

Synoptic_CB_Nutrients_2023_AnalysisTemplate

Month of Data Being Analyzed

2025-06-23

Contents

0.1	Import Data & Clean	2
0.2	Assessing standard Curves	2
0.3	Dilution Corrections - ensure the latest dilution is kept	7
0.4	Analyze the Check Standards	8
0.5	Analyze Blanks	9
0.6	Analyze Duplicates	10
0.7	Spikes	11
0.8	Matrix Effects	12
0.9	Unit Converted Data Column Added (mg/L to uM)	12
0.10	Sample Flagging - Within range of standard curve	12
0.11	Pull out sample id information	12
0.12	Pulling Rhizon Samples	12
0.13	Check to see if samples run match metadata & merge info	13
0.14	Visualize Data	15
0.15	Export Processed Data	17

##Run Information

```
cat("Run Information: Input by User") #lets you know what section you're in
```

Run Information: Input by User

```
#set the run date & user name
run_date <- "20240114"
sample_year <- 2023
sample_month <- 06
user <- "Stephanie Wilson"

#identify the files you want to read in
#read in as a list to accomodate ultiple runs in a month
NOx_files <- c("Raw Data/SEAL_COMPASS_Synoptic_NOx_June2023_1.csv",
               "Raw Data/SEAL_COMPASS_Synoptic_NOx_June2023_2.csv",
```

```

        "Raw Data/SEAL_COMPASS_Synoptic_NOx_June2023_3.csv")
NH3_P04_files <- c("Raw Data/SEAL_COMPASS_Synoptic_NH3_P04_June2023_1.csv",
                  "Raw Data/SEAL_COMPASS_Synoptic_NH3_P04_June2023_2.csv",
                  "Raw Data/SEAL_COMPASS_Synoptic_NH3_P04_June2023_3.csv")

# Define the file path for QAQC log file - NO Need to change just check year
file_path <- "Raw Data/SEAL_COMPASS_Synoptic_QAQC_Log_2023.csv"
final_path <- "Processed Data/COMPASS_Synoptic_Nutrients_202306.csv"

#record any notes about the run or anything other info here:
run_notes <- "There are two sample names we suspect were input incorrectly,
they are listed below and have been checked against metadata. The metadata from Goodwin and Sweethall"

#duplicate sample names to be changed
#list the sample IDs that are messed up and create a list
#with run number as well so that we can change them below
wrong_names <- c("GCW_202304_TR_LysC_45cm", "GCW_202304_TR_LysA_20cm_8",
                "GWI_202304_UP_LysA_20cm", "GWI_202304_UP_LysA_20cm")
wrong_nums <- c(20, 16, 46, 44)
correct_names <- c("GCW_202304_TR_LysB_45cm", "GCW_202304_TR_LysA_20cm",
                  "GWI_202304_UP_LysA_10cm", "GWI_202304_UP_LysA_10cm")

#can't determine from metadata - for now unsure
remove_names <- c("GCW_202304_TR_LysA_20cm", "GCW_202304_TR_LysA_20cm",
                  "GCW_202304_TR_LysB_20cm_13", "GCW_202304_TR_LysB_20cm_13")
#couldn't tell which one this is from the metadata, no A_10cm which is what we thought
#marked on the sheet, need to check sample vials in freezer
#to see if we have a A_10cm from GCW_TR to be sure
remove_nums <- c(15, 13, 21, 19 )

#Set up file path for metadata
#downloaded metadata csv - downloaded from Google drive as csv for this year
#https://docs.google.com/spreadsheets/d/1HCANO_q6y17x0RUXVzID09hVal-RfwWc/edit?usp=sharing&ouid=10899
Raw_Metadata = "Raw Data/COMPASS_SynopticCB_PW_SampleLog_2023.csv"

```

```
##Setup
```

```
##Read in metadata and create similar sample IDs for matching to samples
```

0.1 Import Data & Clean

0.2 Assessing standard Curves

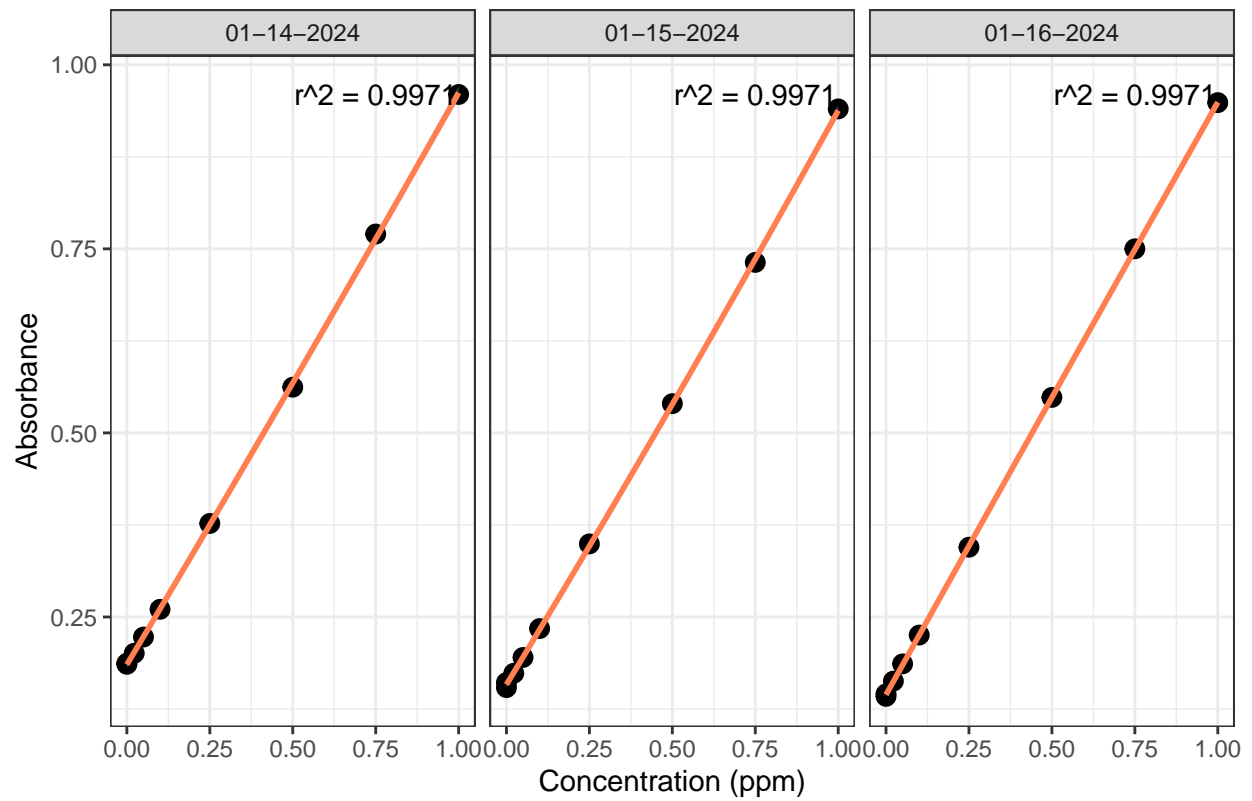
```
#Pull out standards data
```

```
## Assess Standard Curves
```

```
#Plot standards data
```

