

Synoptic_CB_Nutrients_2023_October

Month of Data Being Analyzed

2025-07-01

Contents

0.1	Import Data & Clean	2
0.2	Assessing standard Curves	3
0.3	Dilution Corrections - ensure the latest dilution is kept	7
0.4	Performance Check	7
0.5	Analyze the Check Standards	9
0.6	Analyze Blanks	10
0.7	Analyze Duplicates	11
0.8	Spikes	12
0.9	Matrix Effects	13
0.10	Unit Converted Data Column Added (mg/L to uM)	13
0.11	Sample Flagging - Within range of standard curve	13
0.12	Pull out sample id information	13
0.13	Pulling Rhizon Samples	13
0.14	Check to see if samples run match metadata & merge info	14
0.15	Visualize Data	15
0.16	Export Processed Data	17

```

##Run Information

cat("Run Information: Input by I. Van Benschoten") #lets you know what section you're in

## Run Information: Input by I. Van Benschoten

#set the run date & user name
run_date <- "DATE"
sample_year <- 2023
sample_month <- "October"
user <- "Stephanie Wilson"

#identify the files you want to read in
#read in as a list to accommodate multiple runs in a month
NOx_files <- c("Raw Data/SEAL_COMPASS_Synoptic_NOx_Oct2023_1.csv",
              "Raw Data/SEAL_COMPASS_Synoptic_NOx_Oct2023_2.csv",
              "Raw Data/SEAL_COMPASS_Synoptic_NOx_Oct2023_3.csv")
NH3_P04_files <- c("Raw Data/SEAL_COMPASS_Synoptic_NH3_P04_Oct2023_1.csv",
                   "Raw Data/SEAL_COMPASS_Synoptic_NH3_P04_Oct2023_2.csv",
                   "Raw Data/SEAL_COMPASS_Synoptic_NH3_P04_Oct2023_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_202310.csv"

#record any notes about the run or anything other info here:
run_notes <- ""

#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

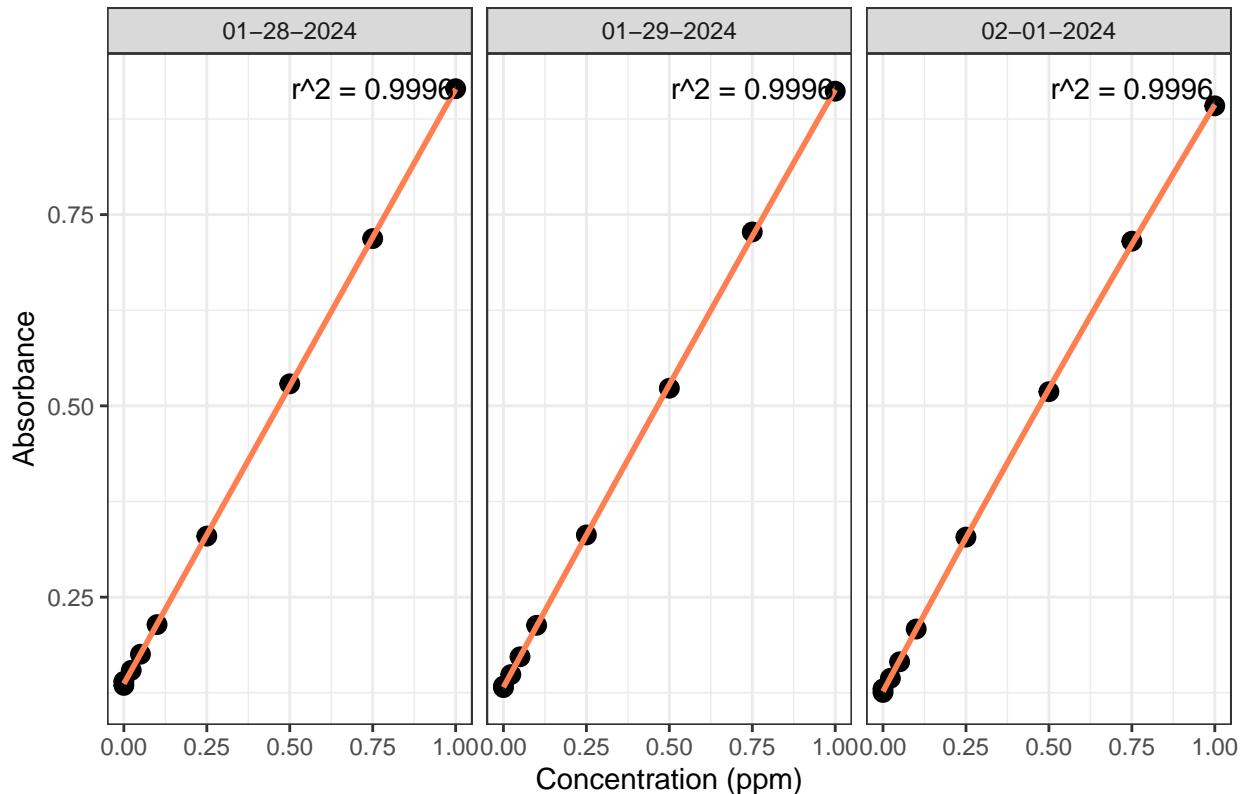
```

