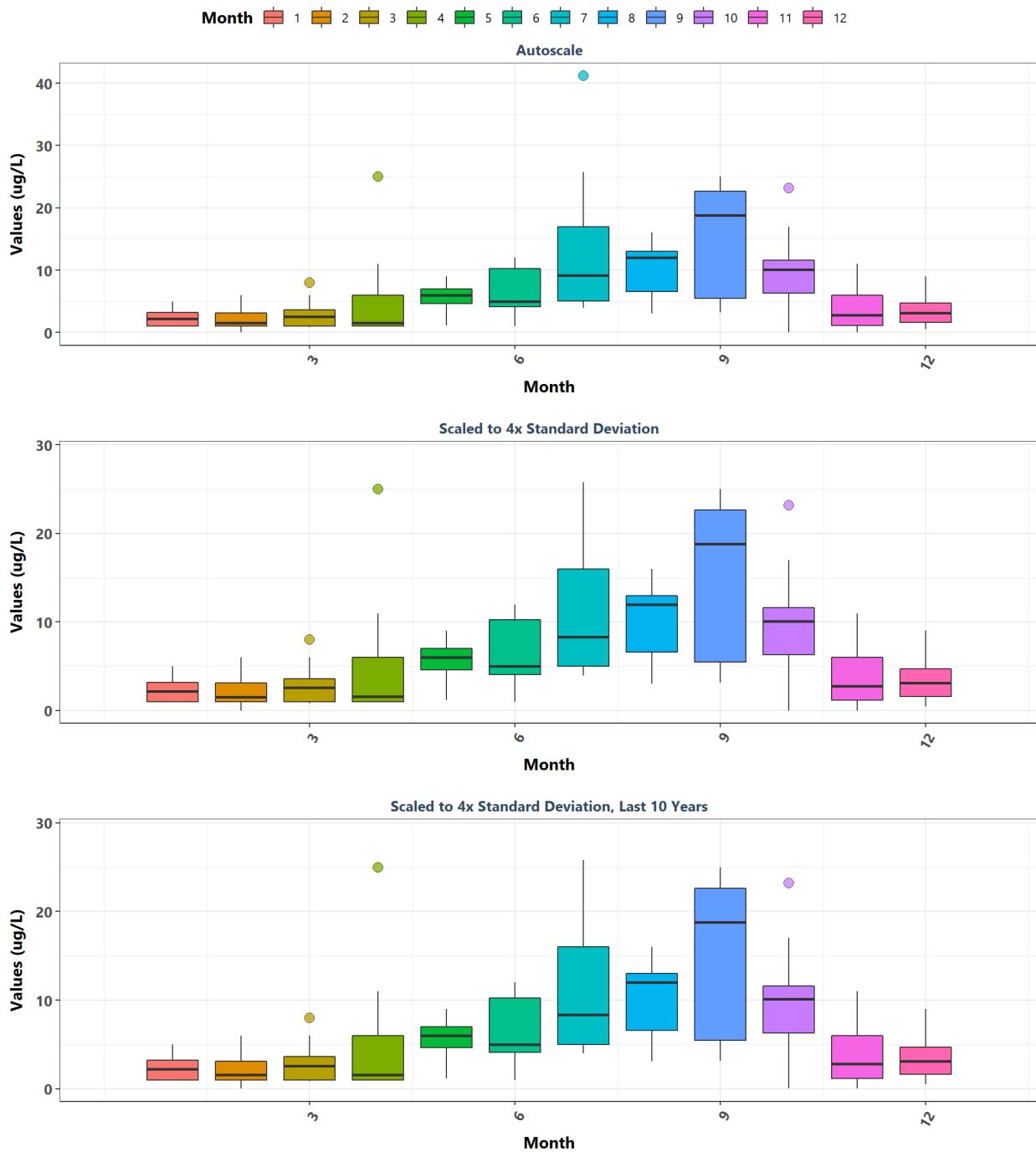
**Pine Island Sound Aquatic Preserve**  
By Year



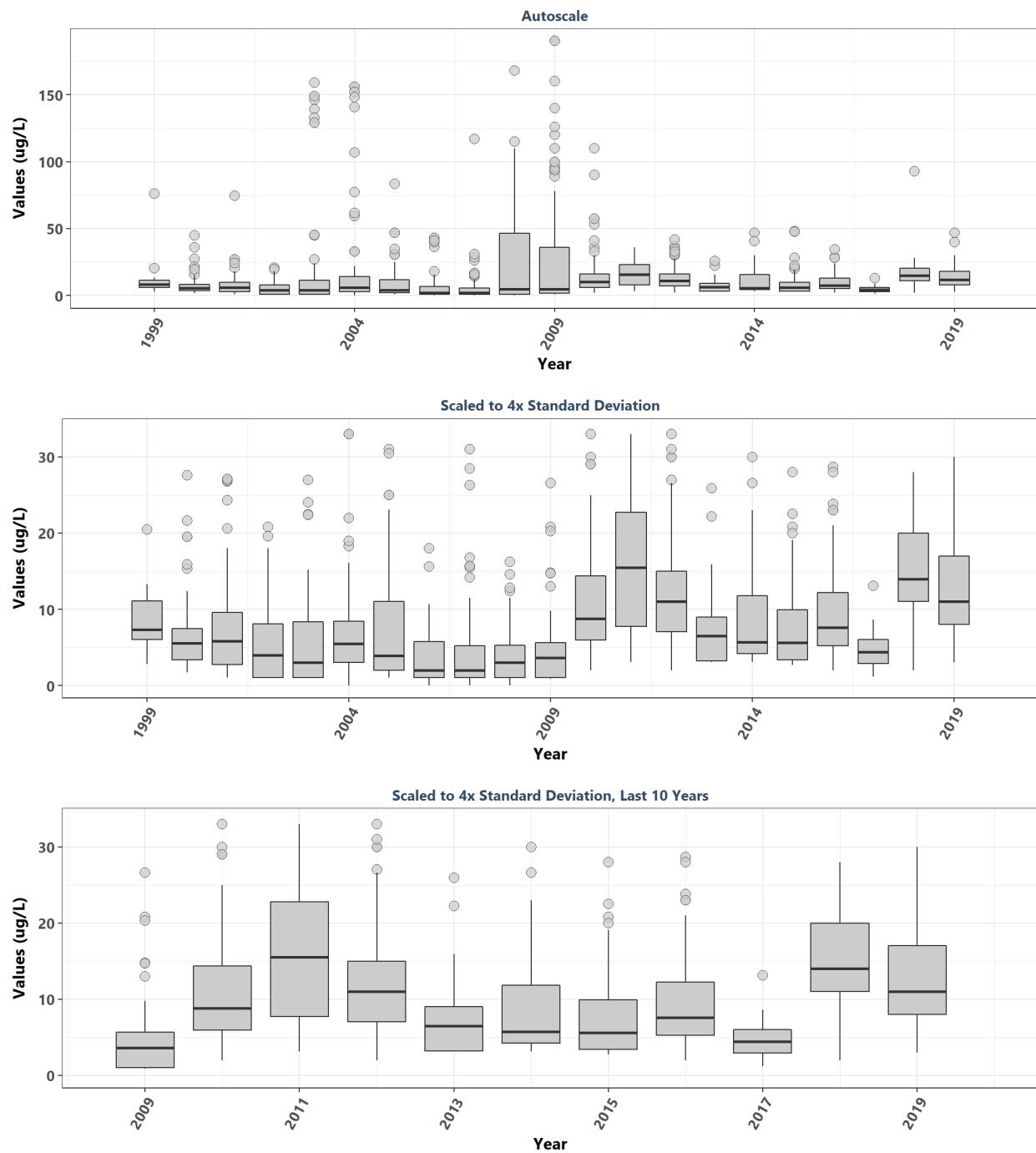
**Pine Island Sound Aquatic Preserve**  
By Year & Month



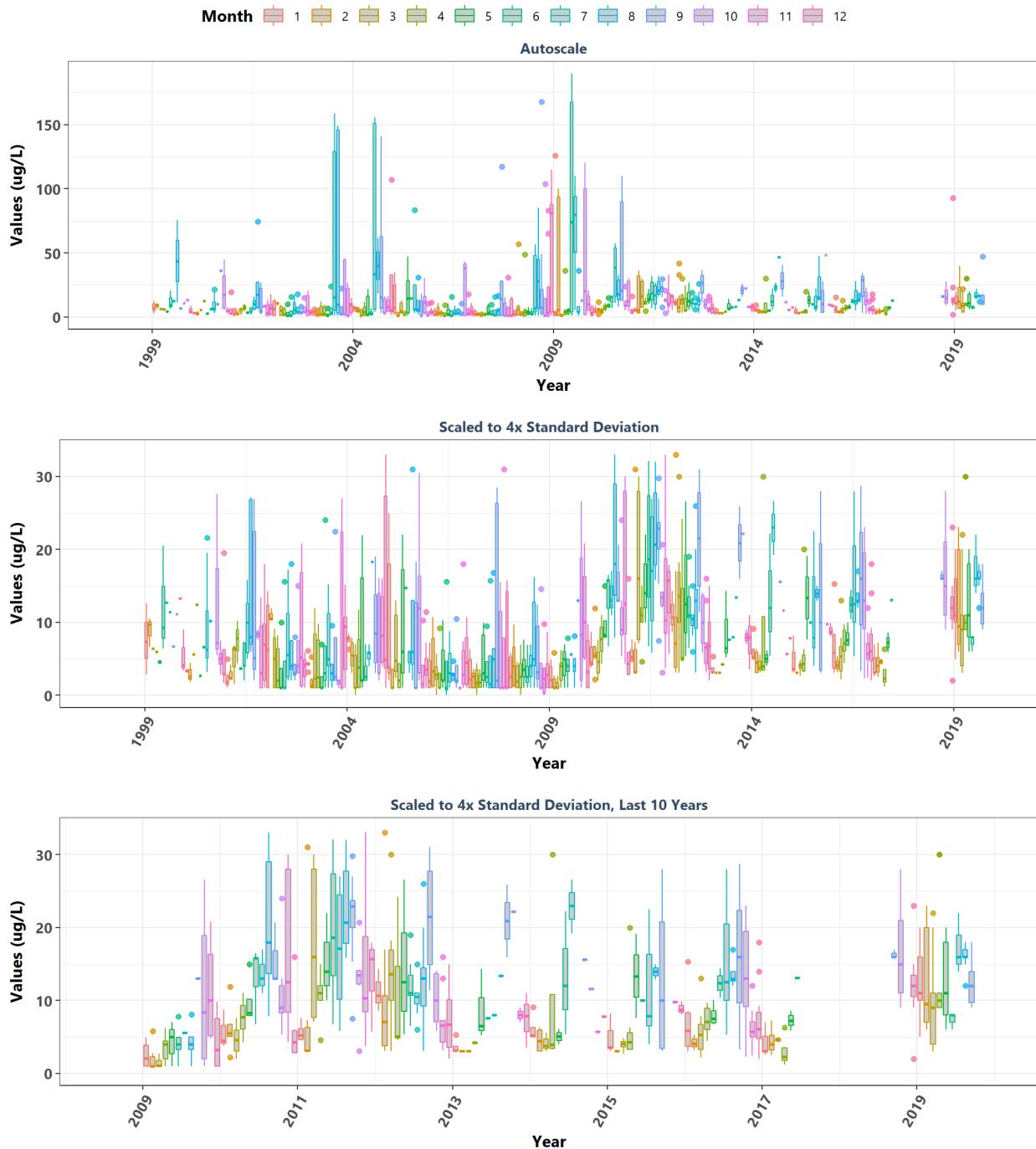
**Pine Island Sound Aquatic Preserve**  
By Month