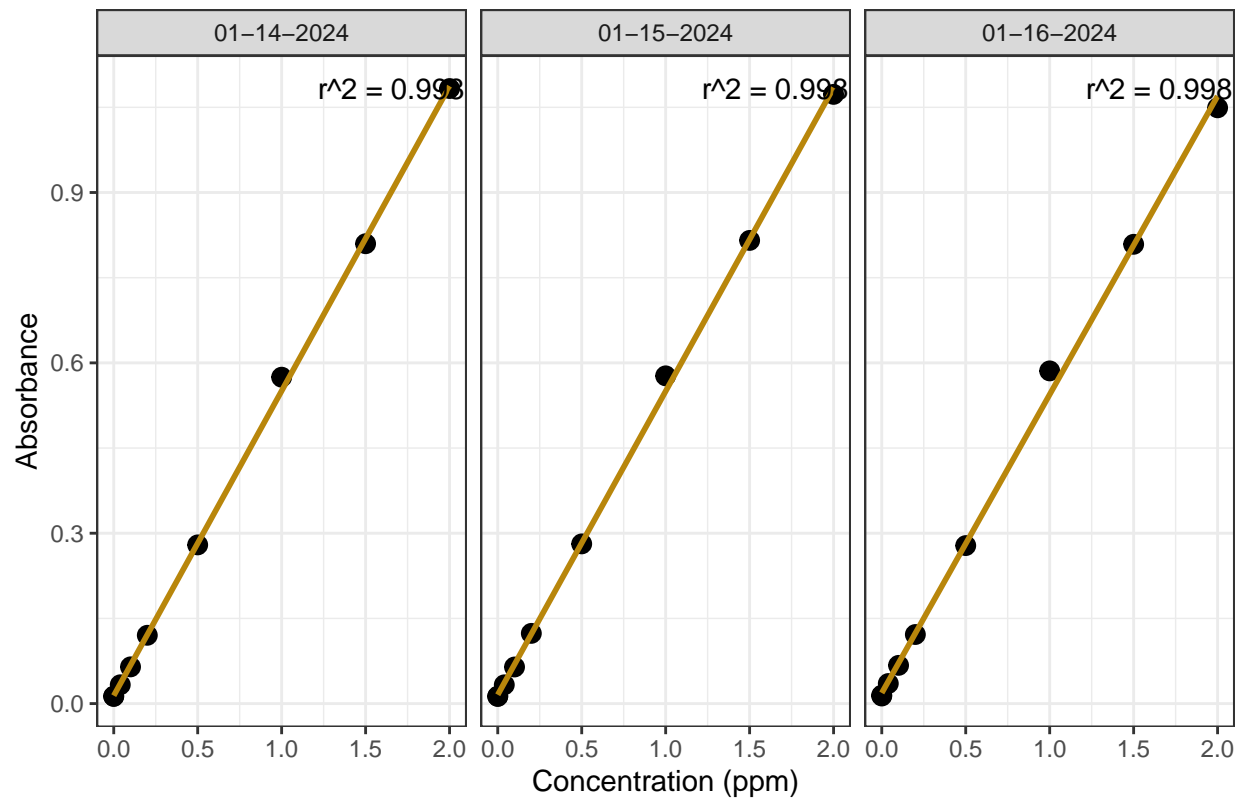
```
## Assess Standard Curves
```

NOx Standard Curve



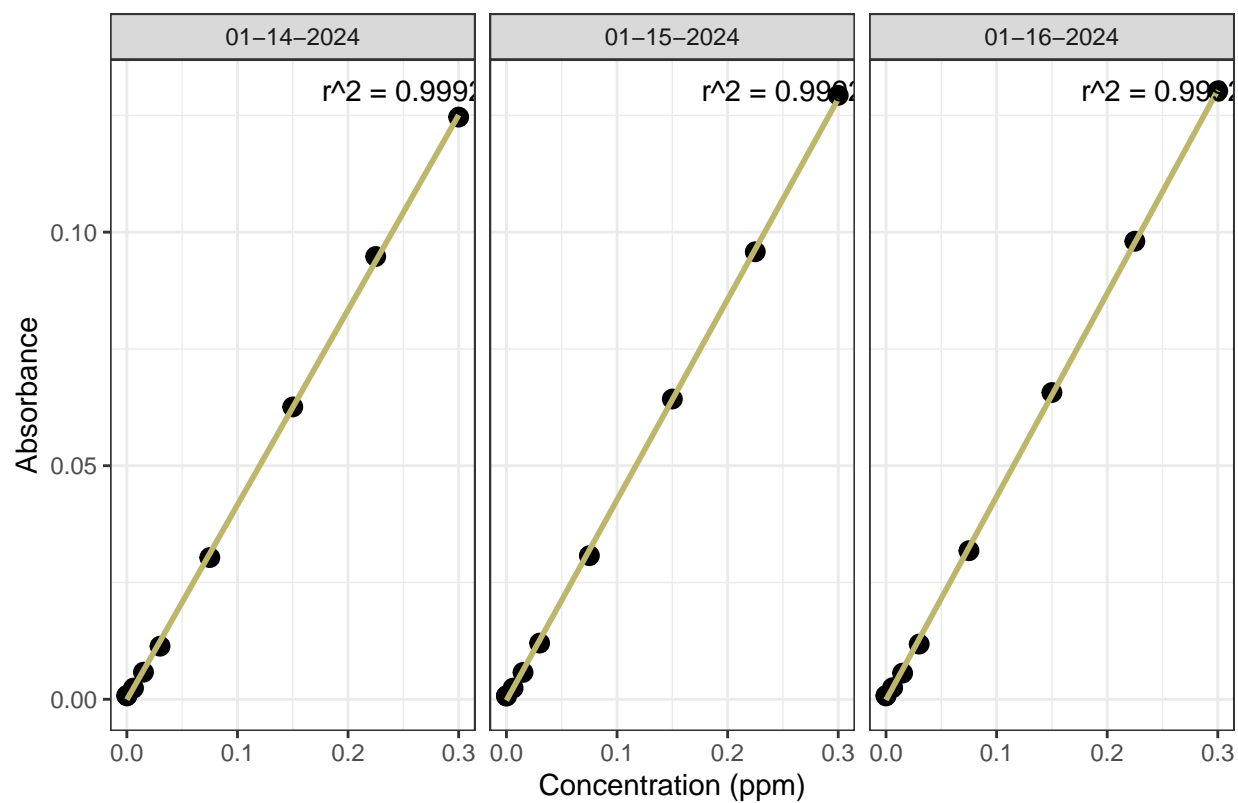
```
## 'geom_smooth()' using formula = 'y ~ x'
```