0.1 Import Data & Clean

0.2 Assessing standard Curves

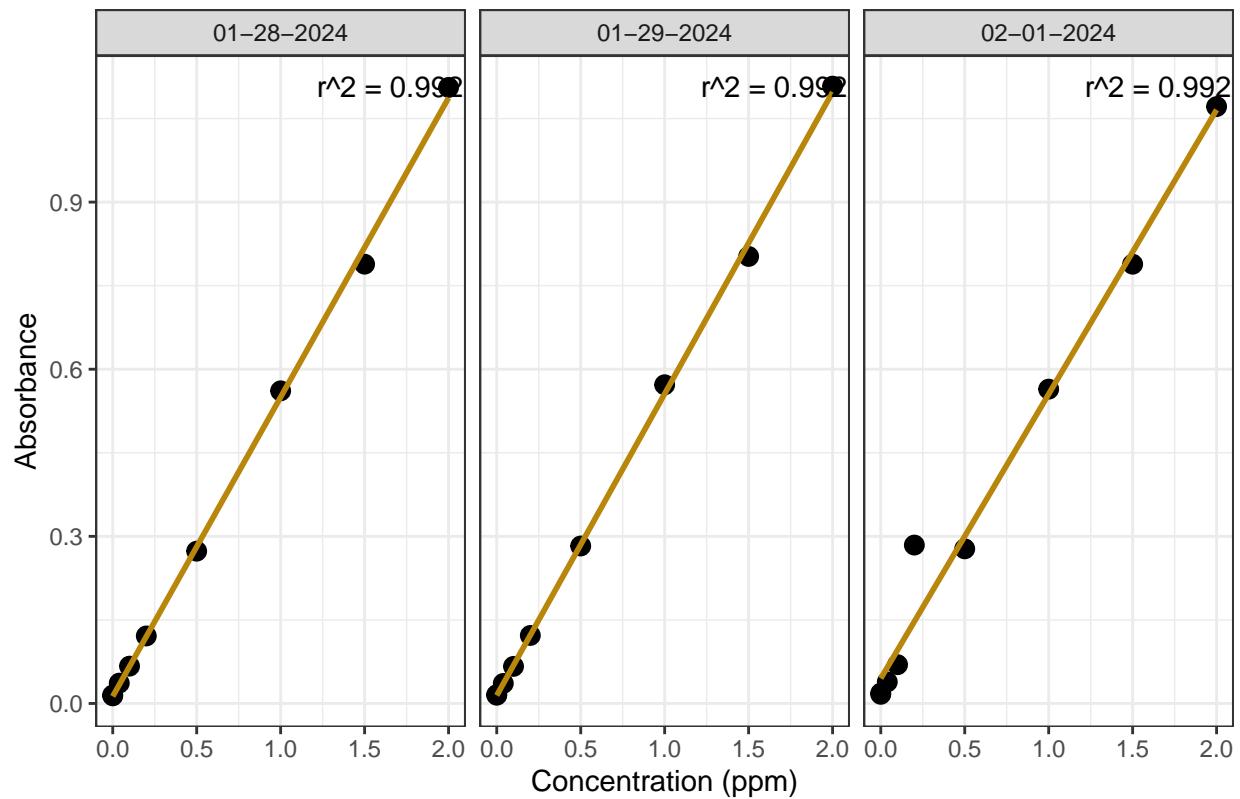
```
#Pull out standards data  
  
## Assess Standard Curves  
  
#Plot standards data  
  
## Assess Standard Curves
```

NOx Standard Curve



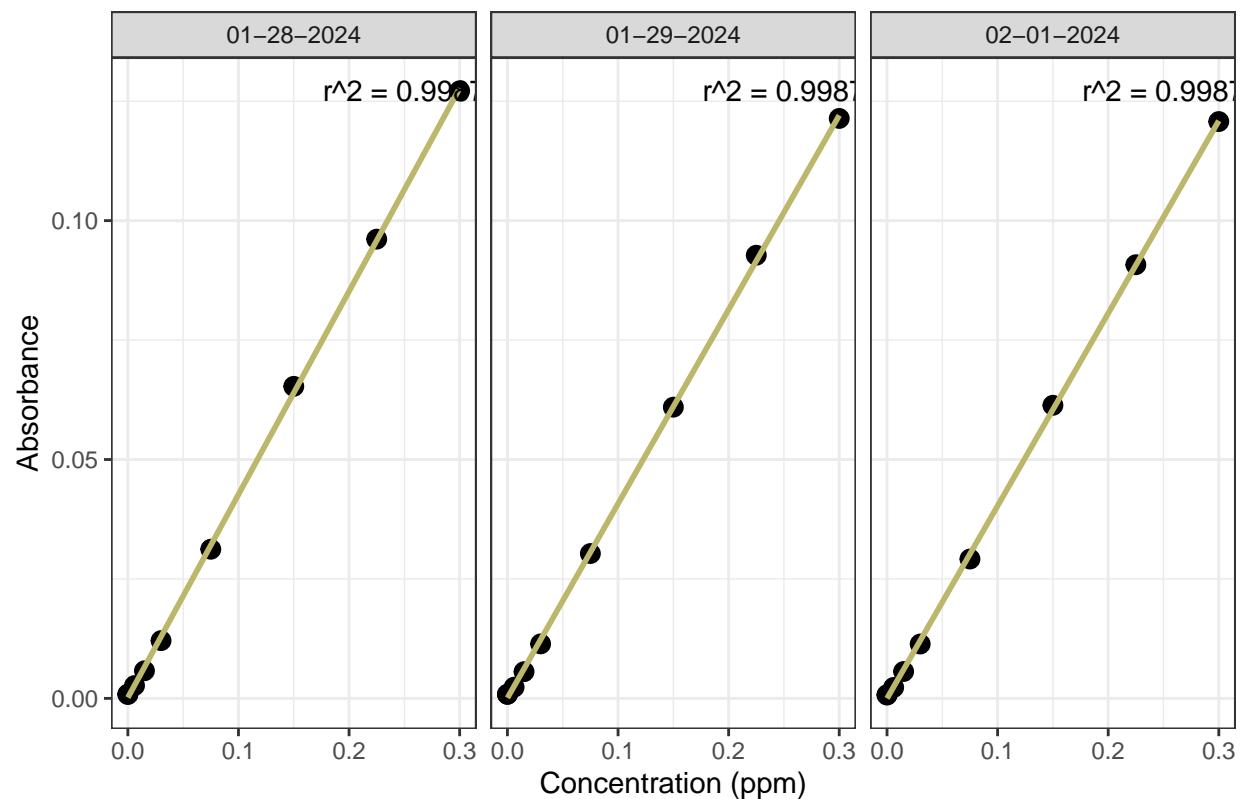
```
## `geom_smooth()` using formula = 'y ~ x'
```

