COMPASS_TEMPEST_SGW_2023: June

Stephanie J. Wilson

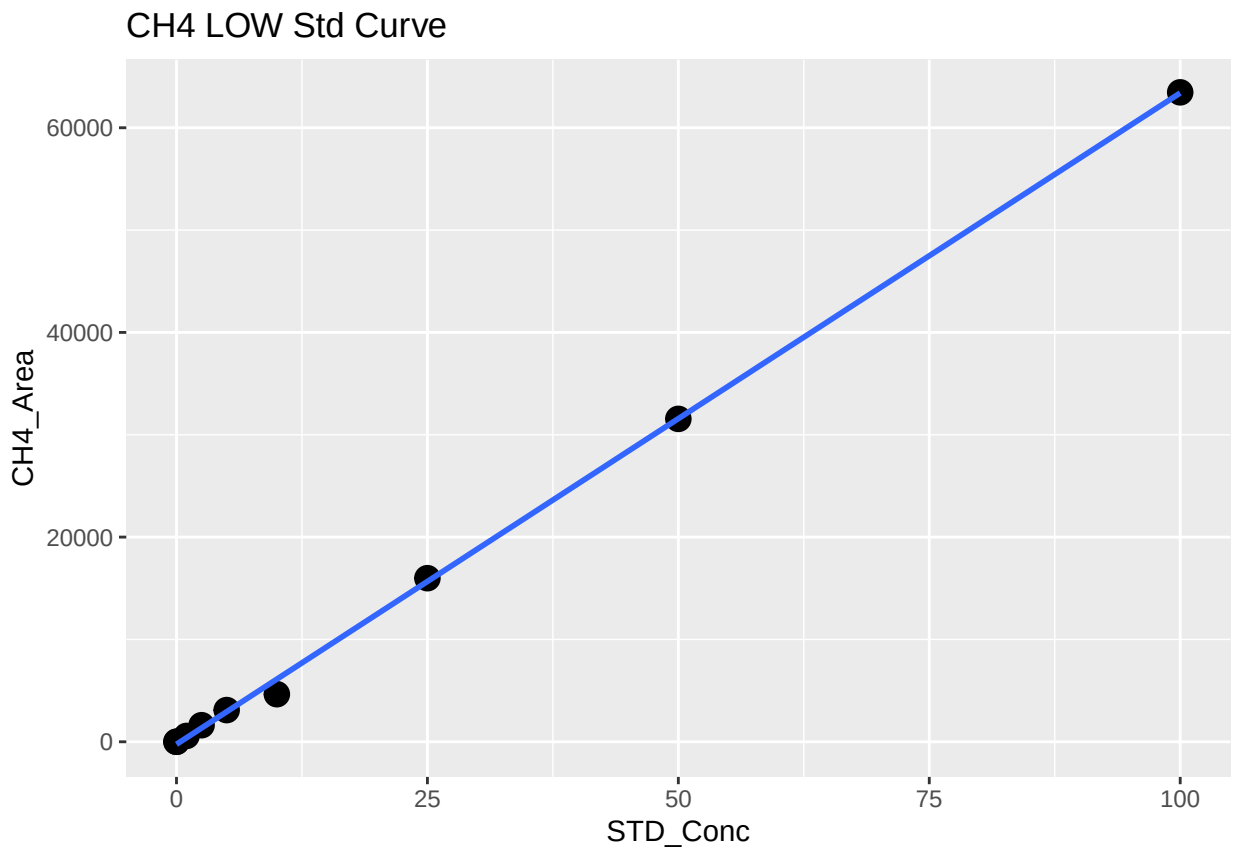
2023-01-14

##Set Up

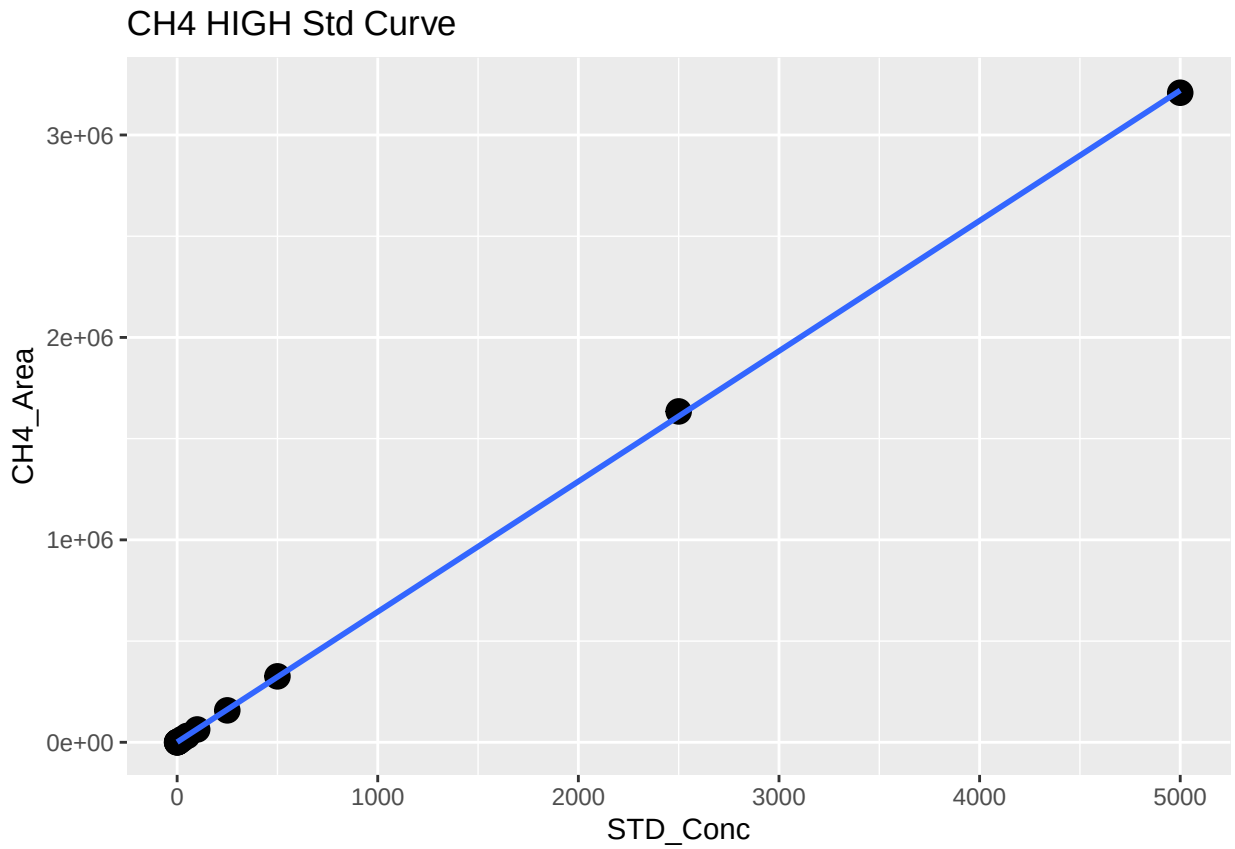
Read in first data file and assess standard curves

```
##     Machine               User  Run_Date Sample_Type Type1 Sample_Year
## 1 Varian GC Stephanie J. Wilson 9/10/2023     Unknown   SGW        2023
## 2 Varian GC Stephanie J. Wilson 9/10/2023     Unknown   SGW        2023
## 3 Varian GC Stephanie J. Wilson 9/10/2023     Unknown   SGW        2023
## 4 Varian GC Stephanie J. Wilson 9/10/2023     Unknown   SGW        2023
## 5 Varian GC Stephanie J. Wilson 9/10/2023     Unknown   SGW        2023
## 6 Varian GC Stephanie J. Wilson 