NH3 Standard Curve



```
## 'geom_smooth()' using formula = 'y ~ x'
```

PO4 Standard Curve



```
## [1] "NOx Curve r2 GOOD - PROCEED"
```

```
## [1] "NH3 Curve r2 GOOD - PROCEED"
```

```
## [1] "PO4 Curve r2 GOOD - PROCEED"
```

```
## [1] "QAQC log file exists and has been read into the code."
```

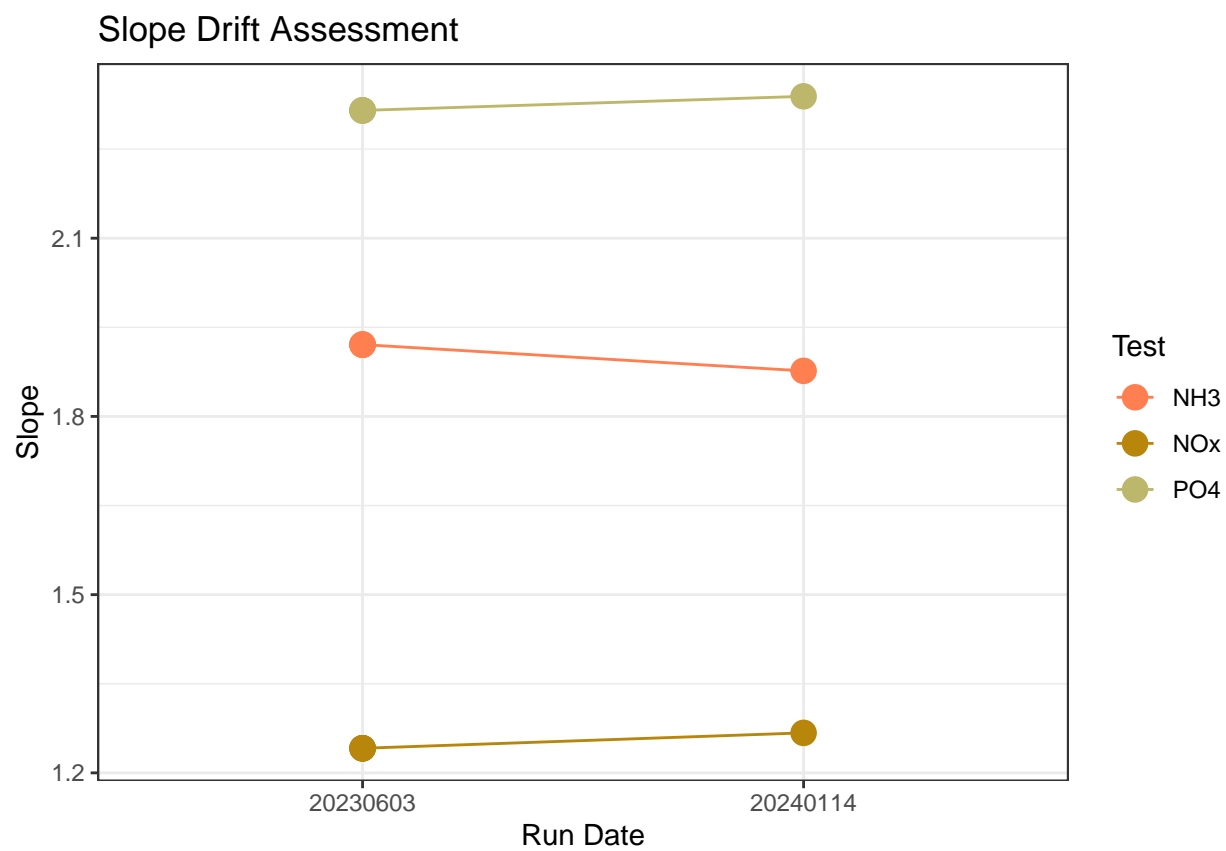


Table 1: Average Slope by Analyte

Test	avg_slope
NH3	1.912
NOx	1.247
PO4	2.320

0.3 Dilution Corrections - ensure the latest dilution is kept

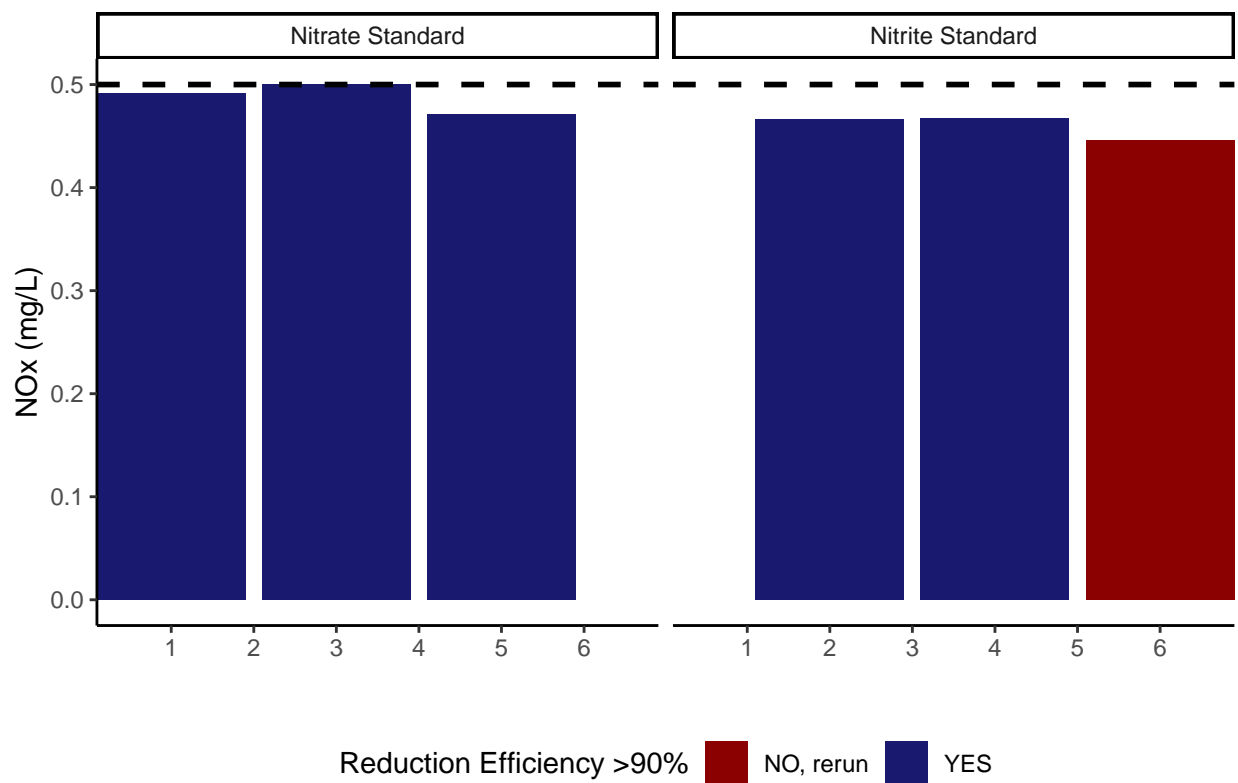
```
## Dilution Corrections
```

```
## Duplicated samples: MSM_202306_TR_LysB_20cm, GCW_202306_WC_LysA_45cm, SWH_202306_TR_LysB_45cm, SWH_202306_WC_LysA_45cm
```

```
## Dilution Present, Need to Correct
```

```