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."
```

Slope Drift Assessment

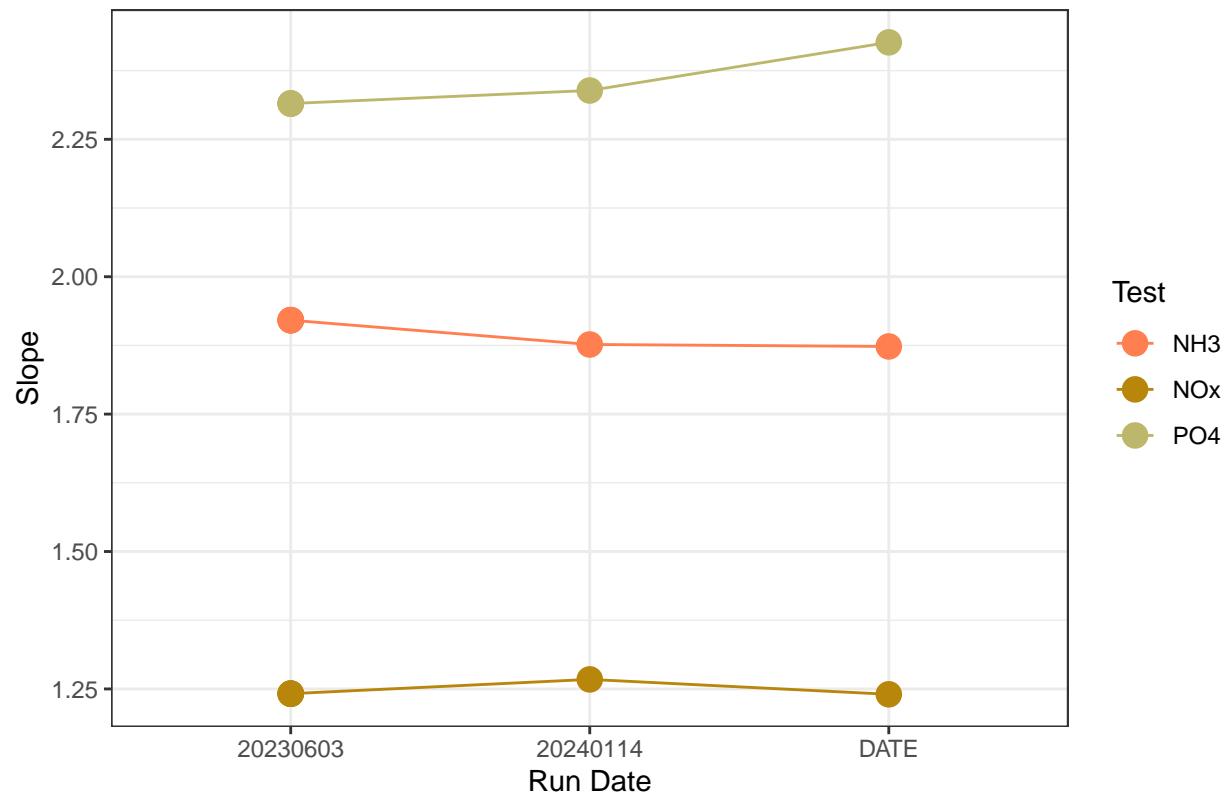


Table 1: Average Slope by Analyte

Test	avg_slope
NH3	1.906
NOx	1.246
PO4	2.338

0.3 Dilution Corrections - ensure the latest dilution is kept

```
## Dilution Corrections

## Duplicated samples: GCW_202310_TR_LysB_20cm, GCW_202310_WC_LysA_45cm, MSM_202310_TR_LysA_20cm, MSM_202310_WC_LysB_45cm