9/10/2023     Unknown   SGW        2023
##   Sample_Month        Sample_ID Dilution_Factor STD_Conc CO2_Area CH4_Area
## 1         June       GWI_TR_ATM               1       NA   198043     1864
## 2         June GWI_TR_SgwA_10cm               1       NA 10917112   324546
## 3         June GWI_TR_SgwA_20cm               1       NA 13406605  1114638
## 4         June GWI_TR_SgwA_45cm               1       NA  9177712    25696
## 5         June GWI_TR_SgwB_10cm               1       NA  5079048   574732
## 6         June GWI_TR_SgwB_20cm               1       NA 22269870  1112553
##   Lab_Notes
## 1        NA
## 2        NA
## 3        NA
## 4        NA
## 5        NA
## 6        NA
```

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

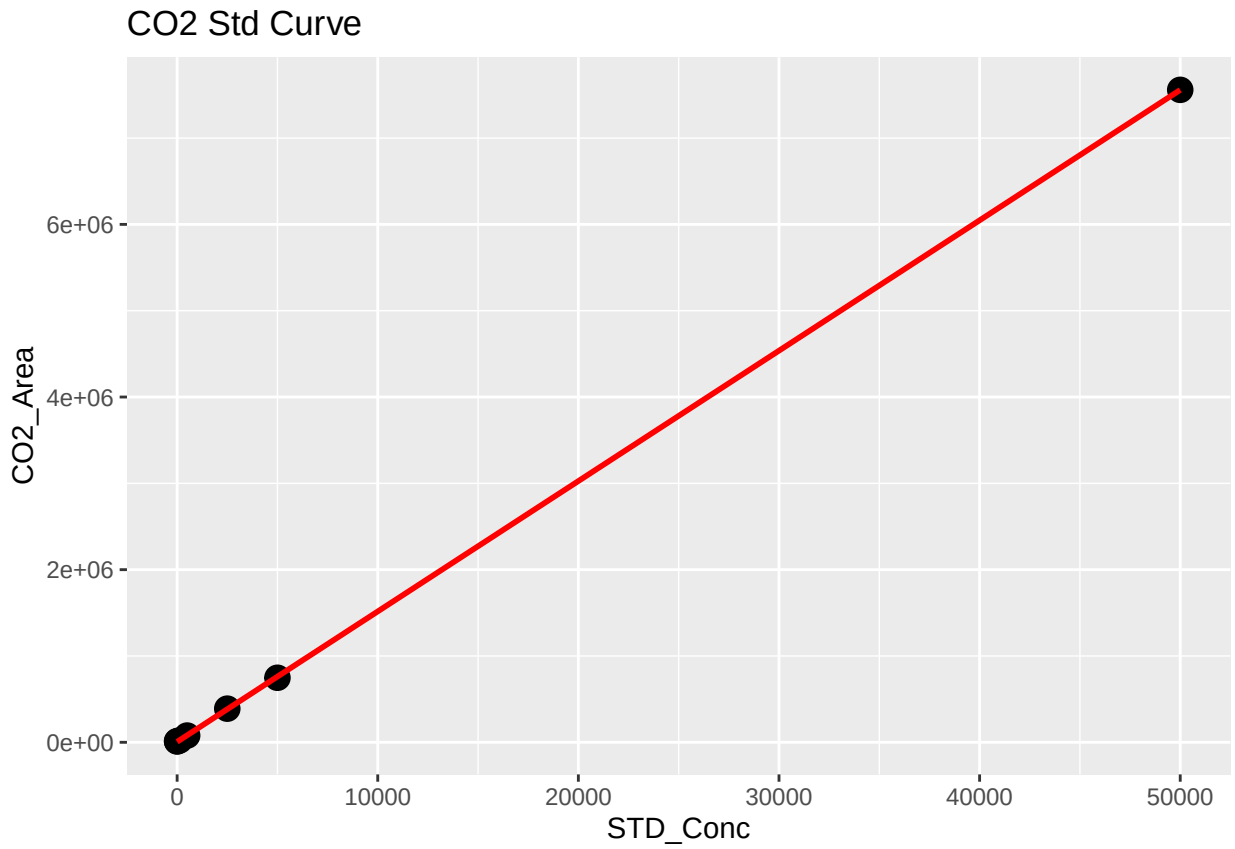



```
##
## Call:
## lm(formula = stds_ch4_low$CH4_Area ~ stds_ch4_low$STD_Conc)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1468.42    62.52   180.69   247.14   319.74
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)             -247.14     241.41  -1.024     0.34
## stds_ch4_low$STD_Conc    636.46       6.29 101.187 2.43e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 599.9 on 7 degrees of freedom
## Multiple R-squared:  0.9993, Adjusted R-squared:  0.9992
## F-statistic: 1.024e+04 on 1 and 7 DF,  p-value: 2.426e-12

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