#Check NOx Reduction Efficiency
```

```
## Assess Reduction Efficiency
```



```
## [1] "Mean NOx Reduction Efficiency <95% - REASSESS"
```

```
## [1] 94.7757
```

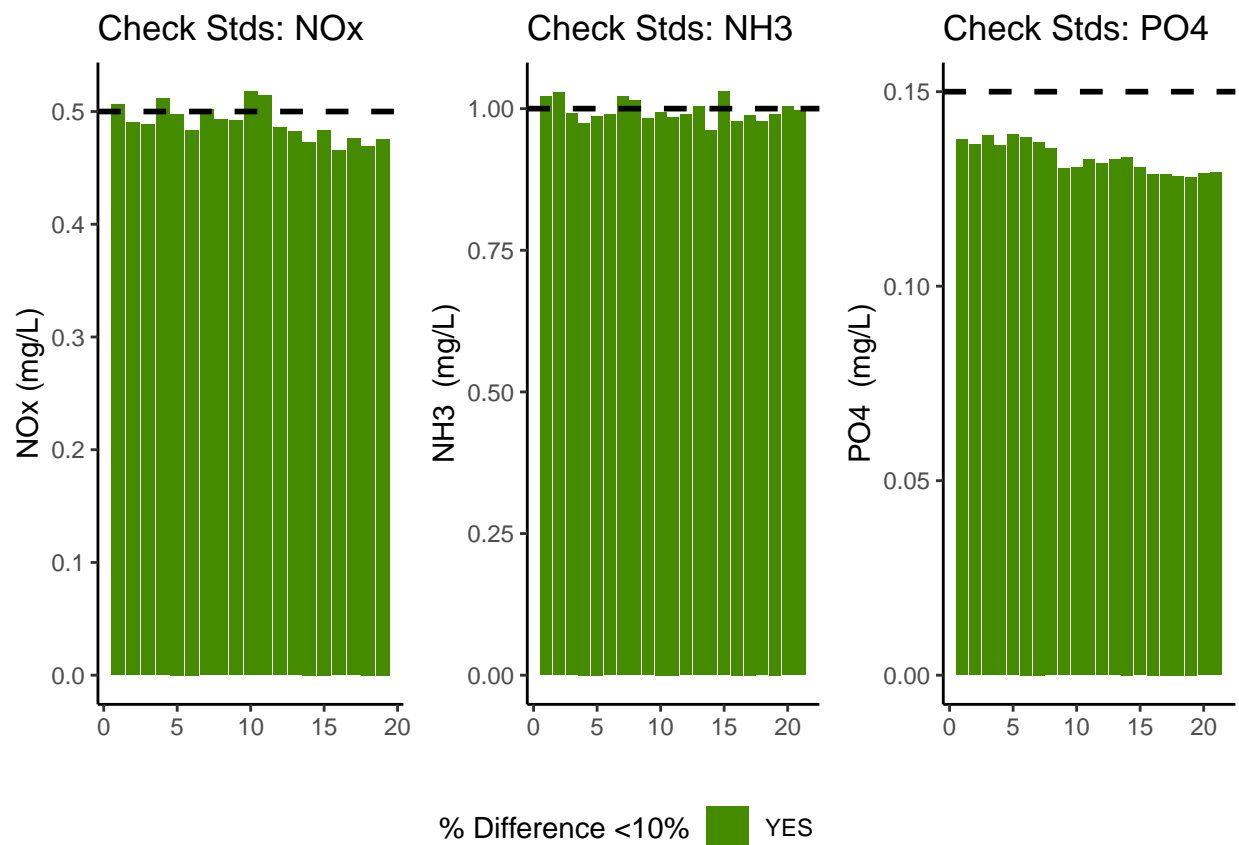
0.4 Analyze the Check Standards

```
## Analyze Check Standards
```

```
## [1] "NOx Check Standard RSD within Range - PROCEED"
```

```
## [1] "NH3 Check Standard RSD within Range - PROCEED"
```

```
## [1] "PO4 Check Standard RSD within Range - PROCEED"
```



```
## [1] ">60% of NOx Check Standards are within range of expected concentration - PROCEED"
```

```
## [1] ">60% of NH3 Check Standards are within range of expected concentration - PROCEED"
```

```
## [1] ">60% of PO4 Check Standards are within range of expected concentration - PROCEED"
```