## All duplicated samples have valid dilutions. No naming issues detected.
```

0.4 Performance Check

```
## [1] "NOx pe Check has a % Difference <10% - PROCEED"
```

```
## Run mean = 0.7126627
```

```
## Expected = 0.706
```

```
## [1] "NH3 pe Check has a % Difference >10% - REASSESS"
```

```
## Run mean = 1.482162
```

```
## Expected = 0.948
```

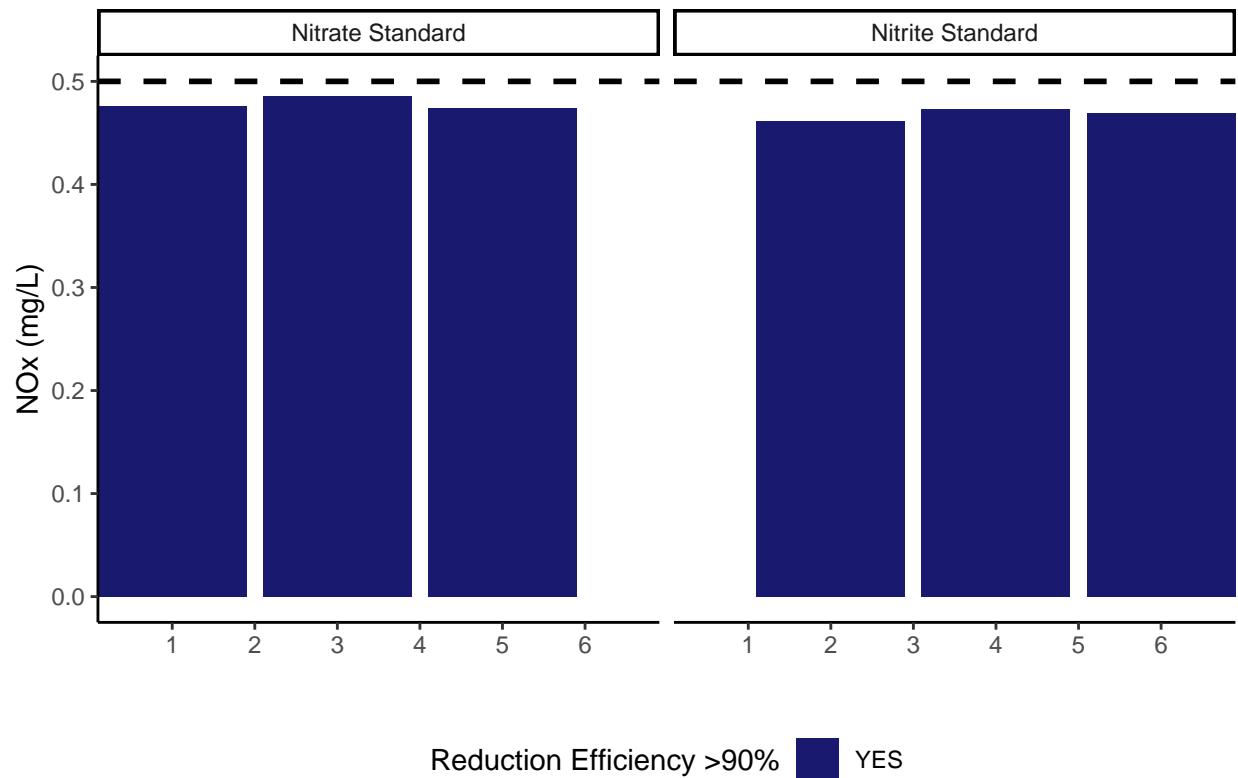
```
## [1] "PO4 pe Check has a % Difference >10% - REASSESS"
```

```
## Run mean = 0.9255853
```

```
## Expected = 0.818
```

```
#Check NOx Reduction Efficiency
```

```
## Assess Reduction Efficiency
```



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

```
## [1] 94.6084
```

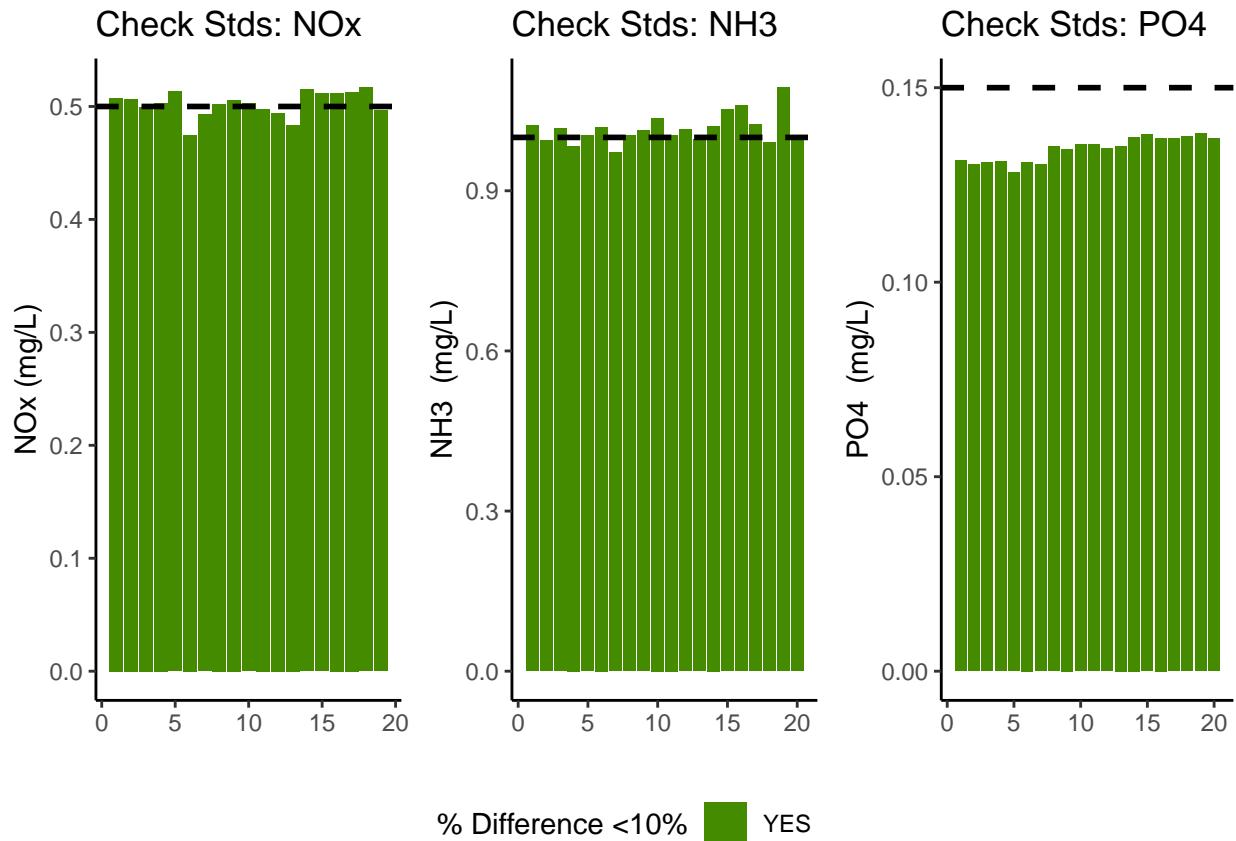
0.5 Analyze the Check Standards