## Warning: Removed 5 rows containing non-finite outside the scale range
## (`stat_smooth()`).

## Warning: Removed 5 rows containing missing values or values outside the scale range
## (`geom_point()`).
```



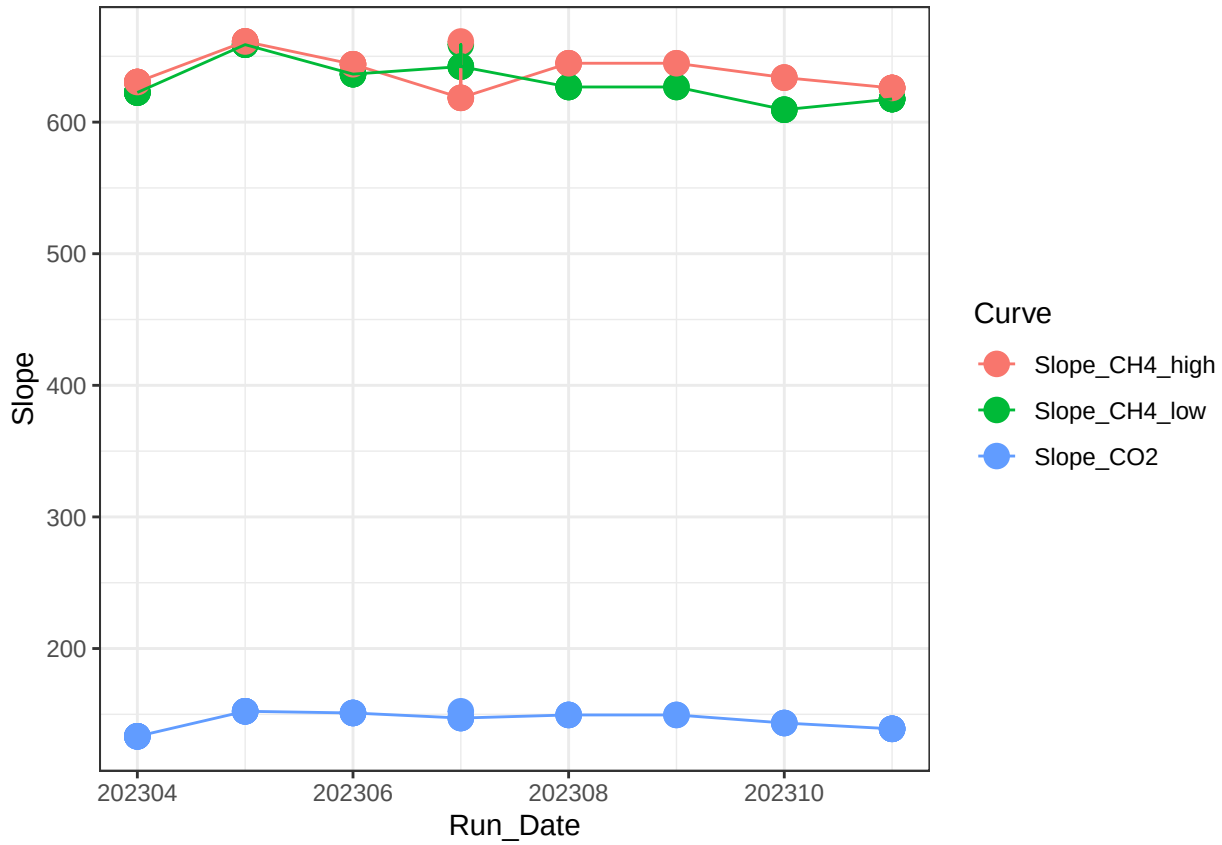

```
##
## Call:
## lm(formula = stds_ch4$CH4_Area ~ stds_ch4$STD_Conc)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -11926.8  -1726.2   -906.7   -791.8  23782.2
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        791.804   2517.397   0.315    0.759
## stds_ch4$STD_Conc  643.954      1.615 398.665   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8251 on 11 degrees of freedom
##   (5 observations deleted due to missingness)
## Multiple R-squared:  0.9999, Adjusted R-squared:  0.9999
## F-statistic: 1.589e+05 on 1 and 11 DF,  p-value: < 2.2e-16

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