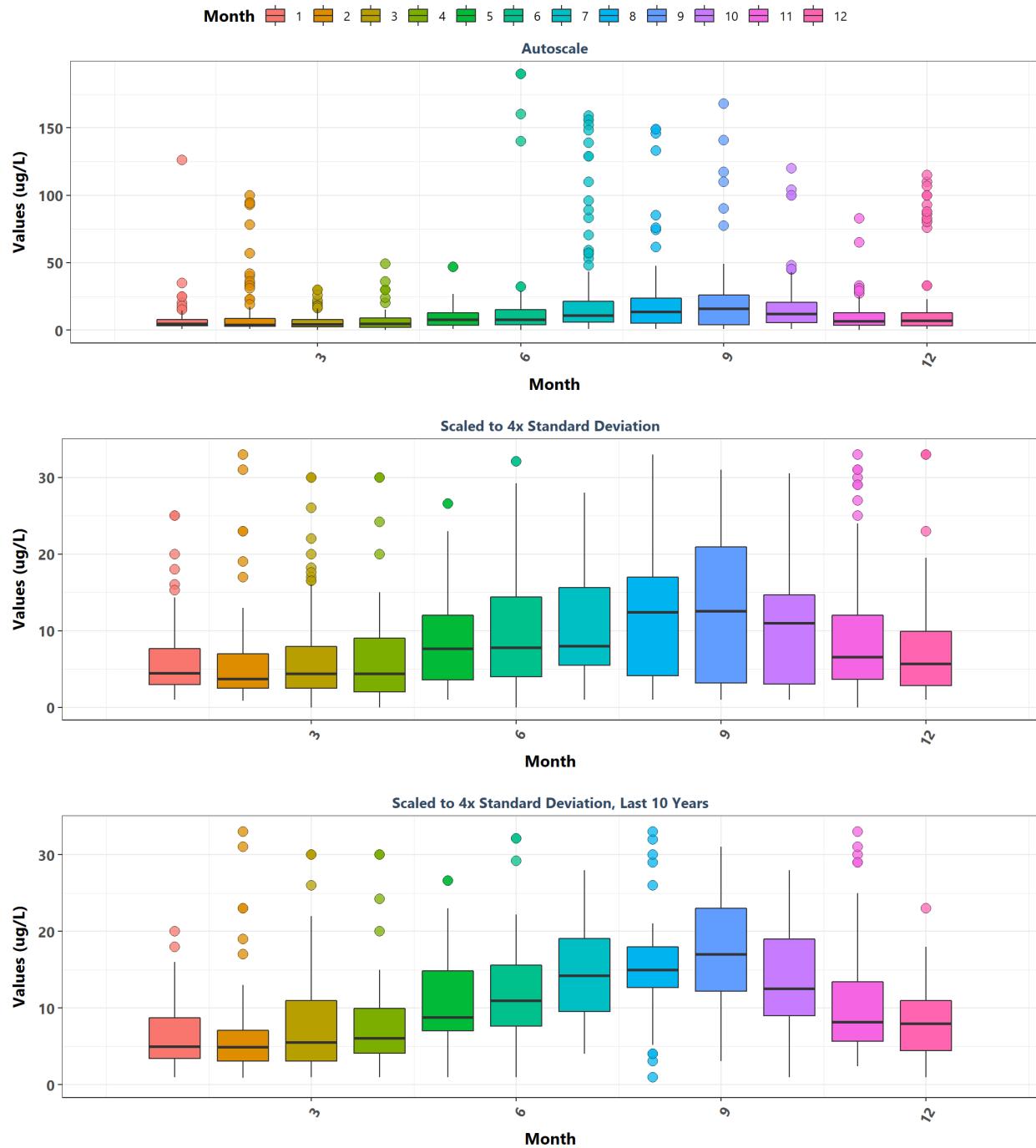
**Pinellas County Aquatic Preserve**  
By Year



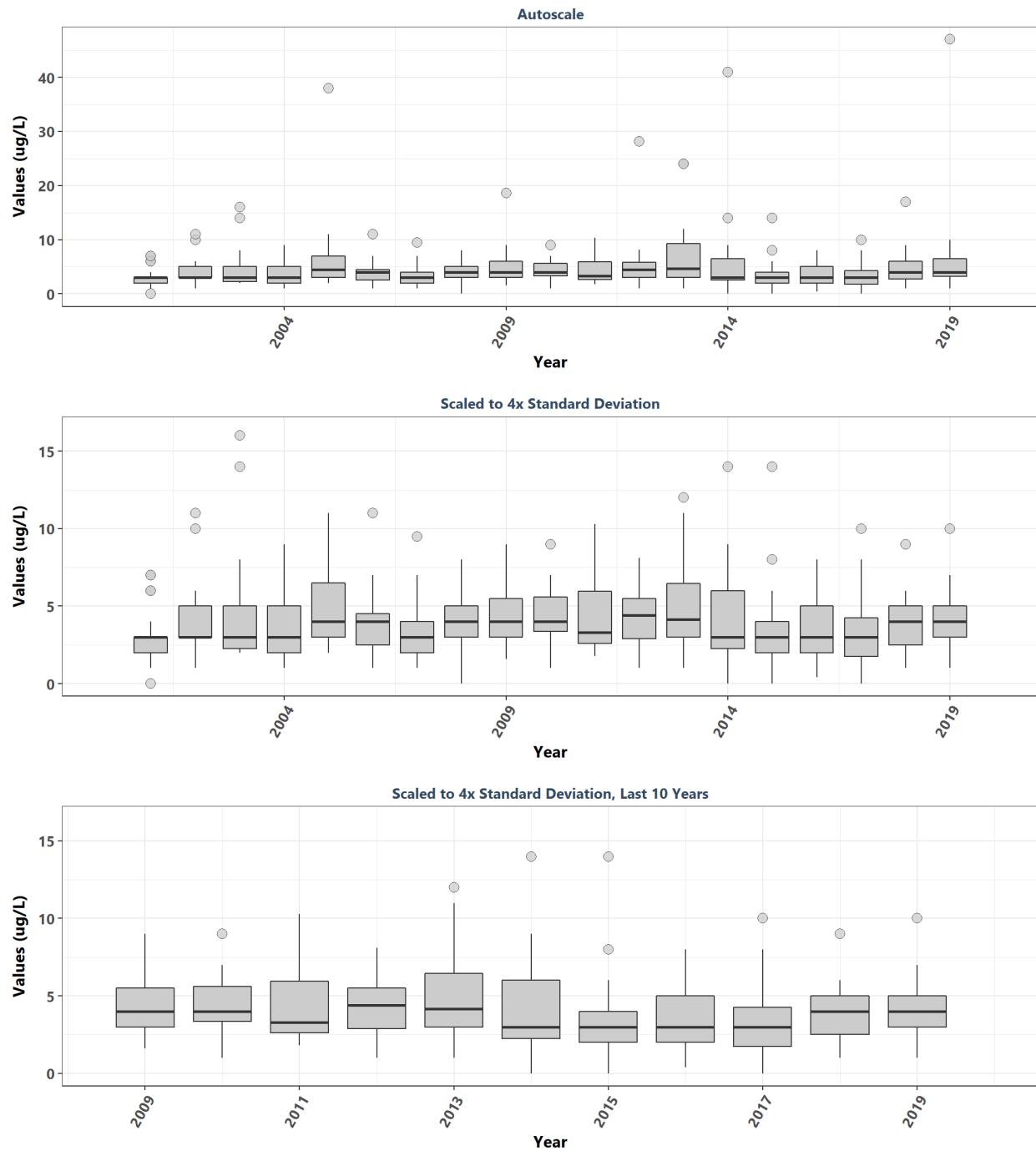
**Pinellas County Aquatic Preserve**  
By Year & Month



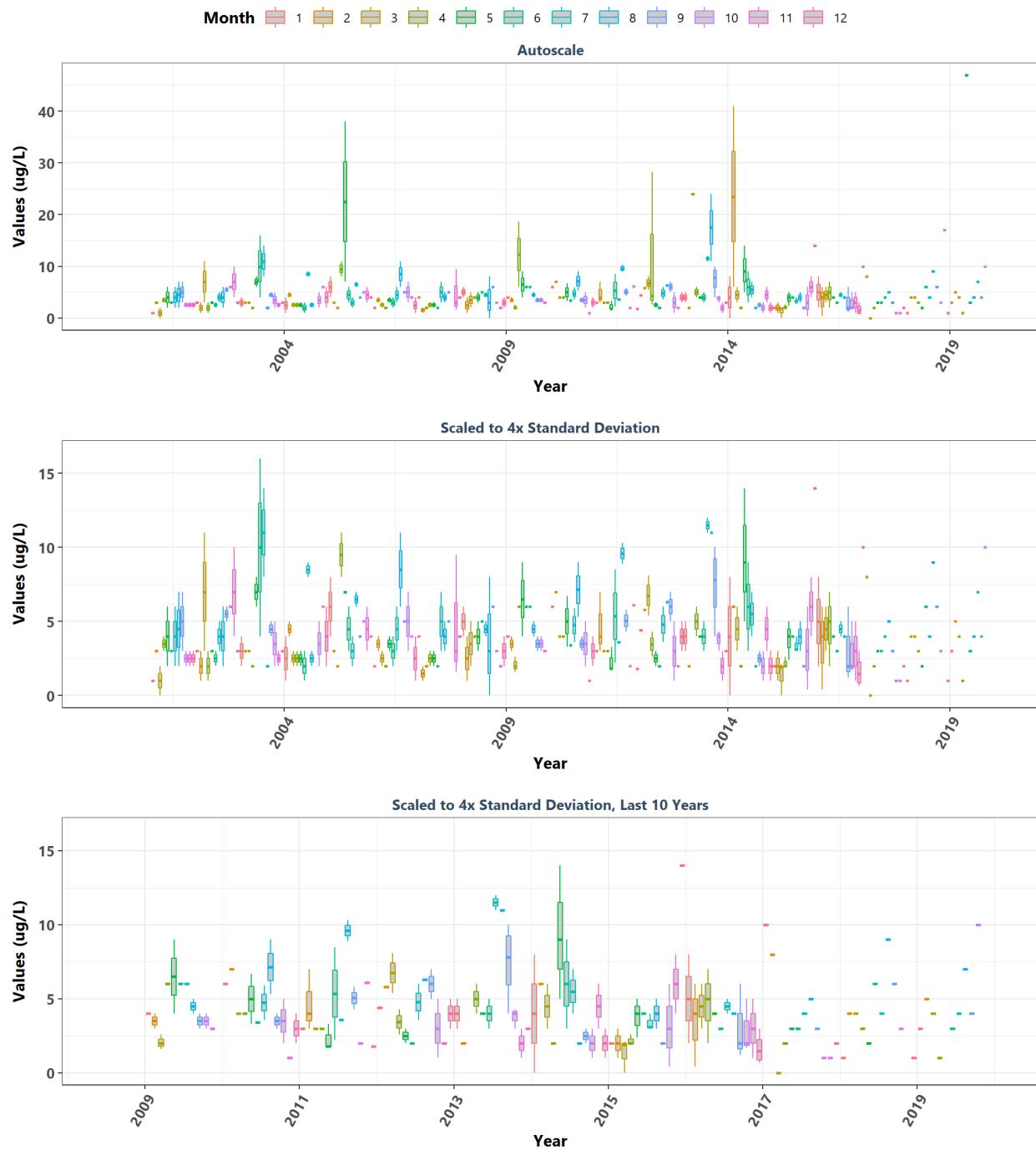
**Pinellas County Aquatic Preserve**  
By Month