```
## 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"
```

0.6 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"
```

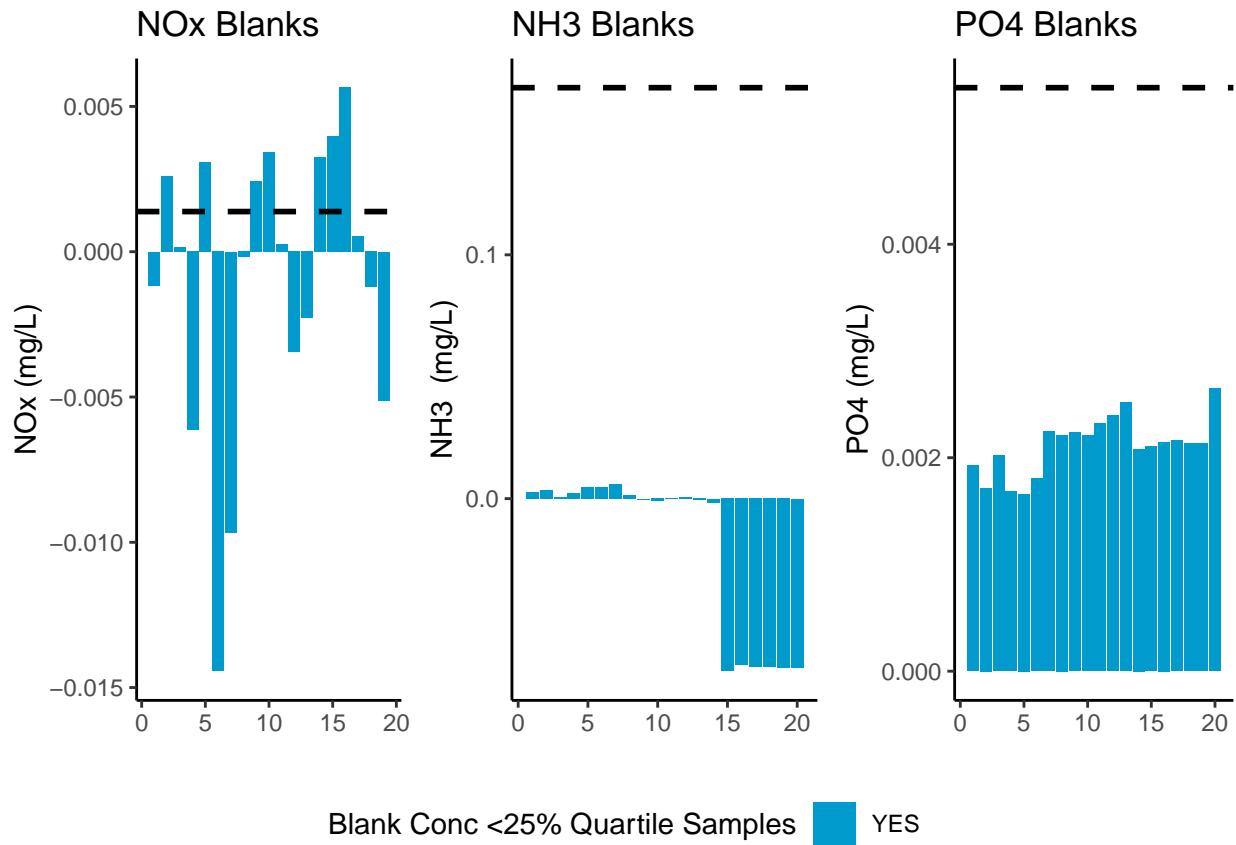


Table 2: Mean Concentration of Blanks

Test	Blank_Mean_Conc
NOx	-0.0010
NH3	-0.0196
PO4	0.0021

0.7 Analyze Duplicates

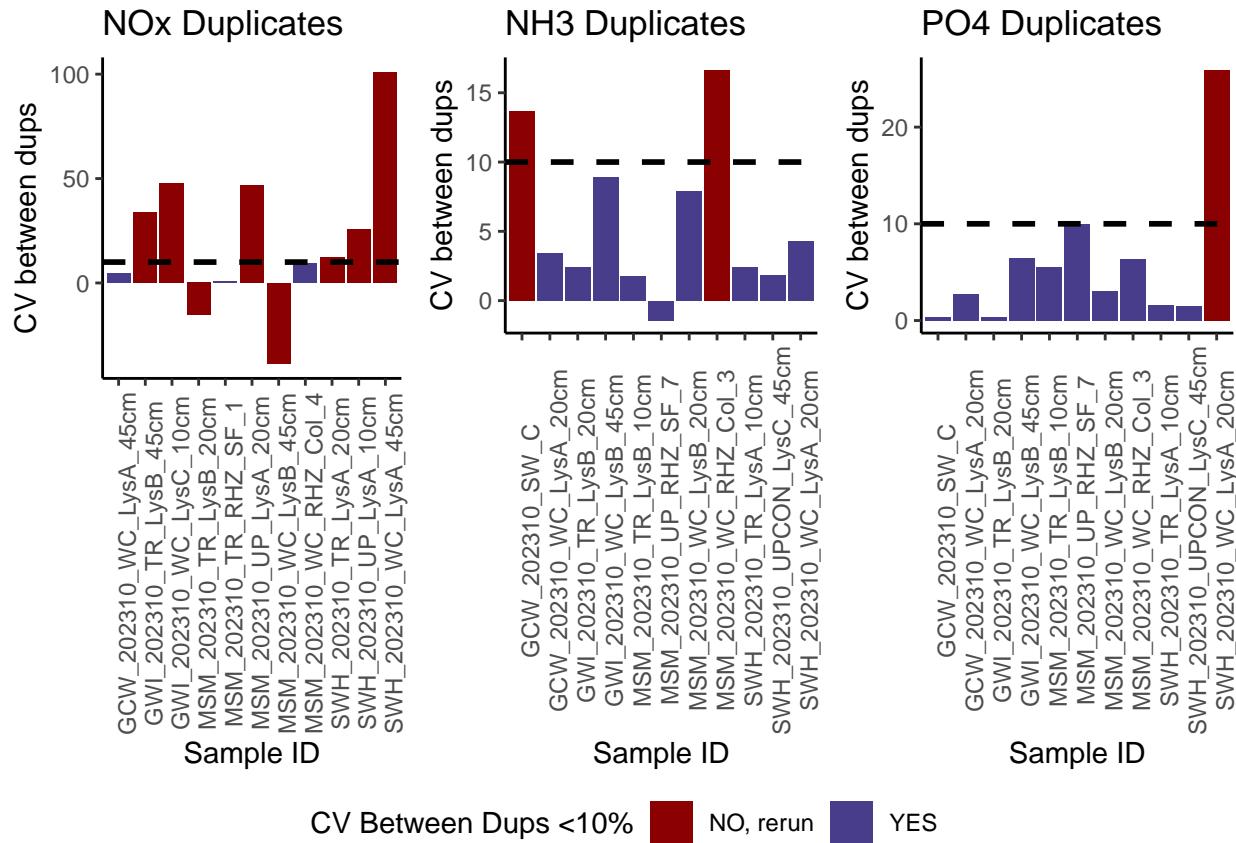
```
## Analyze Duplicates

## [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.
```

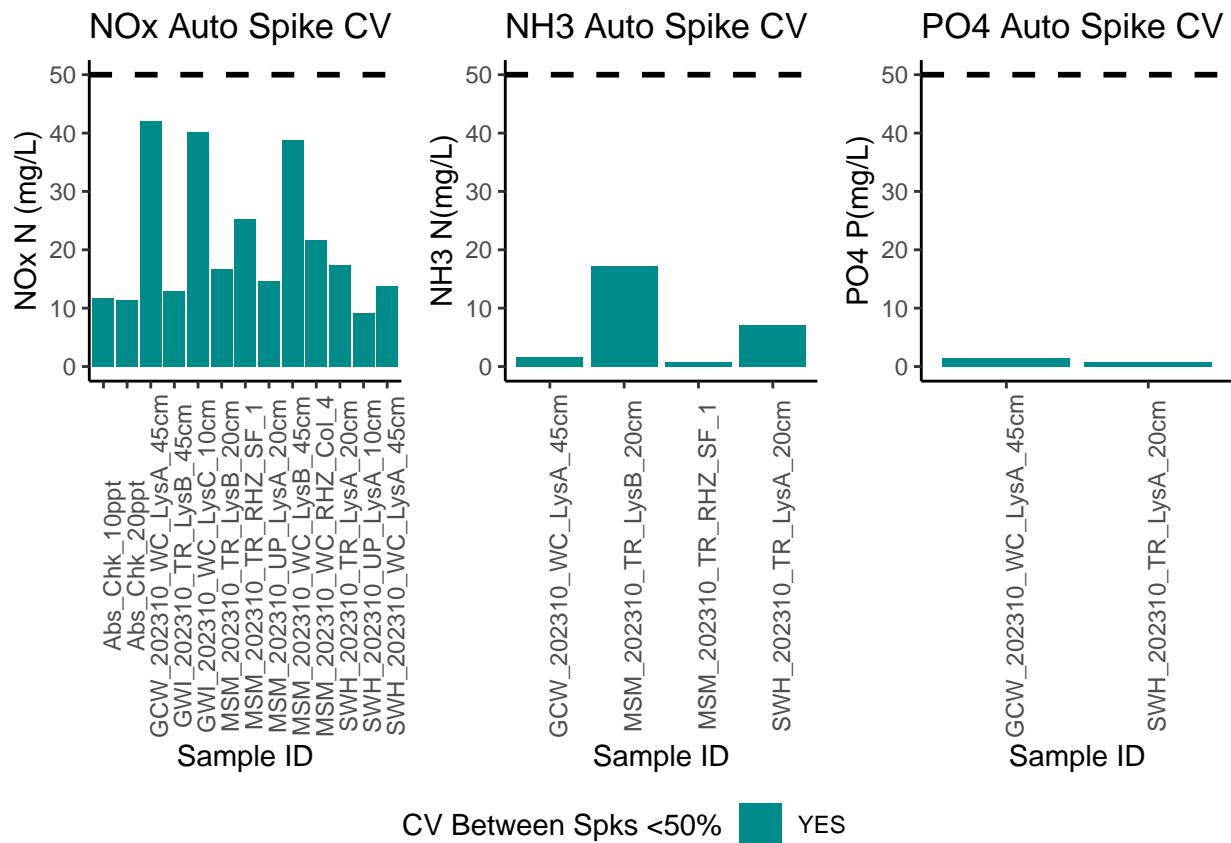


0.8 Spikes

```
## [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"
```



0.9 Matrix Effects

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

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

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

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

0.11 Sample Flagging - Within range of standard curve

```
## Sample Flagging
```

0.12 Pull out sample id information

```
## Sample Processing
```

```
## Warning: Expected 5 pieces. Missing pieces filled with 'NA' in 36 rows [17, 18, 19, 44,
## 45, 46, 83, 84, 85, 107, 108, 109, 124, 125, 126, 146, 147, 148, 163, 164,
## ...].
```