0.5 Analyze Blanks

Assess Blanks

[1] ">60% of NOx Blank concentrations are lower than the lower 25% quartile of samples - PROCEED"

[1] ">60% of NH3 Blank concentrations are lower than the lower 25% quartile of samples - PROCEED"

[1] ">60% of PO4 Blank concentrations are lower than the lower 25% quartile of samples- PROCEED"

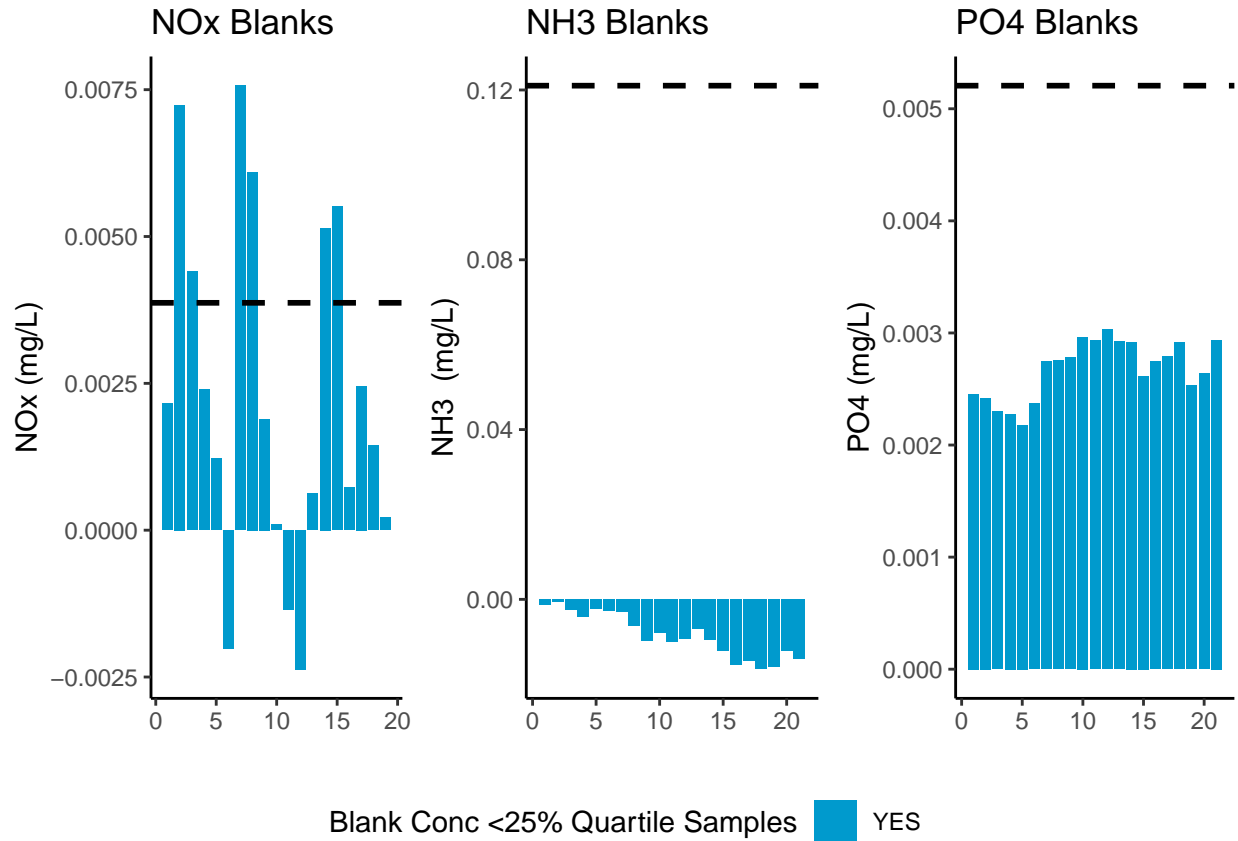


Table 2: Mean Concentration of Blanks

Test	Blank_Mean_Conc
NOx	0.0023
NH3	-0.0084
PO4	0.0027

0.6 Analyze Duplicates

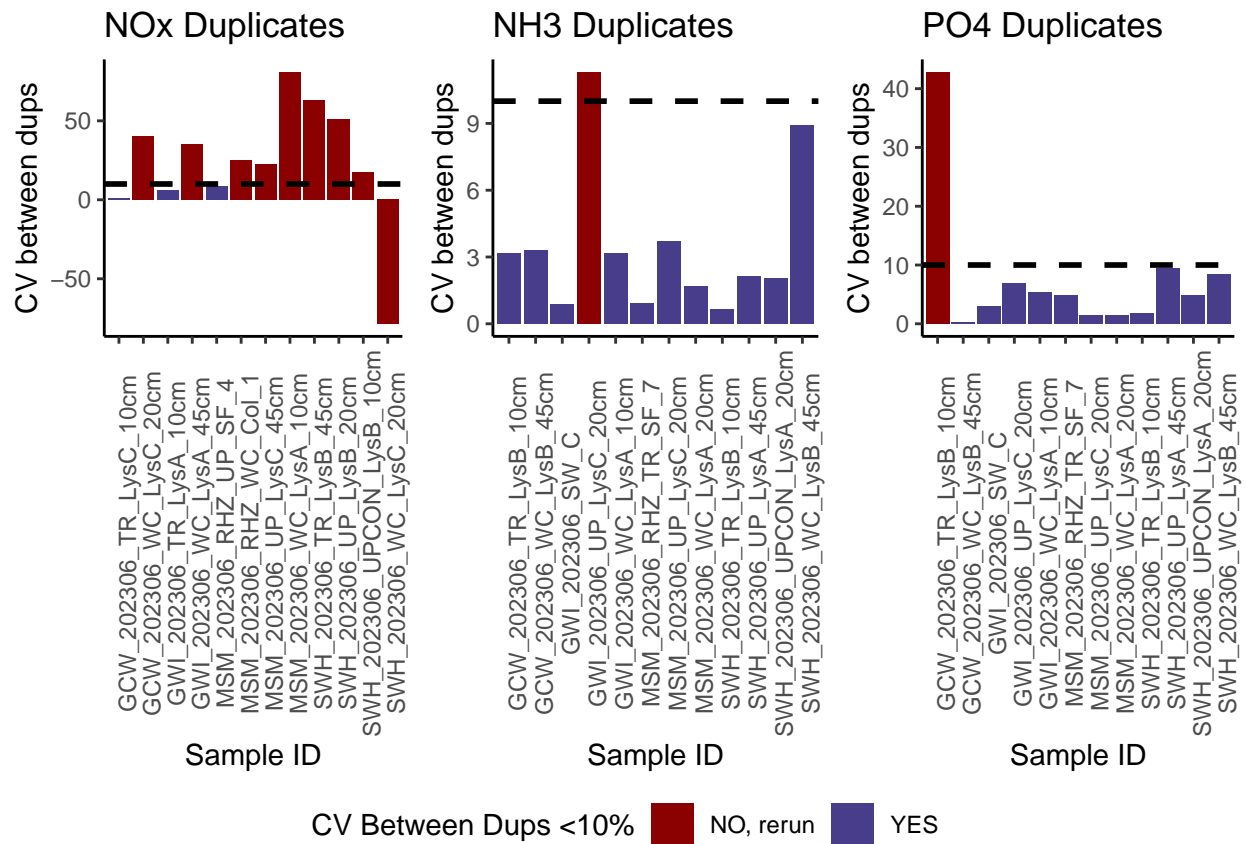
```
## Analyze Duplicates

## [1] "<60% of NOx Duplicates have a CV <10% - REASSESS"

## [1] ">60% of NH3 Duplicates have a CV <10% - PROCEED"

## [1] ">60% of PO4 Duplicates have a CV <10% - PROCEED"

## Warning: Using 'size' aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use 'linewidth' instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```

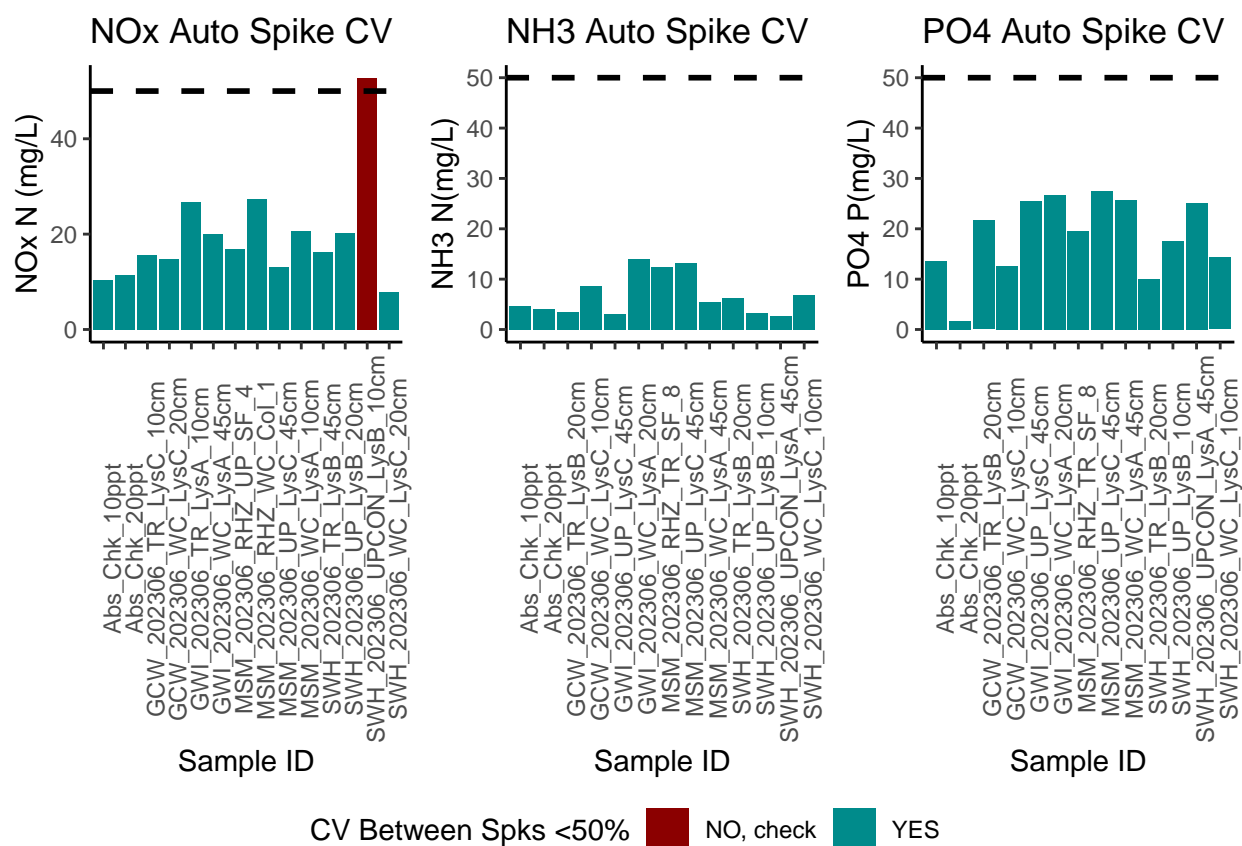


0.7 Spikes

[1] ">60% of Spikes have a CV <50% - PROCEED"

[1] ">60% of Spikes have a CV <50% - PROCEED"

[1] ">60% of Spikes have a CV <50% - PROCEED"



0.8 Matrix Effects

```
## [1] "NO NOx Matrix Effect, PROCEED"
```

```
## [1] "NO NH3 Matrix Effect, PROCEED"
```

```
## [1] "NO PO4 Matrix Effect, PROCEED"
```

0.9 Unit Converted Data Column Added (mg/L to uM)

0.10 Sample Flagging - Within range of standard curve

```
## Sample Flagging
```

0.11 Pull out sample id information

```
## Sample Processing
```

```
## Warning: Expected 5 pieces. Missing pieces filled with 'NA' in 36 rows [22, 23, 24, 49,
## 50, 51, 87, 88, 89, 115, 116, 117, 139, 140, 141, 181, 182, 183, 209, 210,
## ...].
```