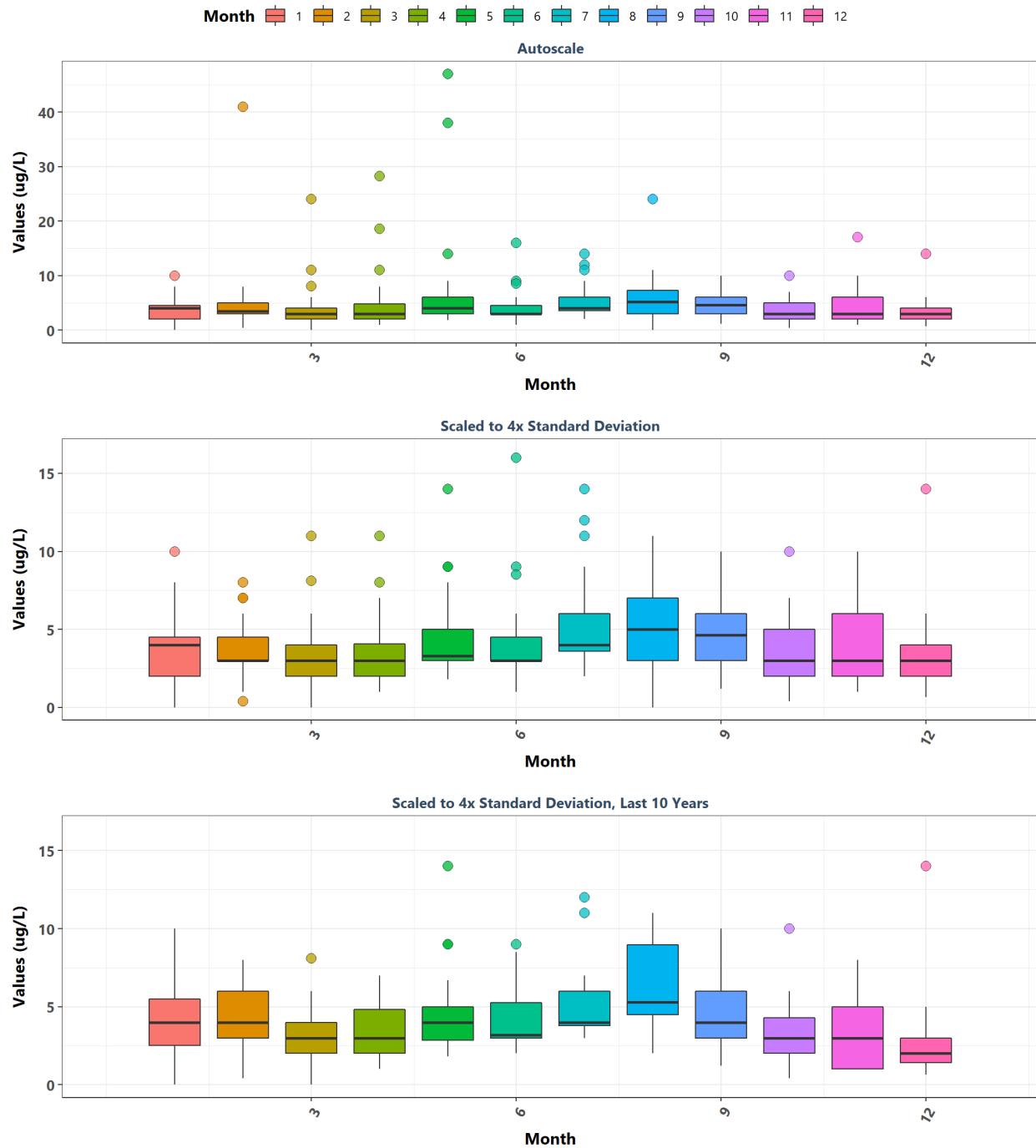
**Rocky Bayou State Park Aquatic Preserve**  
By Year



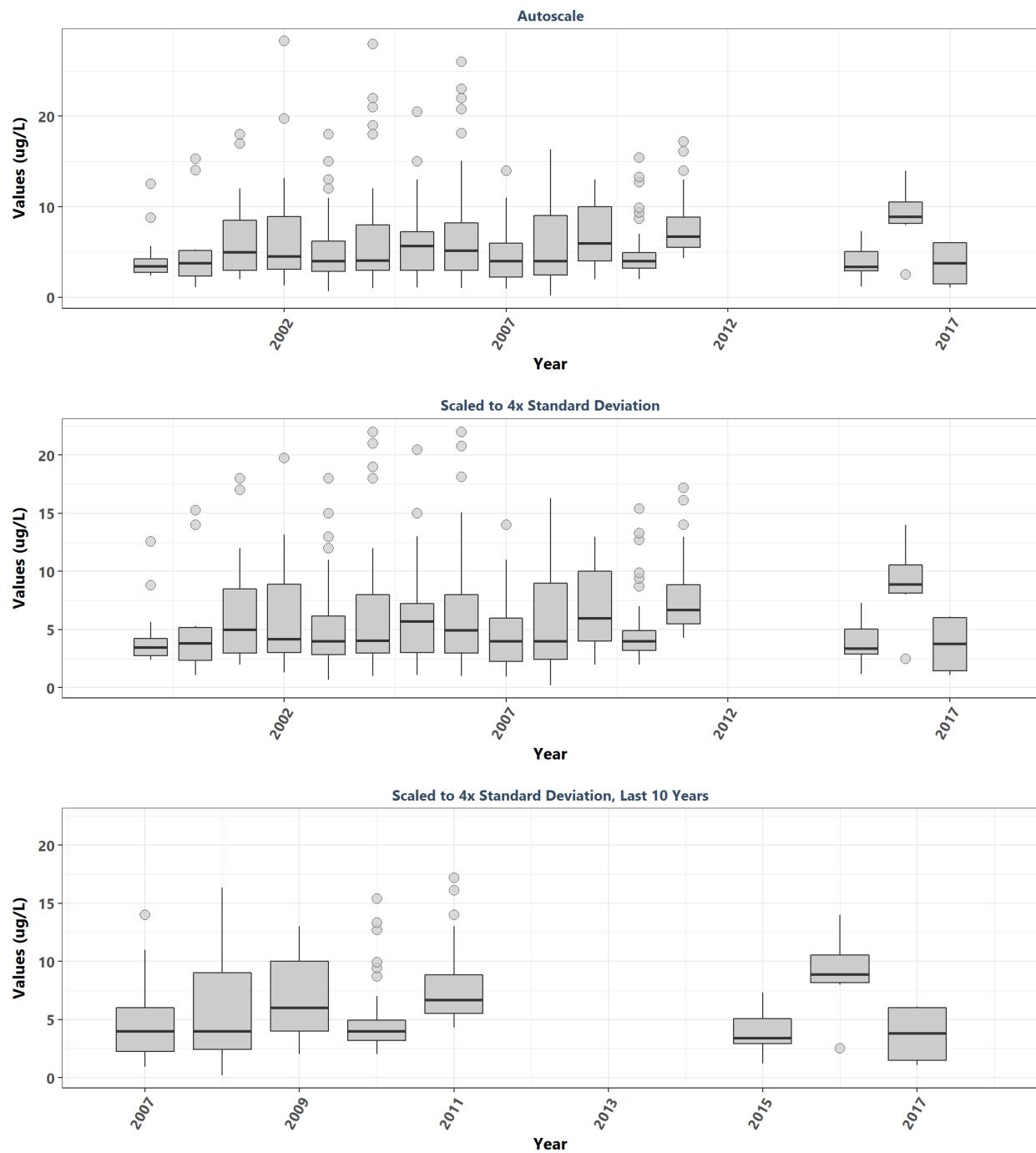
**Rocky Bayou State Park Aquatic Preserve**  
By Year & Month



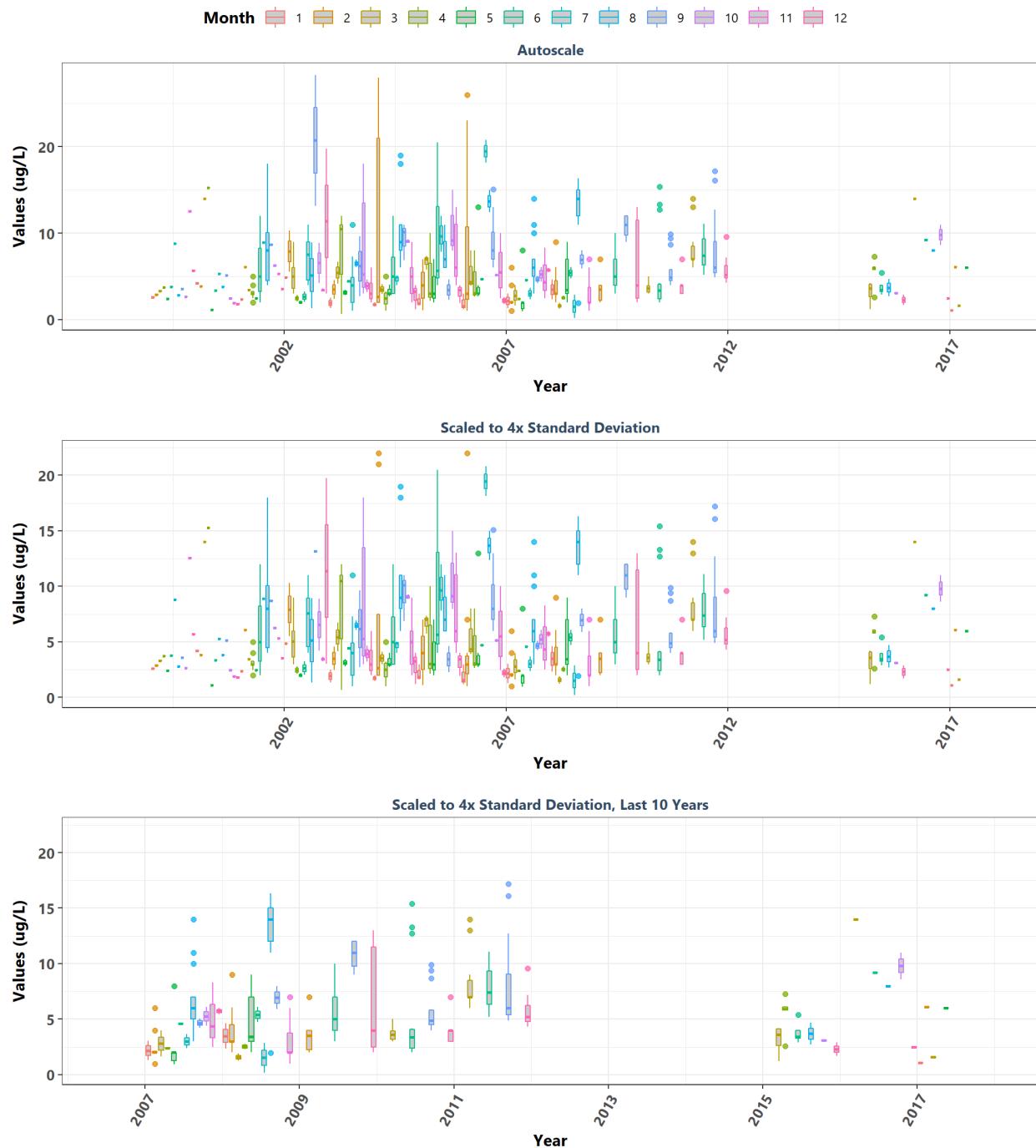
**Rocky Bayou State Park Aquatic Preserve**  
By Month