0.13 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 ^^

```

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

```

## Check Sample IDs with Metadata

## Some sample IDs are missing from metadata.

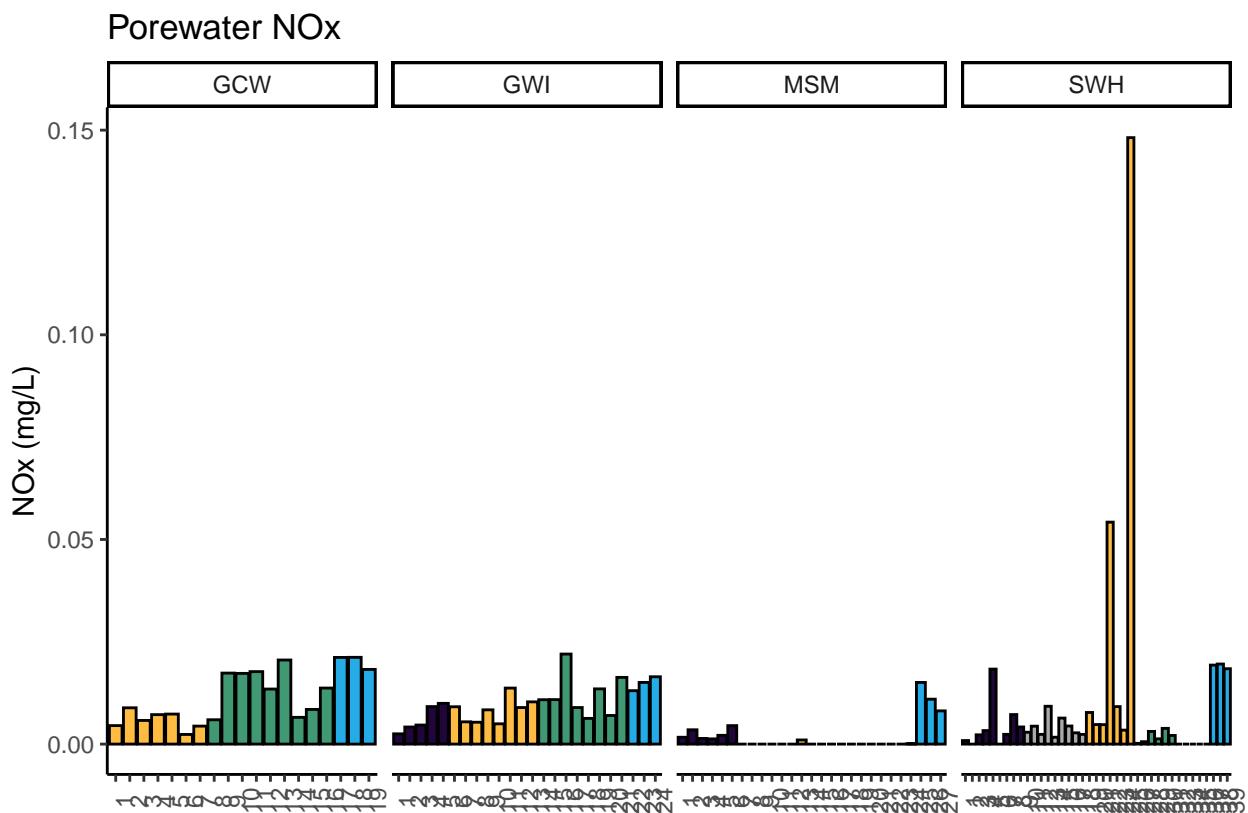
## [1] "GCW_202310_TR_LysA_45cm"      "GCW_202310_TR_LysB_45cm"
## [3] "GCW_202310_TR_LysC_20cm"      "GCW_202310_TR_LysC_45cm"
## [5] "SWH_202310_UPCON_LysA_10cm"    "SWH_202310_UPCON_LysA_20cm"
## [7] "SWH_202310_UPCON_LysA_45cm"    "SWH_202310_UPCON_LysB_10cm"
## [9] "SWH_202310_UPCON_LysB_20cm"    "SWH_202310_UPCON_LysB_45cm"
## [11] "SWH_202310_UPCON_LysC_10cm"    "SWH_202310_UPCON_LysC_20cm"
## [13] "SWH_202310_UPCON_LysC_45cm"    "SWH_202310_WC_LysA_10cm"
## [15] "SWH_202310_WC_LysA_20cm"       "SWH_202310_WC_LysA_45cm"
## [17] "SWH_202310_WC_LysB_10cm"       "SWH_202310_WC_LysB_20cm"
## [19] "SWH_202310_WC_LysB_45cm"

## Warning: Expected 5 pieces. Missing pieces filled with 'NA' in 36 rows [17, 18, 19, 44,
## 45, 46, 83, 84, 85, 107, 108, 109, 124, 125, 126, 146, 147, 148, 163, 164,
## ...].