```
## Warning: There was 1 warning in 'mutate()'.
## i In argument: 'Samp_Time = ym(Samp_Time)'.
## Caused by warning:
## ! 2 failed to parse.
```

0.12 Pulling Rhizon Samples

```
# Filter rhizon and peeper samples
df_rhizon <- df_all %>%
  filter(str_detect(Sample_Name, "RHZ"))
df_peep <- df_all %>%
  filter(str_detect(Sample_Name, "PPR"))

# Timestamp for backups
timestamp <- format(Sys.time(), "%Y-%m-%d_%H%M")

# Paths
folder_path <- file.path("Raw Data", "Rhizon+Peeper")
dir.create(folder_path, recursive = TRUE, showWarnings = FALSE)

rhizon_main <- file.path(folder_path, "rhizon_data.csv")
peeper_main <- file.path(folder_path, "peeper_data.csv")

rhizon_backup <- file.path(folder_path, paste0("rhizon_data_", timestamp, ".csv"))
peeper_backup <- file.path(folder_path, paste0("peeper_data_", timestamp, ".csv"))

# Write timestamped backups
```

```

write.csv(df_rhizon, rhizon_backup, row.names = FALSE)
write.csv(df_peep, peeper_backup, row.names = FALSE)

# Overwrite the main files with latest data
write.csv(df_rhizon, rhizon_main, row.names = FALSE)
write.csv(df_peep, peeper_main, row.names = FALSE)

## ^^ I think there is a cleaner way to write this out, but this should work for now ^^

```

0.13 Check to see if samples run match metadata & merge info