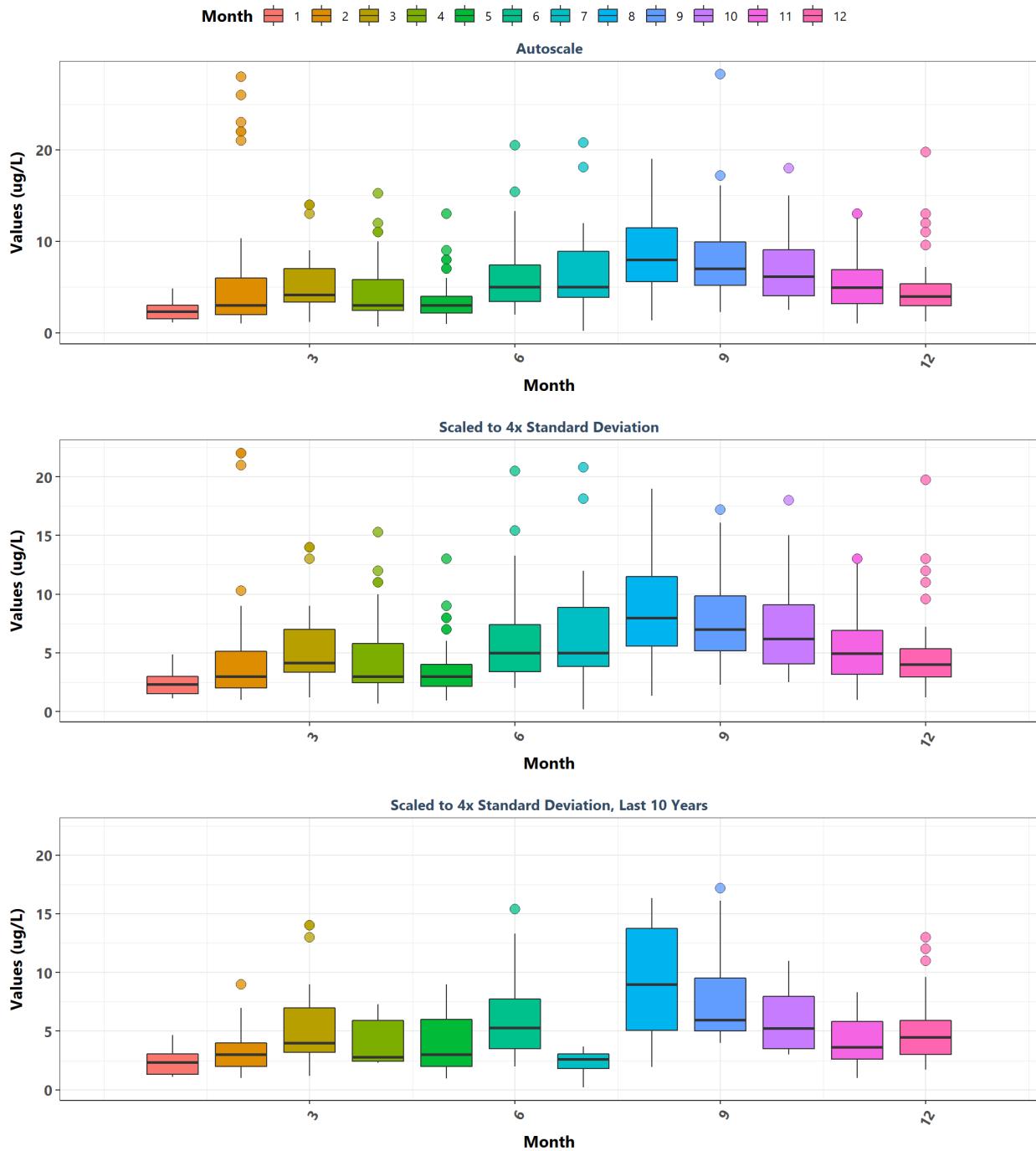
**Rookery Bay Aquatic Preserve**  
By Year



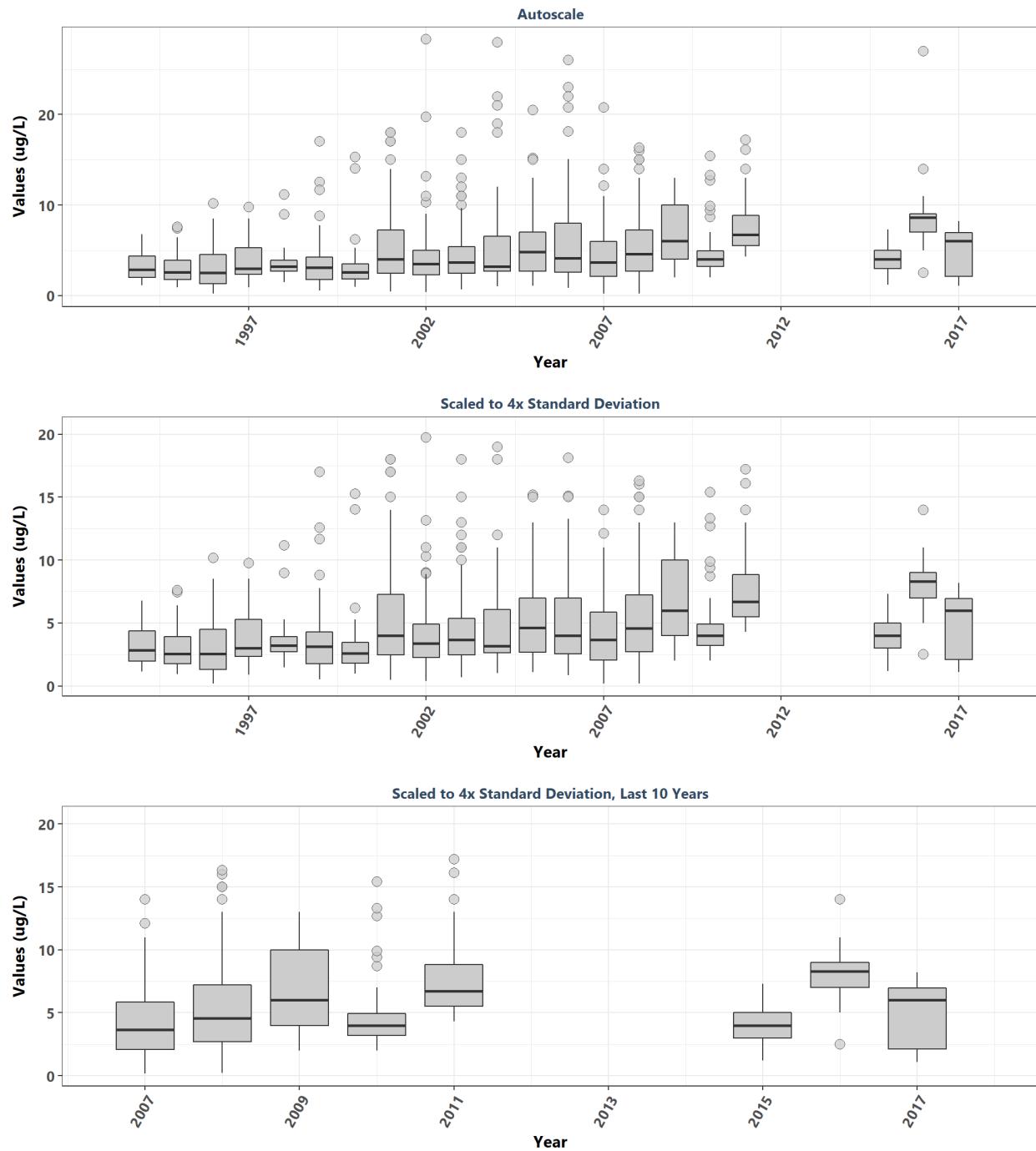
**Rookery Bay Aquatic Preserve**  
By Year & Month



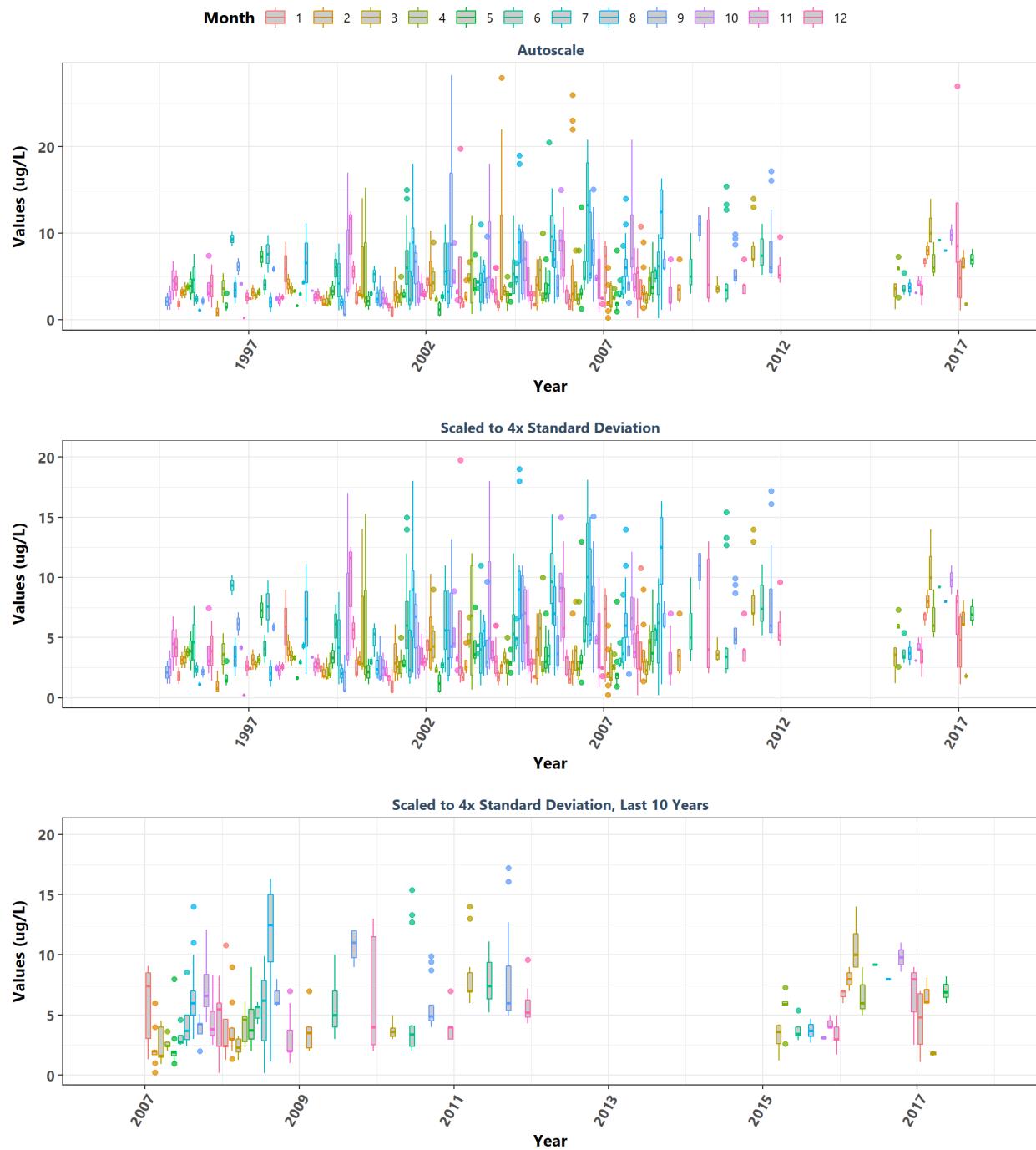
**Rookery Bay Aquatic Preserve**  
By Month