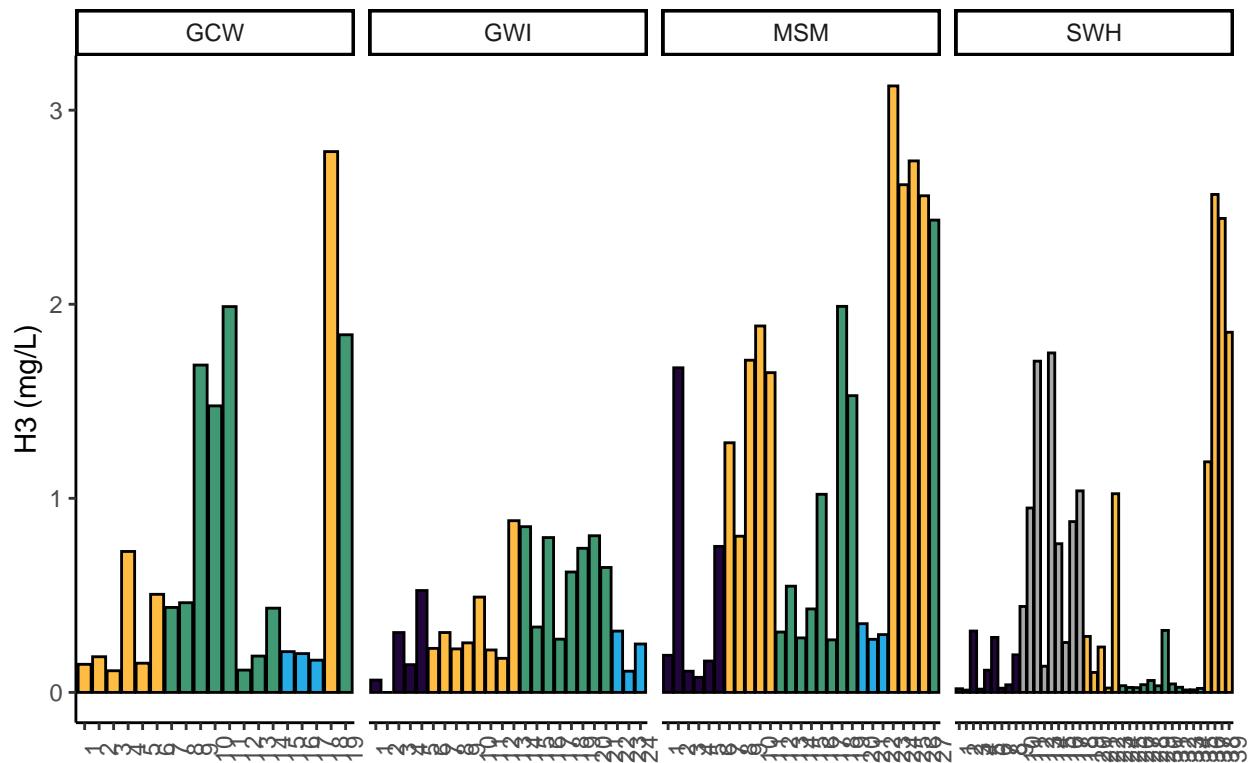
```

0.15 Visualize Data

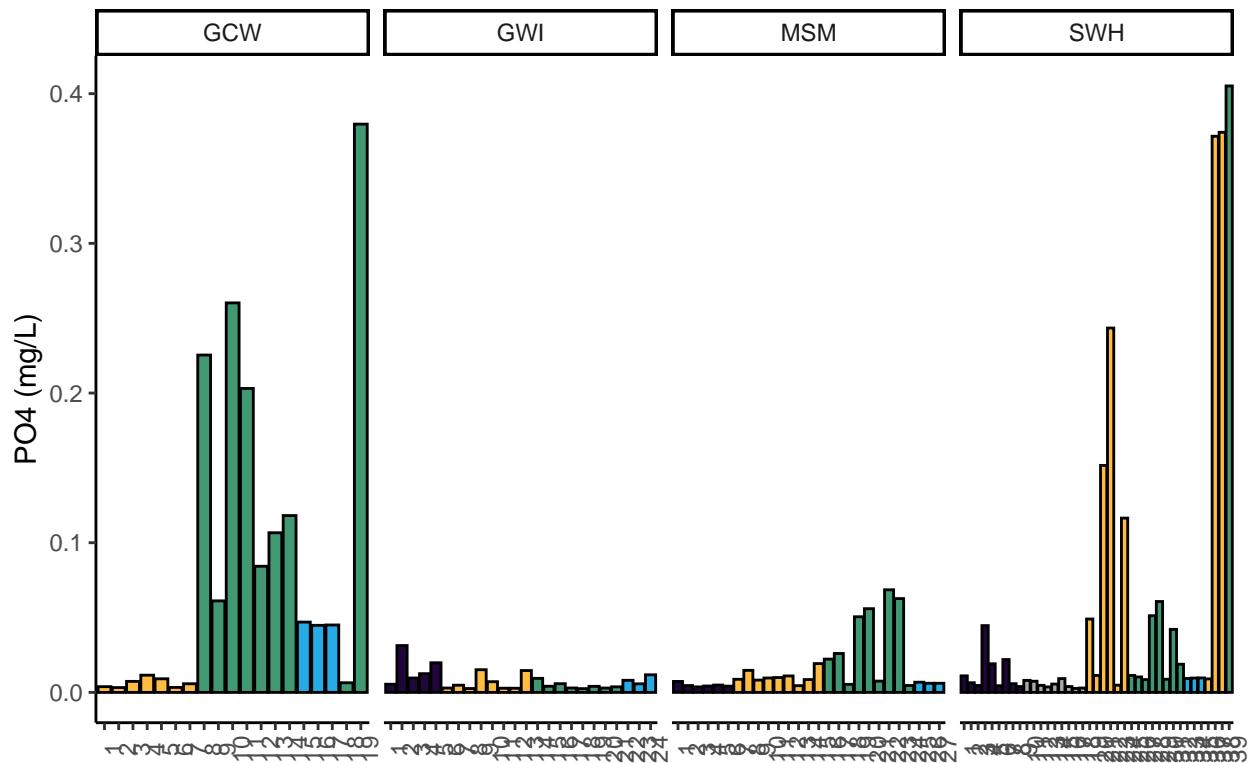
```
## Visualize Data
```



Porewater NH₃



Porewater PO4



0.16 Export Processed Data

#end