## Warning: Removed 2 rows containing non-finite outside the scale range
## (`stat_smooth()`).
```

```
## Warning: Removed 2 rows containing missing values or values outside the scale range
## (`geom_point()`).
```




```
##
## Call:
## lm(formula = stds_co2$CO2_Area ~ stds_co2$STD_Conc)
##
## Residuals:
##      1      2      3      5      7      8      9
##   6491   4889  -2414  -3004   6347 -13363   1055
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)       5838.1736  3258.9619   1.791    0.133
## stds_co2$STD_Conc  151.0364     0.1714 881.341 3.57e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7757 on 5 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:      1,  Adjusted R-squared:      1
## F-statistic: 7.768e+05 on 1 and 5 DF,  p-value: 3.569e-14

##   X          Curve        R2    Slope   Intercept Run_Date
## 1 1  Slope_CH4_low 0.9997982 642.2013  -176.61917   202307
```

```
## 2 2 Slope_CH4_high 0.9995317 618.4551 11202.40792   202307
## 3 3      Slope_CO2 0.9999842 147.1737 10330.39781   202307
## 4 4  Slope_CH4_low 0.9990979 658.8841   -87.56451   202307
## 5 5 Slope_CH4_high 0.9999901 661.2523   176.62351   202307
## 6 6      Slope_CO2 0.9999978 152.3154  8787.66721   202307
```



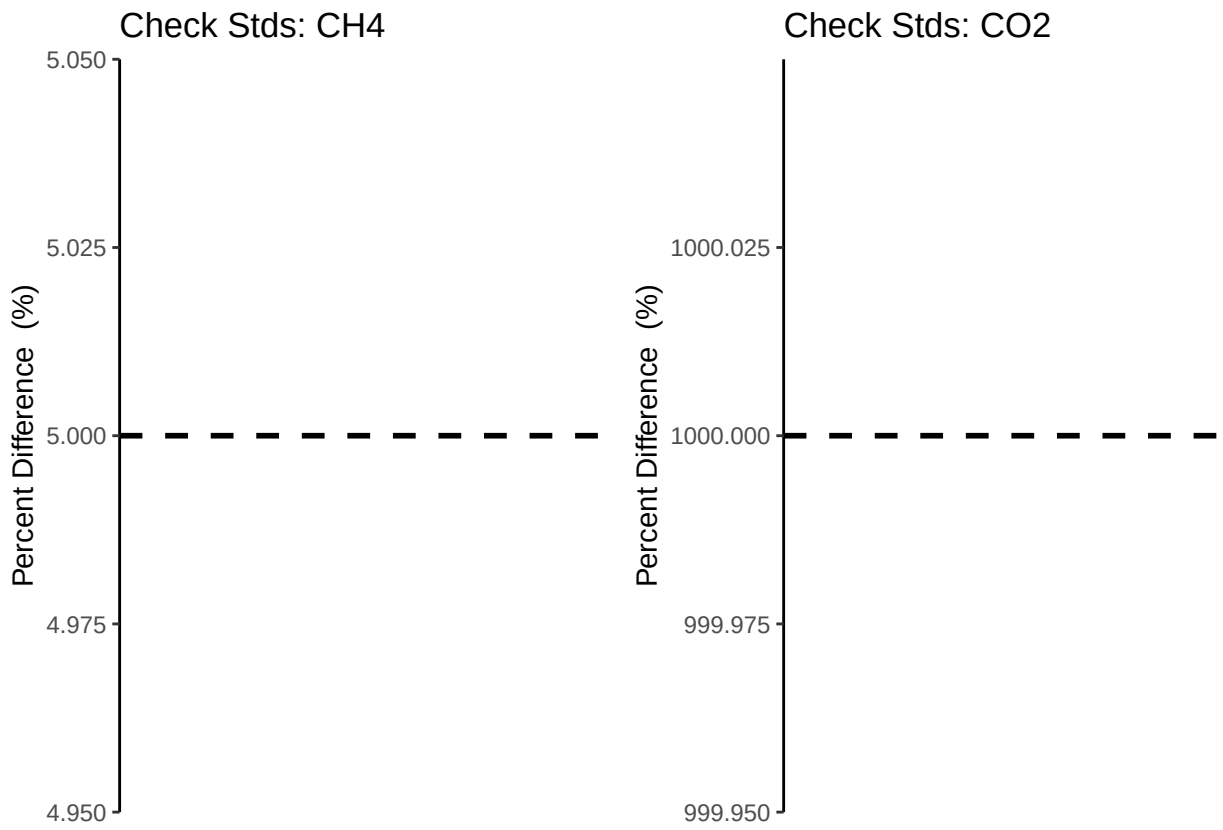
Now calculate the CH4 & CO2 concentrations in ppm