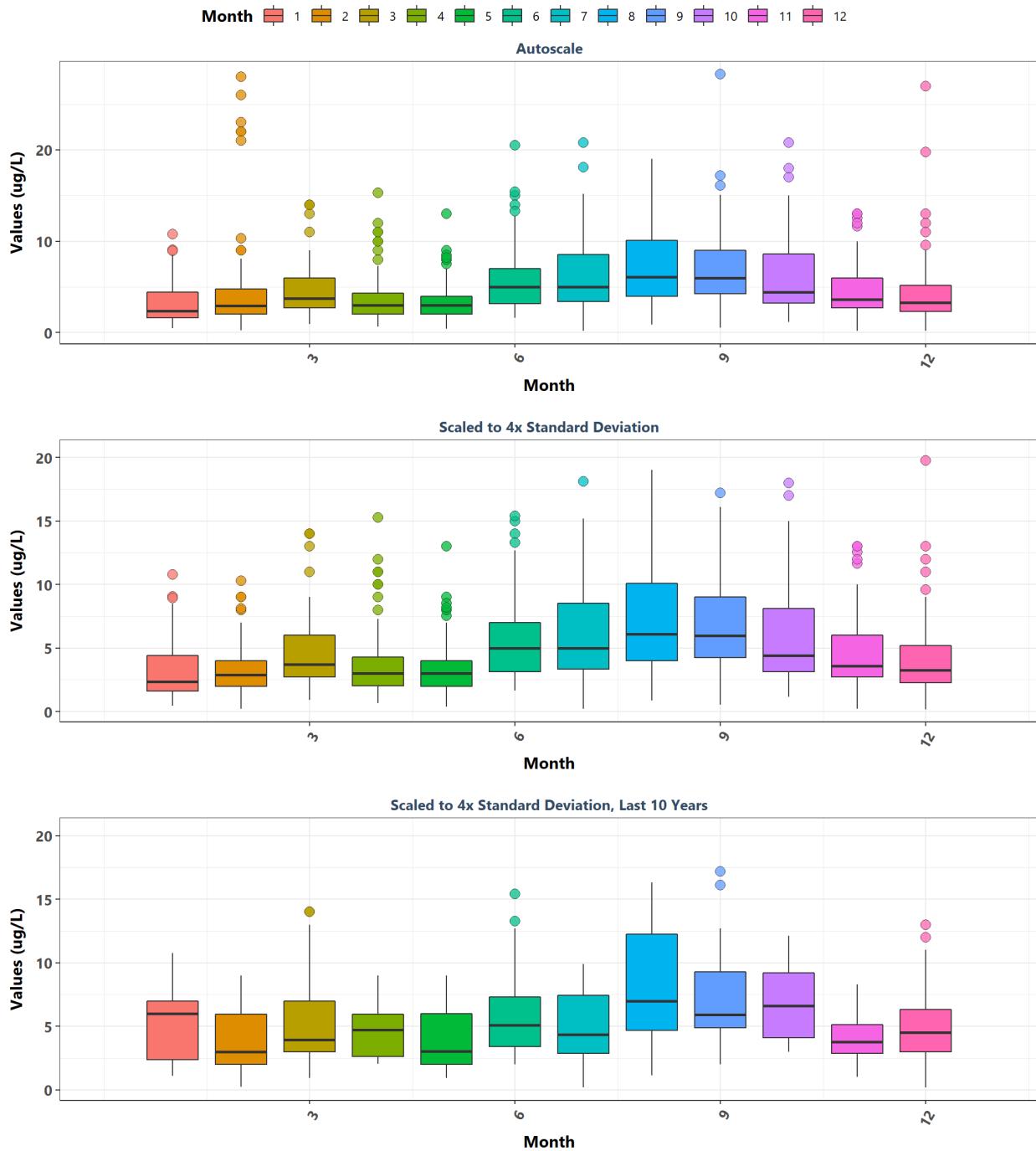
**Rookery Bay National Estuarine Research Reserve**  
By Year



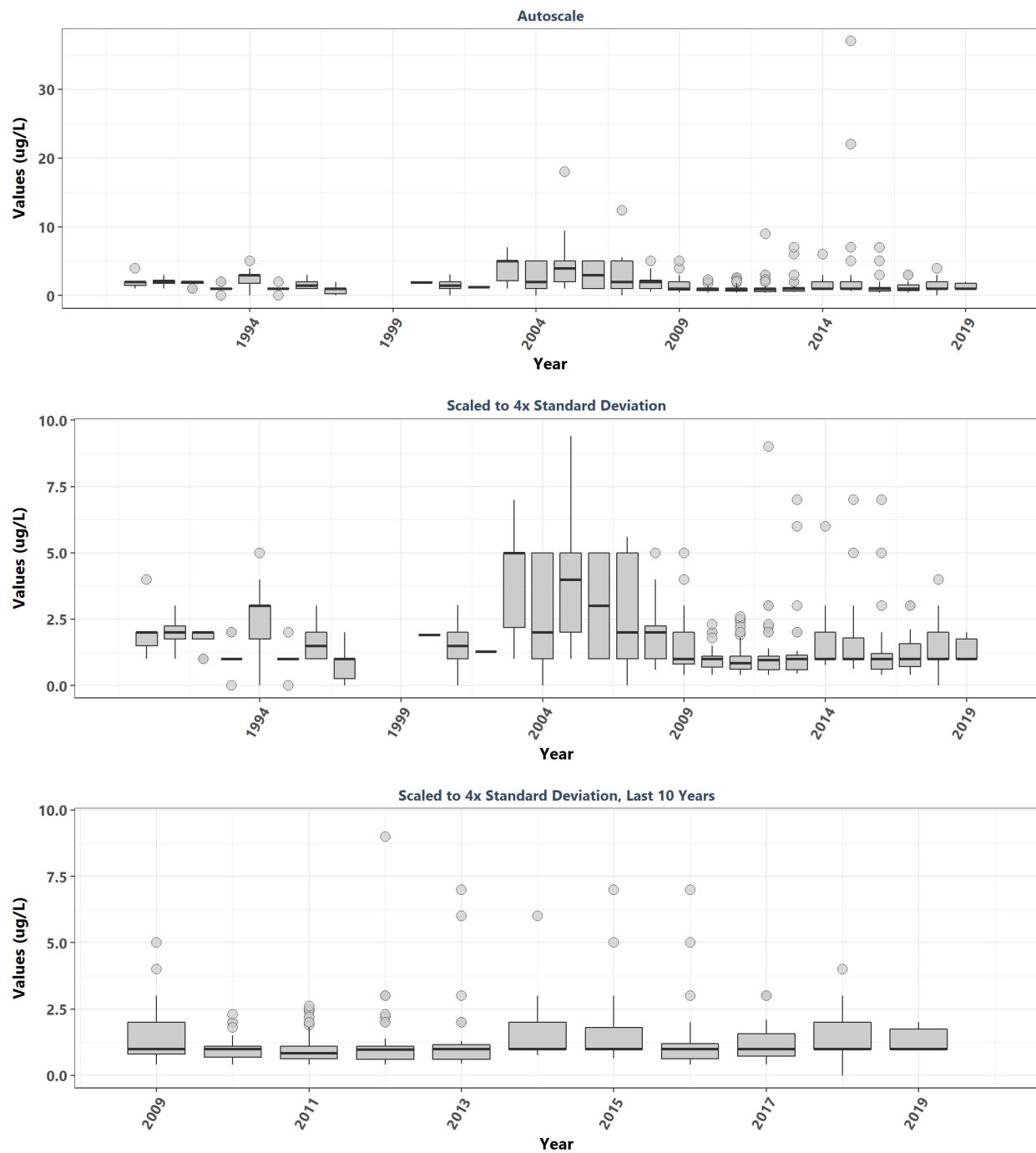
**Rookery Bay National Estuarine Research Reserve**  
By Year & Month



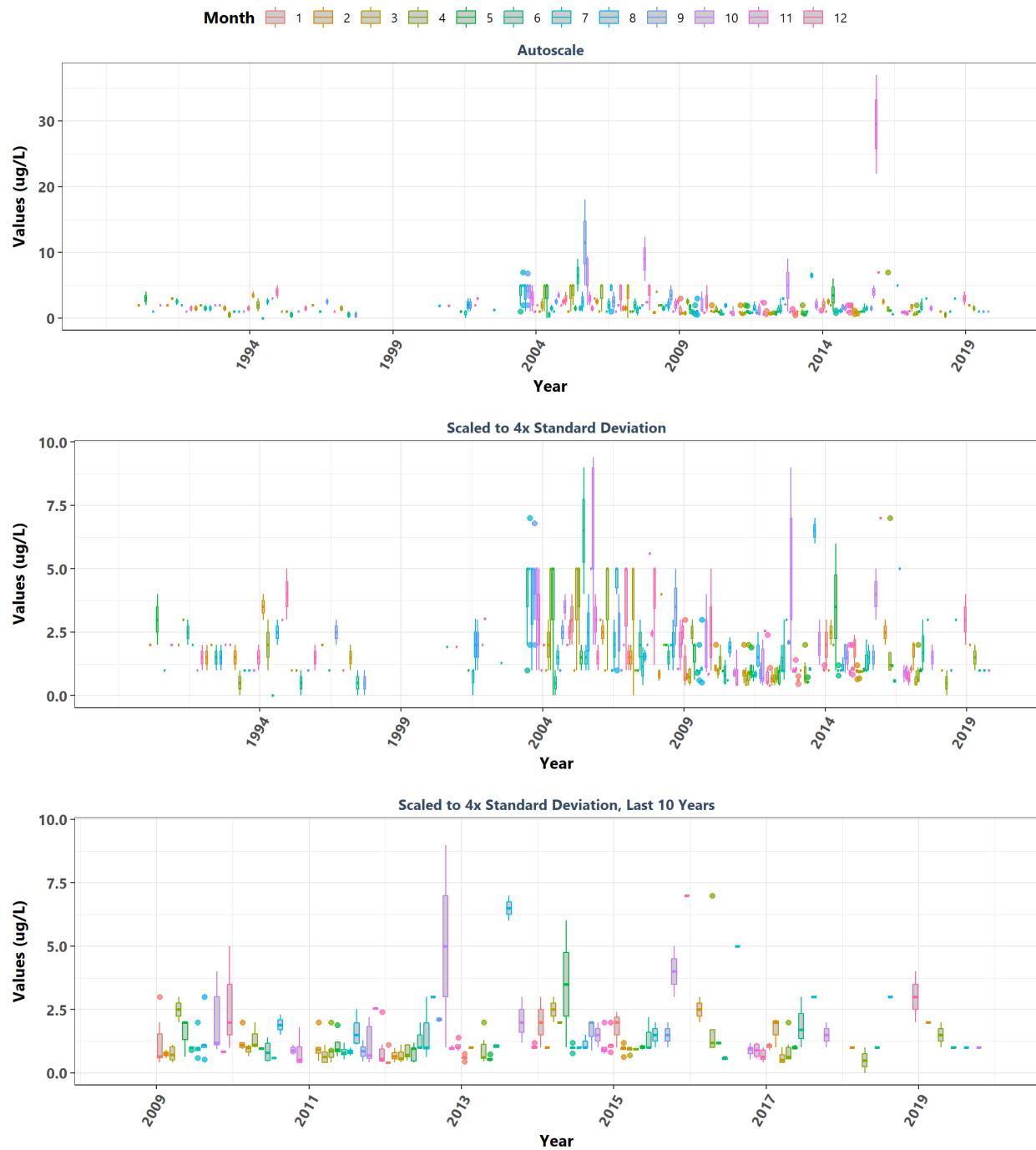
**Rookery Bay National Estuarine Research Reserve**  
By Month