```
## Check Sample IDs with Metadata
```

```
## Some sample IDs are missing from metadata.
```

```

## [1] "MSM_202306_UP_LysA_45cm"      "MSM_202306_UP_LysC_20cm"
## [3] "MSM_202306_UP_LysC_45cm"      "MSM_202306_TR_LysA_45cm"
## [5] "SWH_202306_UPCON_LysA_10cm"    "SWH_202306_UPCON_LysA_20cm"
## [7] "SWH_202306_UPCON_LysA_45cm"    "SWH_202306_UPCON_LysB_10cm"
## [9] "SWH_202306_UPCON_LysB_20cm"    "SWH_202306_UPCON_LysB_45cm"
## [11] "SWH_202306_UPCON_LysC_10cm"    "SWH_202306_UPCON_LysC_20cm"
## [13] "SWH_202306_UPCON_LysC_45cm"    "SWH_202306_UP_LysA_10cm"
## [15] "SWH_202306_UP_LysA_20cm"      "SWH_202306_UP_LysA_45cm"
## [17] "SWH_202306_UP_LysB_10cm"      "SWH_202306_UP_LysB_20cm"
## [19] "SWH_202306_UP_LysB_45cm"      "SWH_202306_UP_LysC_10cm"
## [21] "SWH_202306_UP_LysC_20cm"      "SWH_202306_UP_LysC_45cm"
## [23] "SWH_202306_TR_LysA_10cm"      "SWH_202306_TR_LysA_20cm"
## [25] "SWH_202306_TR_LysA_45cm"      "SWH_202306_TR_LysB_10cm"
## [27] "SWH_202306_TR_LysB_20cm"      "SWH_202306_TR_LysB_45cm"
## [29] "SWH_202306_TR_LysC_10cm"      "SWH_202306_TR_LysC_20cm"
## [31] "SWH_202306_TR_LysC_45cm"      "SWH_202306_WC_LysA_10cm"
## [33] "SWH_202306_WC_LysA_45cm"      "SWH_202306_WC_LysB_10cm"
## [35] "SWH_202306_WC_LysB_20cm"      "SWH_202306_WC_LysB_45cm"
## [37] "SWH_202306_WC_LysC_10cm"      "SWH_202306_WC_LysC_20cm"
## [39] "SWH_202306_WC_LysC_45cm"      "SWH_202306_SW_A"
## [41] "SWH_202306_SW_B"              "SWH_202306_SW_C"
## [43] "GWI_202306_UP_LysA_10cm"      "GWI_202306_UP_LysA_20cm"
## [45] "GWI_202306_UP_LysA_45cm"      "GWI_202306_UP_LysB_10cm"
## [47] "GWI_202306_UP_LysB_20cm"      "GWI_202306_UP_LysB_45cm"
## [49] "GWI_202306_UP_LysC_10cm"      "GWI_202306_UP_LysC_20cm"
## [51] "GWI_202306_UP_LysC_45cm"      "GWI_202306_TR_LysA_10cm"
## [53] "GWI_202306_TR_LysA_20cm"      "GWI_202306_TR_LysB_10cm"
## [55] "GWI_202306_TR_LysB_20cm"      "GWI_202306_TR_LysB_45cm"
## [57] "GWI_202306_TR_LysC_10cm"      "GWI_202306_TR_LysC_20cm"
## [59] "GWI_202306_TR_LysC_45cm"      "GWI_202306_WC_LysA_10cm"
## [61] "GWI_202306_WC_LysA_20cm"      "GWI_202306_WC_LysA_45cm"
## [63] "GWI_202306_WC_LysB_10cm"      "GWI_202306_WC_LysB_20cm"
## [65] "GWI_202306_WC_LysC_10cm"      "GWI_202306_WC_LysC_20cm"
## [67] "GWI_202306_WC_LysB_45cm"      "GWI_202306_SW_A"
## [69] "GWI_202306_SW_B"              "GWI_202306_SW_C"
## [71] "MSM_202036_WC_LysB_10cm"      "GWI_202306_WC_LysC_45cm"

```

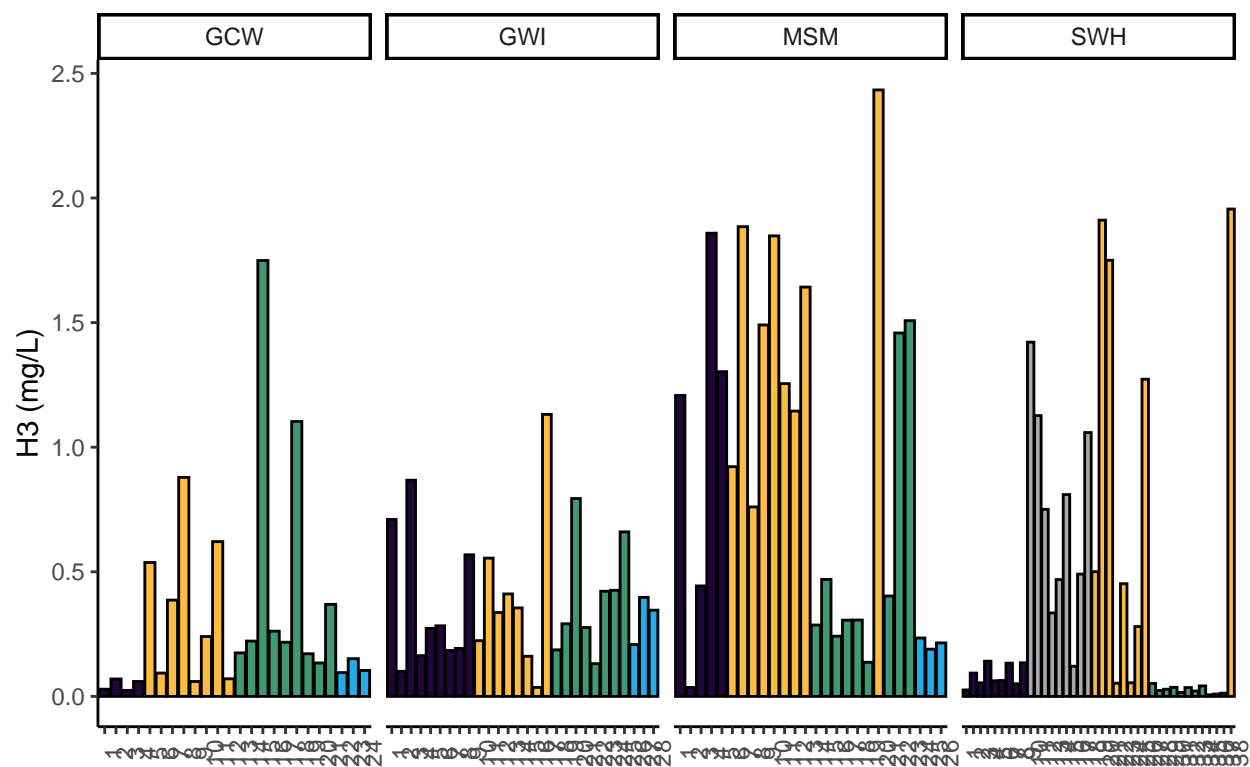
```
## Warning: Expected 5 pieces. Missing pieces filled with 'NA' in 36 rows [22, 23, 24, 49,  
## 50, 51, 87, 88, 89, 115, 116, 117, 139, 140, 141, 181, 182, 183, 209, 210,  
## ...].
```

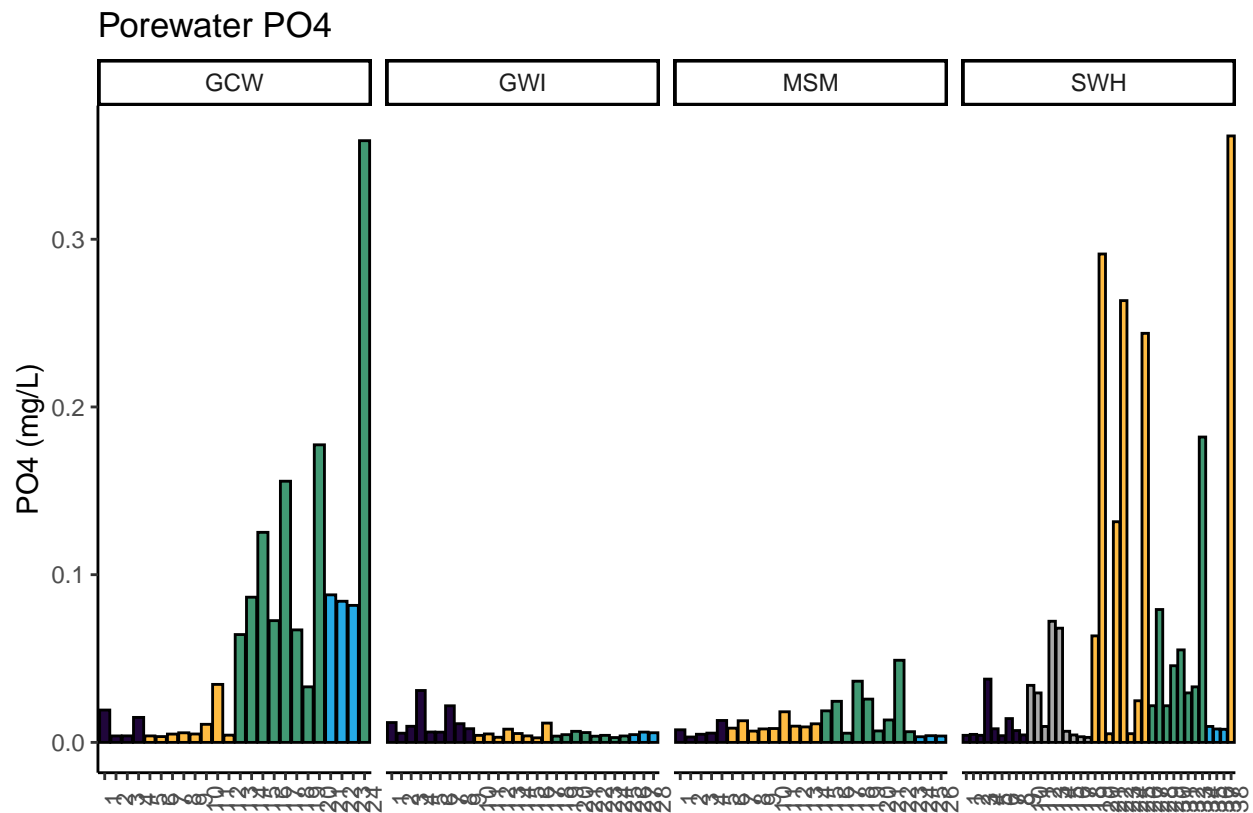
```
## Warning: There was 1 warning in 'mutate()'.  
## i In argument: 'Samp_Time = ym(Samp_Time)'.  
## Caused by warning:  
## ! 2 failed to parse.
```

Visualize Data



Porewater NH3





0.15 Export Processed Data

#end