```
#head(raw)

#pull out methane standards
Samples <-  raw %>%
  filter(!str_detect(Sample_Type, "STD_CH4")) %>%
  filter(!str_detect(Sample_Type, "STD_CO2")) %>%
  filter(!str_detect(Sample_Type, "Blank")) %>%
  filter(!str_detect(Sample_Type, "Chk_STD")) %>%
   filter(!str_detect(Sample_Type, "CHKSTD")) %>%
  filter(!str_detect(Sample_Type, "NA"))
#head(Samples)

#Now flag any areas that are above the 100ppm area for CH4
Samples$CH4_Curve <- ifelse(Samples$CH4_Area >71000, "High", "Low")
#head(Samples)
```

```
#Calculate CH4 concentrations in ppm
Samples$CH4_Conc_ppm <- ifelse(Samples$CH4_Area >71000, (Samples$CH4_Area-Slope_CH4_high$Intercept)/Slo

#Calculate CO2 concentrations in ppm
Samples$CO2_Conc_ppm <- ((Samples$CO2_Area-Slope_CO2$Intercept)/Slope_CO2$Slope)

#head(Samples)

##########make flags for any dilutions needed
#highest CH4 standard = 10000
#highest CO2 standard = 50000

Samples$CH4_Flag <- ifelse(Samples$CH4_Conc_ppm >10000, "Needs Dilution", "Within Range")
Samples$CO2_Flag <- ifelse(Samples$CO2_Conc_ppm >50000, "Needs Dilution", "Within Range")
#head(Samples)
```

Check the Check Standards

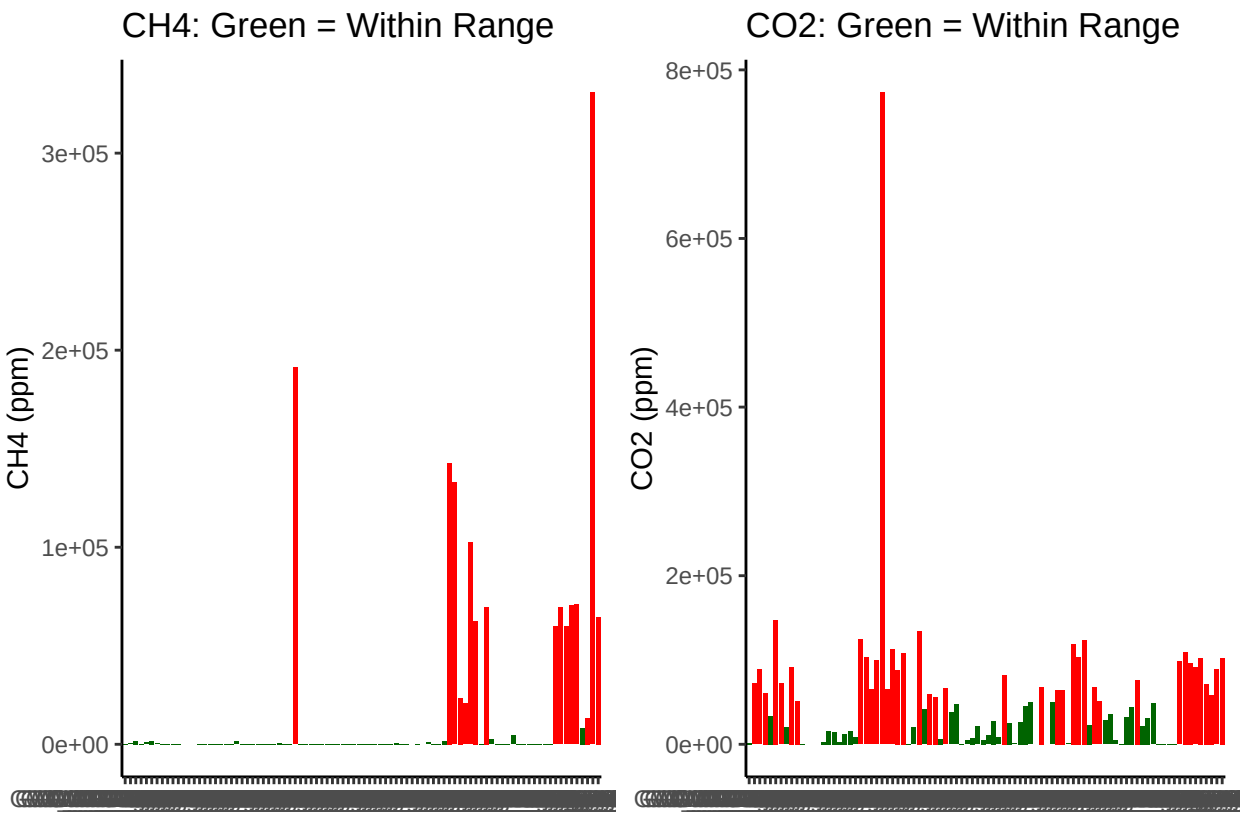
```