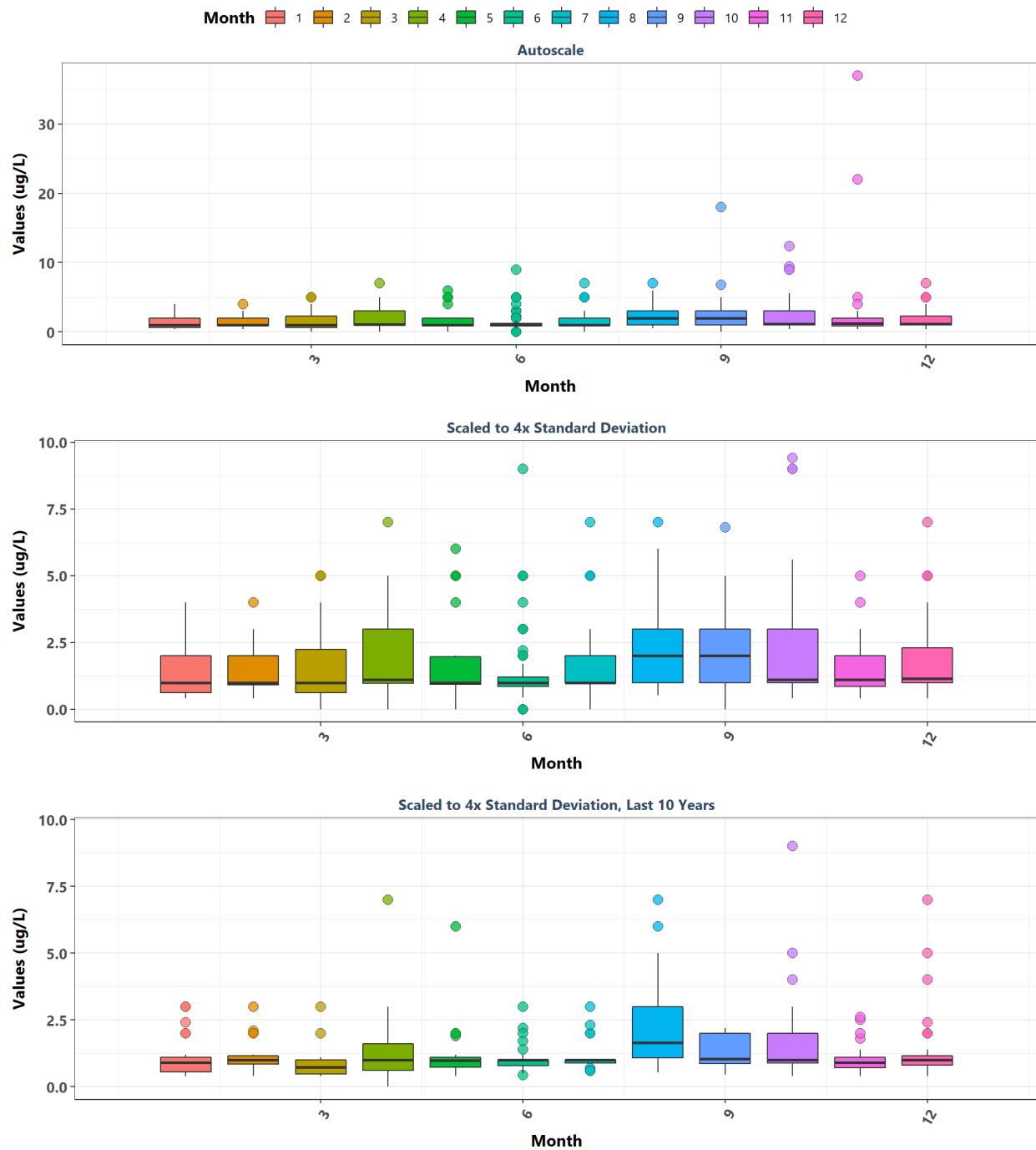
**St. Andrews State Park Aquatic Preserve**  
By Year



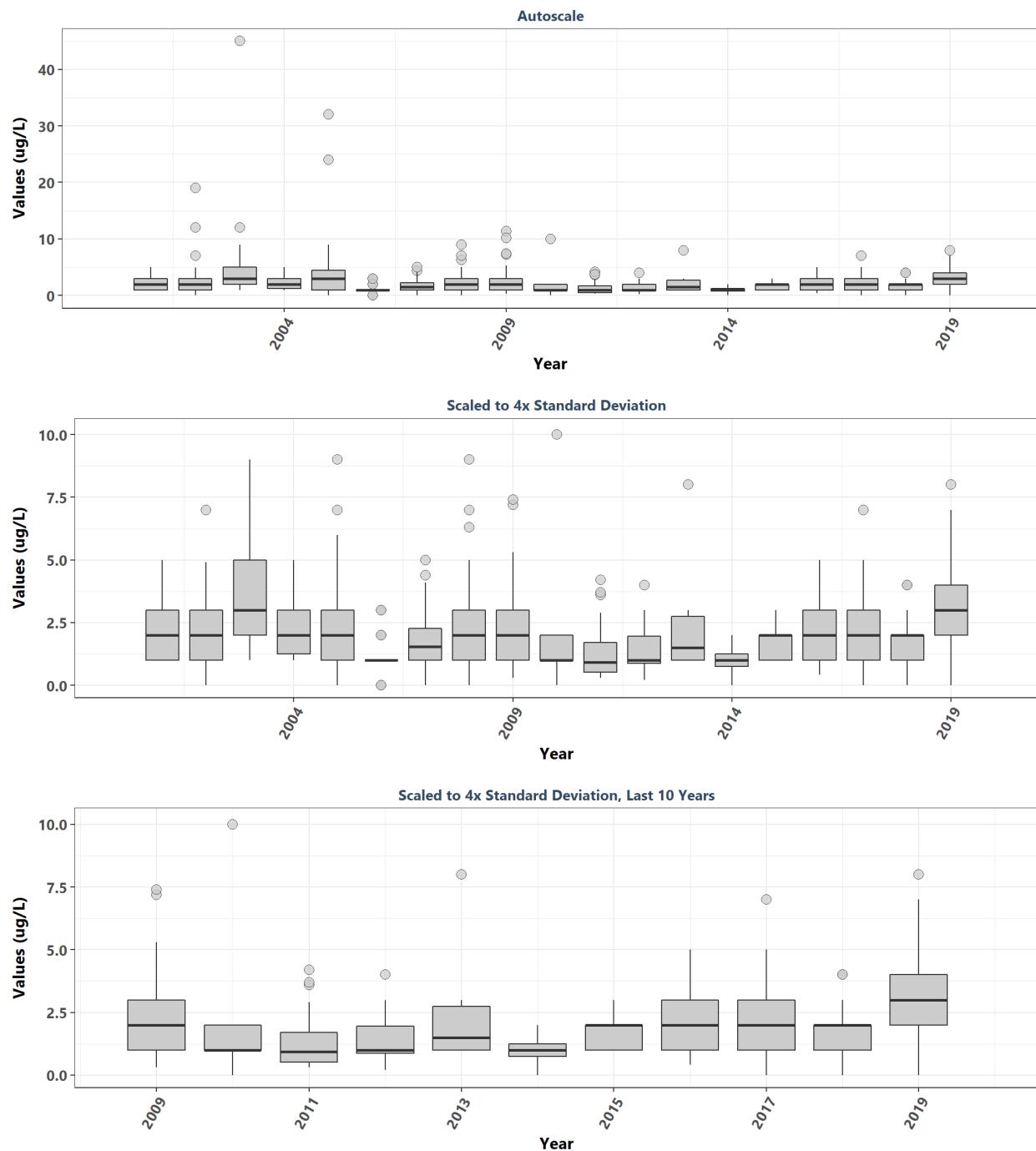
**St. Andrews State Park Aquatic Preserve**  
By Year & Month



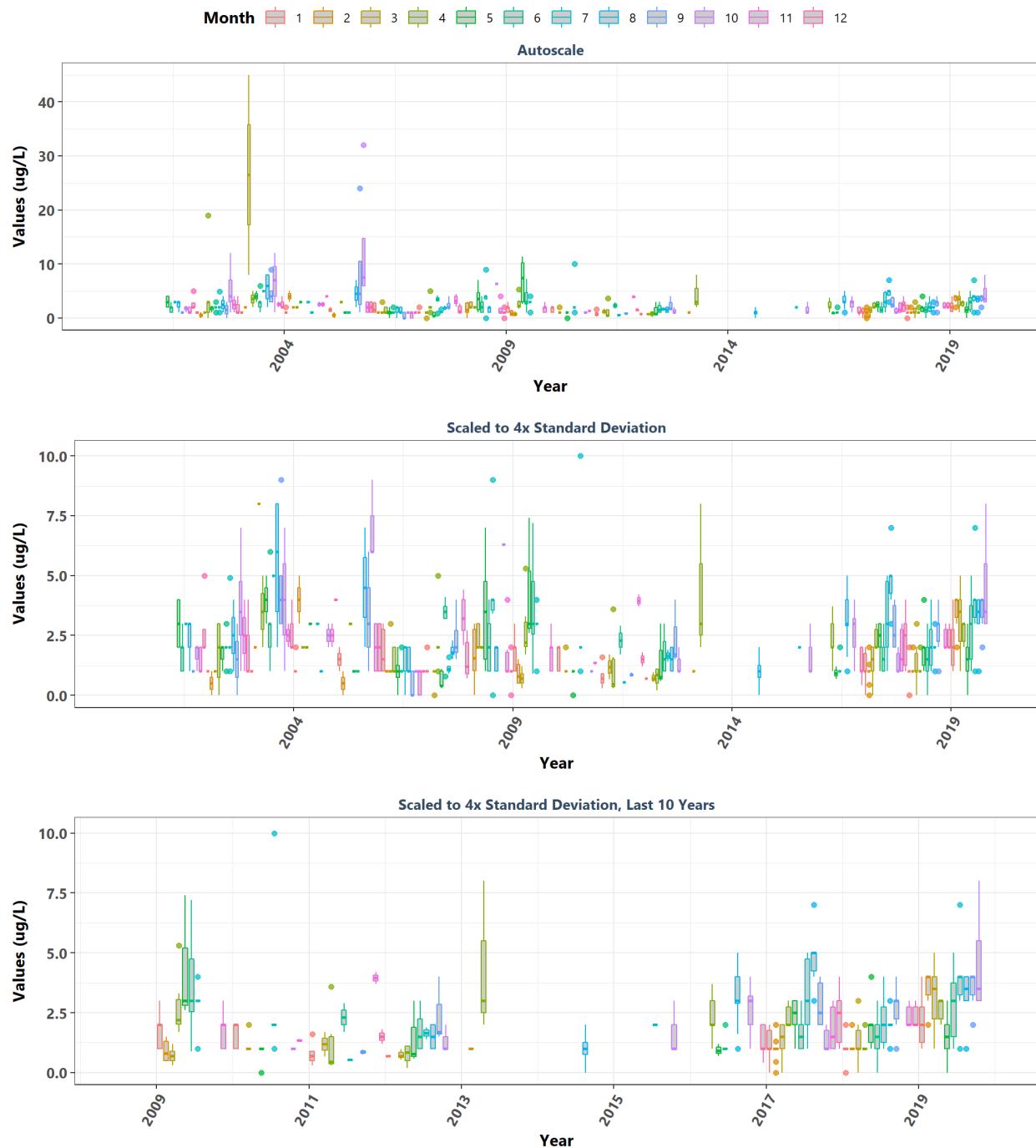
**St. Andrews State Park Aquatic Preserve**  
By Month



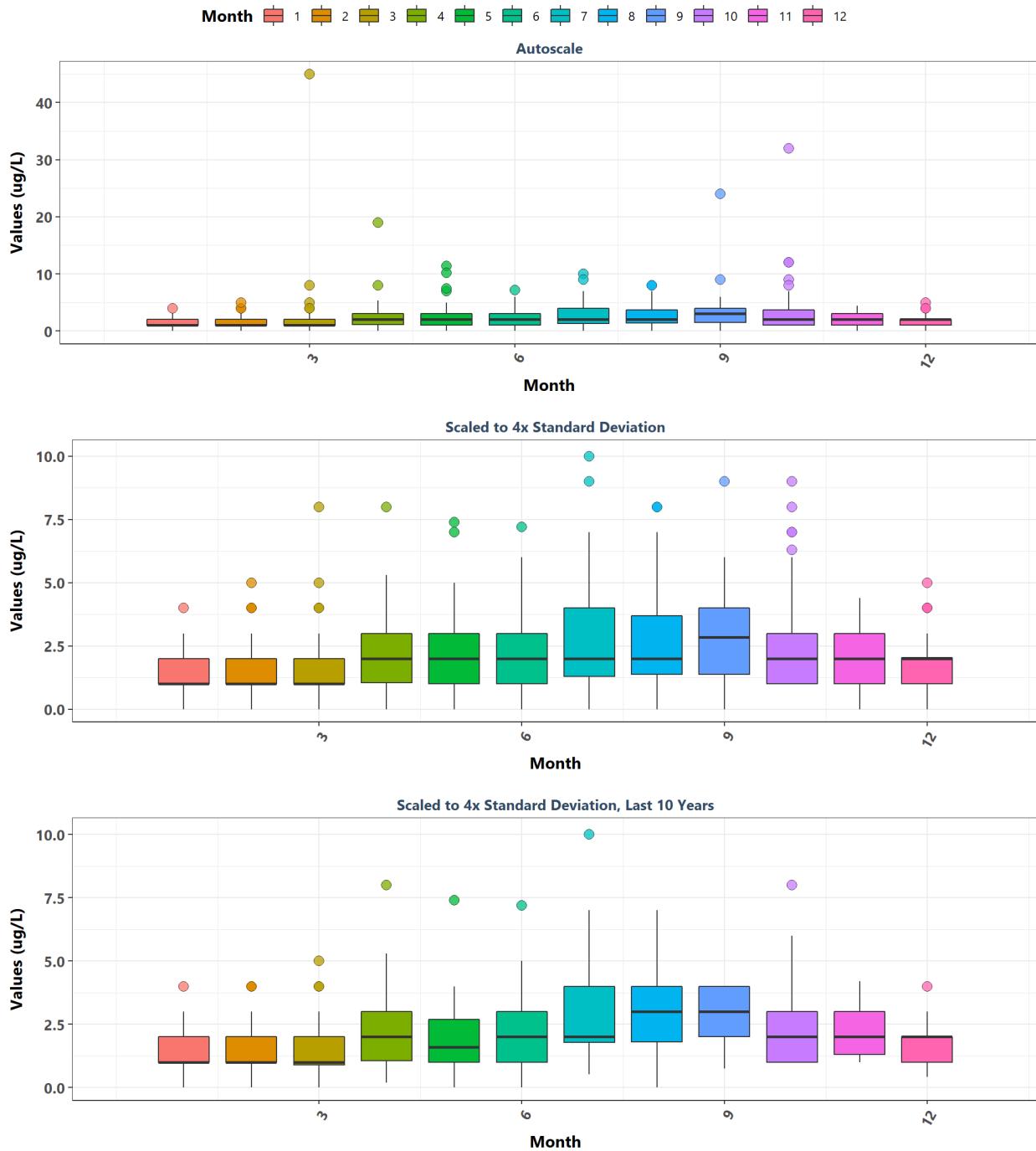
**St. Joseph Bay Aquatic Preserve**  
By Year



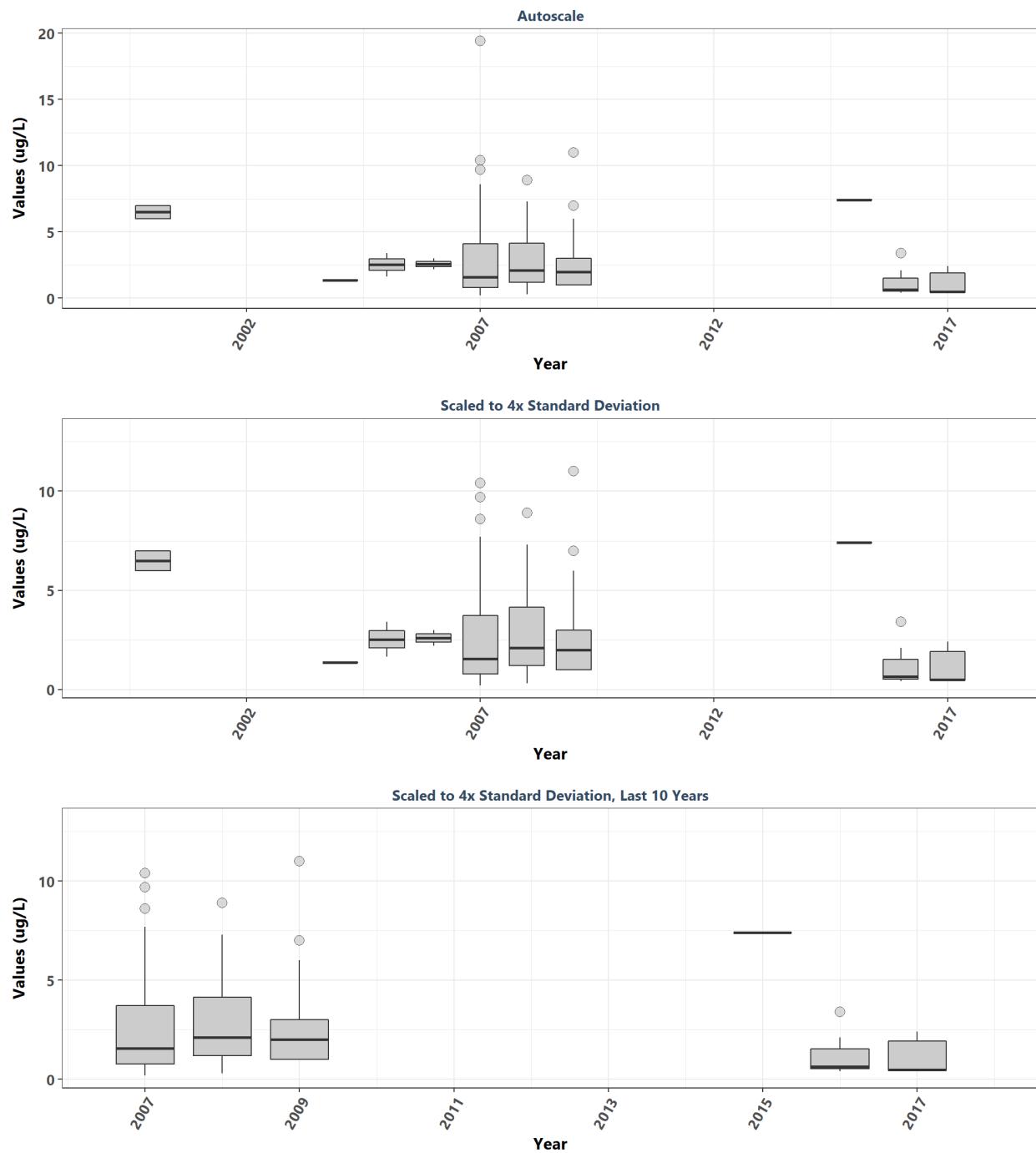
**St. Joseph Bay Aquatic Preserve**  
By Year & Month



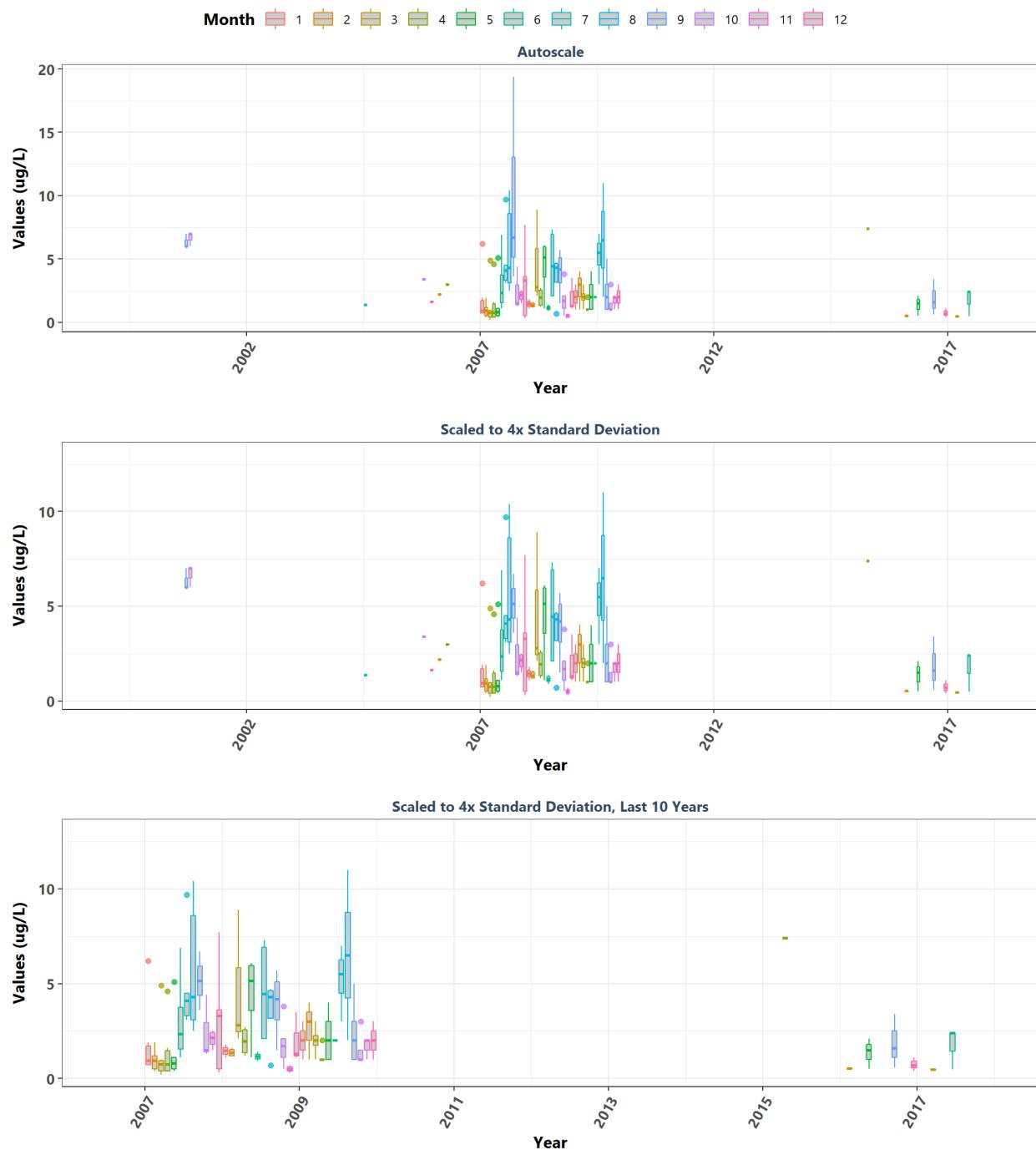
**St. Joseph Bay Aquatic Preserve**  
By Month