## Warning: Using ‘size‘ aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use ‘linewidth‘ instead.
## This warning is displayed once every 8 hours.
## Call ‘lifecycle::last_lifecycle_warnings()‘ to see where this warning was
## generated.
```



Dilution correct samples

```
#multiply the concentration by the dilution factor
Samples$CH4_Conc_ppm_dilcorr <- (Samples$CH4_Conc_ppm * Samples$Dilution_Factor)

Samples$CO2_Conc_ppm_dilcorr <- (Samples$CO2_Conc_ppm * Samples$Dilution_Factor)

#check results
#head(Samples)

#quick first look at the samples
ch4_samples <-  ggplot(data = Samples, aes(x = Sample_ID, y = CH4_Conc_ppm, fill=CH4_Flag)) +
       geom_bar(stat = 'identity') +
        scale_fill_manual(values=c("red","darkgreen"))+
        #scale_fill_gradient2(low='red', mid='white', high='blue', space='Lab') +
        theme_classic() + labs(x= " ", y="CH4 (ppm)", title="CH4: Green = Within Range") +
        theme(legend.position="none")


co2_samples <-  ggplot(data = Samples, aes(x = Sample_ID, y = CO2_Conc_ppm, fill=CO2_Flag)) +
       geom_bar(stat = 'identity') +
        scale_fill_manual(values=c("red", "darkgreen"))+
        #scale_fill_gradient2(low='red', mid='white', high='blue', space='Lab') +
        theme_classic() + labs(x= " ", y="CO2 (ppm)", title="CO2: Green = Within Range") +
        theme(legend.position="none")

ggarrange(ch4_samples, co2_samples, nrow=1, ncol=2)
```

```
## Warning: Removed 5 rows containing missing values or values outside the scale range
## (`geom_bar()`).
## Removed 5 rows containing missing values or values outside the scale range
## (`geom_bar()`).
```



If samples are water calculate gas in water - only need if there is water

Write out processed data & slopes

```
#check results
#head(Samples)

#pull out what we need
Samples1 <- Samples[ ,c(1:3,6:9,17:20)]
#head(Samples1)

IDs <- data.frame(do.call('rbind', strsplit(as.character(Samples1$Sample_ID),'_',fixed=TRUE)))
```

```
## Warning in rbind(c("GWI", "TR", "ATM"), c("GWI", "TR", "SgwA", "10cm"), :
## number of columns of result is not a multiple of vector length (arg 1)
```

```
colnames(IDs) <- c("Site" , "Zone", "Replicate", "Depth")
IDs$Depth1 <- ifelse(IDs$Depth == '10cm', '10',
                ifelse(IDs$Depth == '20cm', '20',
                       ifelse(IDs$Depth == '45cm', '45', '0')))
head(IDs)
```

```
##   Site Zone Replicate Depth Depth1
```

```
## 1  GWI   TR       ATM   GWI      0
## 2  GWI   TR      SgwA  10cm     10
## 3  GWI   TR      SgwA  20cm     20
## 4  GWI   TR      SgwA  45cm     45
## 5  GWI   TR      SgwB  10cm     10
## 6  GWI   TR      SgwB  20cm     20
```

```
#rejoin them to the dataframe
alldat <- cbind(IDs, Samples1)
head(alldat)
```

```
##   Site Zone Replicate Depth Depth1   Machine                User  Run_Date
## 1  GWI   TR       ATM   GWI      0 Varian GC Stephanie J. Wilson 9/10/2023
## 2  GWI   TR      SgwA  10cm     10 Varian GC Stephanie J. Wilson 9/10/2023
## 3  GWI   TR      SgwA  20cm     20 Varian GC Stephanie J. Wilson 9/10/2023
## 4  GWI   TR      SgwA  45cm     45 Varian GC Stephanie J. Wilson 9/10/2023
## 5  GWI   TR      SgwB  10cm     10 Varian GC Stephanie J. Wilson 9/10/2023
## 6  GWI   TR      SgwB  20cm     20 Varian GC Stephanie J. Wilson 9/10/2023
##   Sample_Year Sample_Month        Sample_ID Dilution_Factor     CH4_Flag
## 1        2023         June       GWI_TR_ATM               1 Within Range
## 2        2023         June GWI_TR_SgwA_10cm               1 Within Range
## 3        2023         June GWI_TR_SgwA_20cm               1 Within Range
## 4        2023         June GWI_TR_SgwA_45cm               1 Within Range
## 5        2023         June GWI_TR_SgwB_10cm               1 Within Range
## 6        2023         June GWI_TR_SgwB_20cm               1 Within Range
##         CO2_Flag CH4_Conc_ppm_dilcorr CO2_Conc_ppm_dilcorr
## 1   Within Range             3.317027             1272.573
## 2 Needs Dilution           502.759823            72242.660
## 3 Needs Dilution          1729.698405            88725.391
## 4 Needs Dilution            40.761863            60726.234
## 5   Within Range           891.275155            33589.311
## 6 Needs Dilution          1726.460596           147408.351
```

```
write.csv(alldat, "Processed Data/COMPASS_CBSYN_SGW_202306_Processed.csv")
```

#end