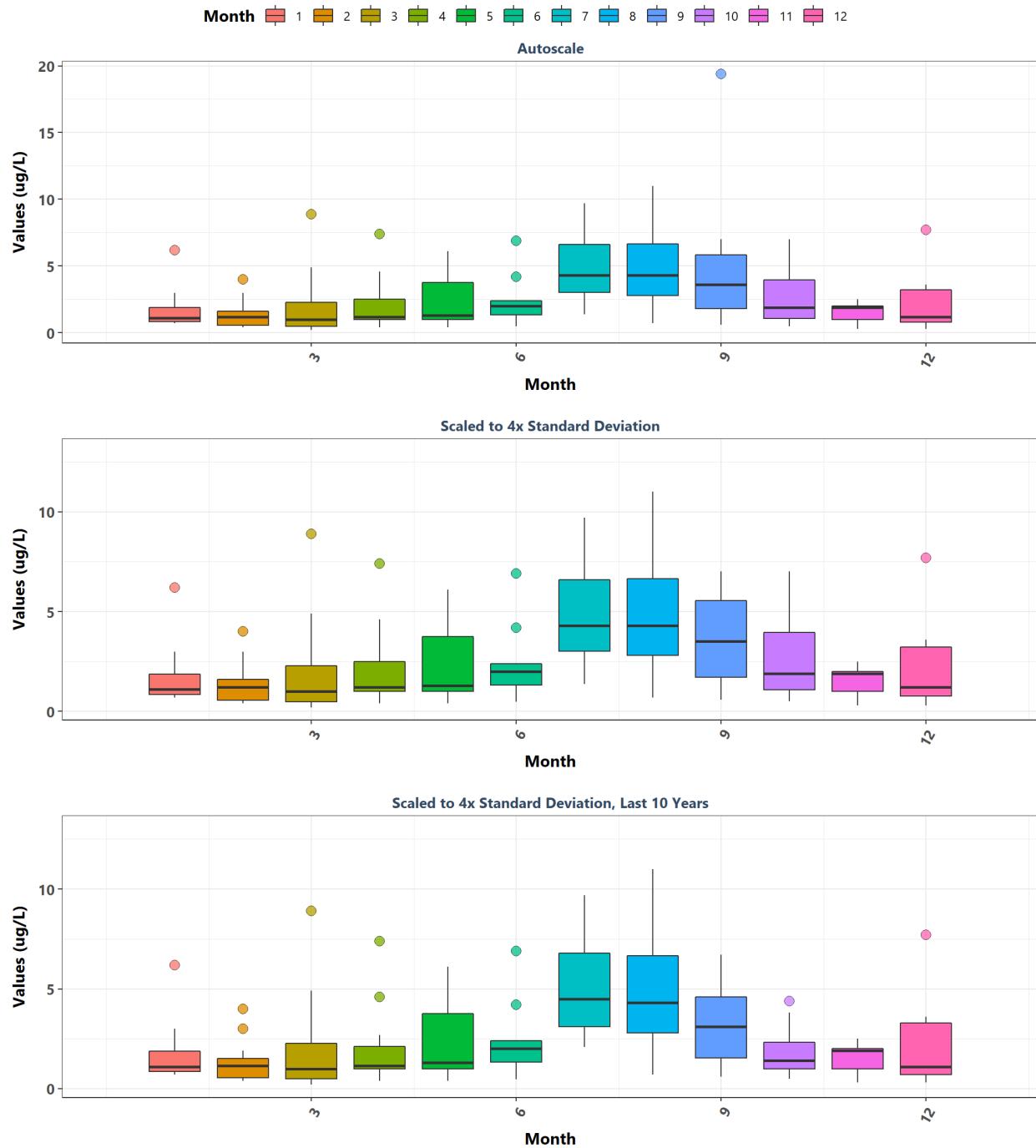
**St. Martins Marsh Aquatic Preserve**  
By Year



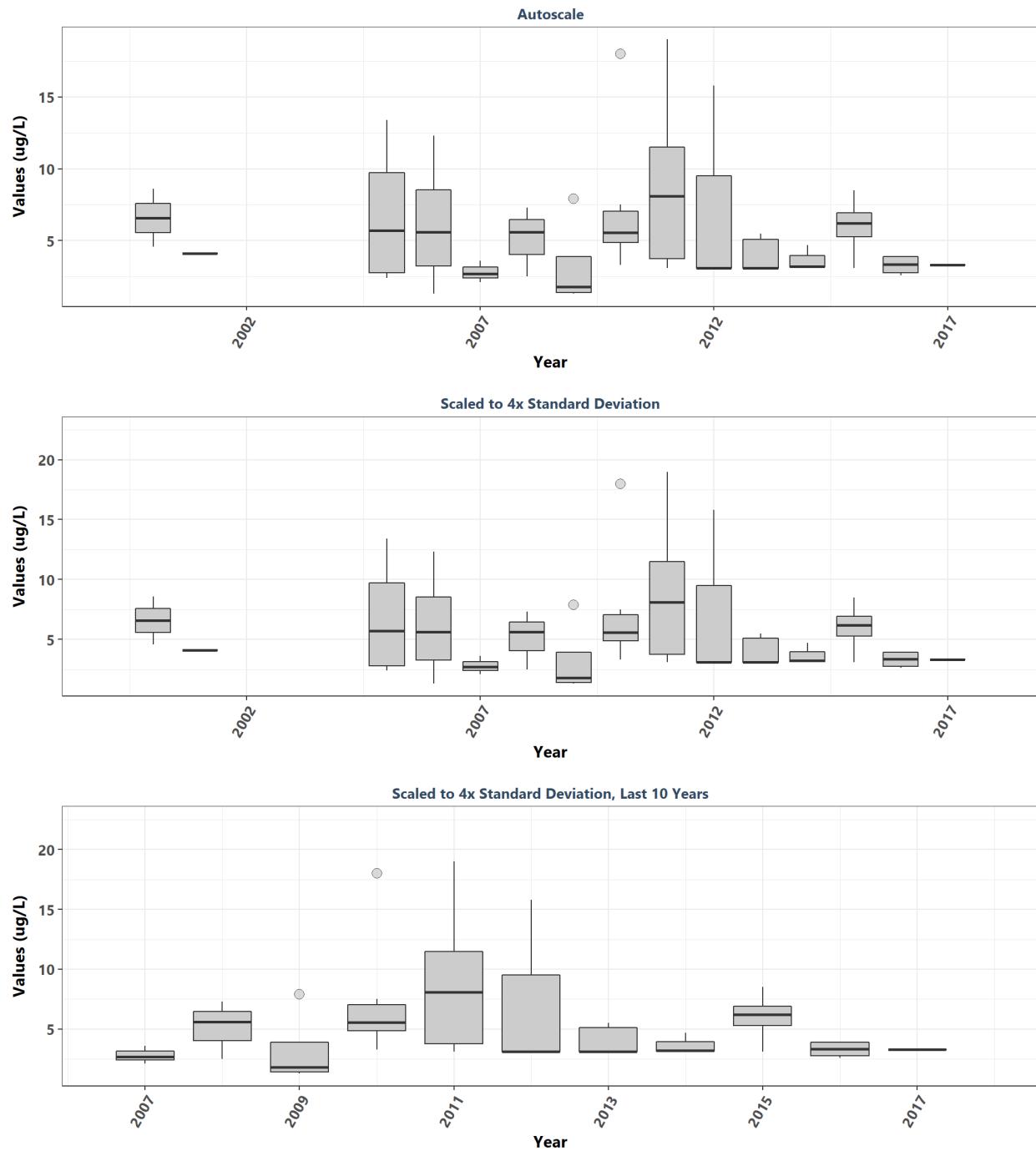
**St. Martins Marsh Aquatic Preserve**  
By Year & Month



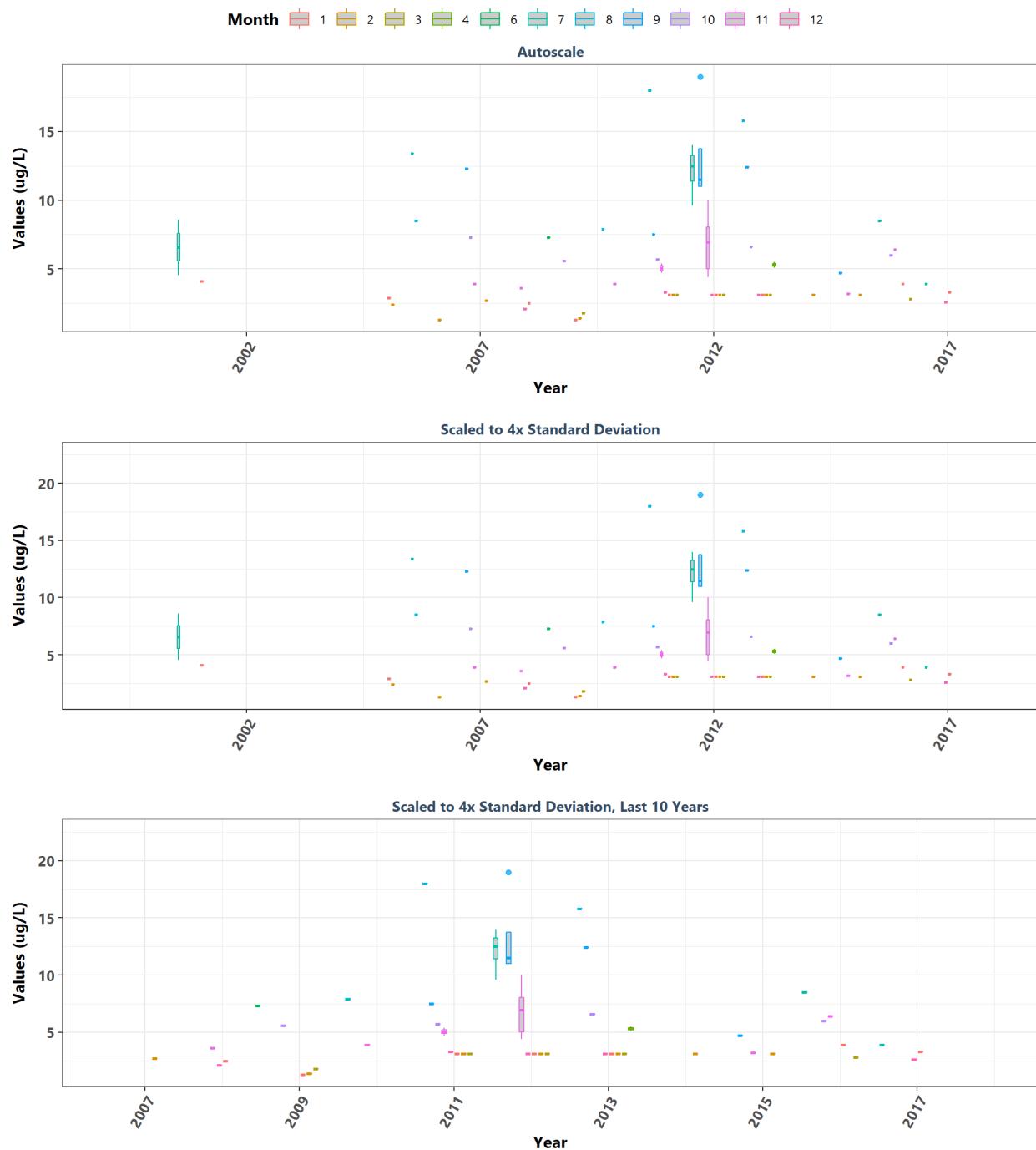
**St. Martins Marsh Aquatic Preserve**  
By Month



**Terra Ceia Aquatic Preserve**  
By Year



**Terra Ceia Aquatic Preserve**  
By Year & Month



**Terra Ceia Aquatic Preserve**  
By Month

