COMPASS: TEMPEST Discrete DOC Data QAQC

Freshwater Well Test: 2025-11-10

2025-11-25

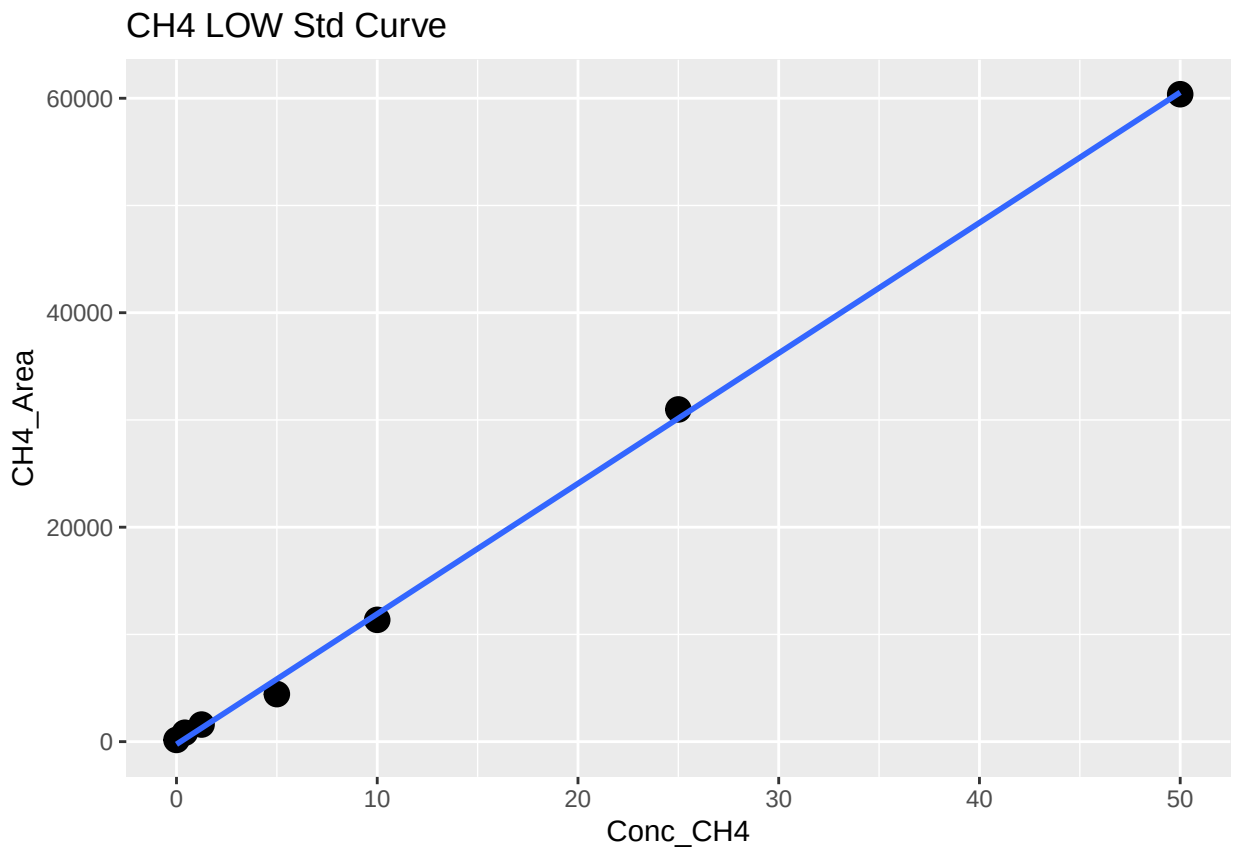
Contents

##Set Up

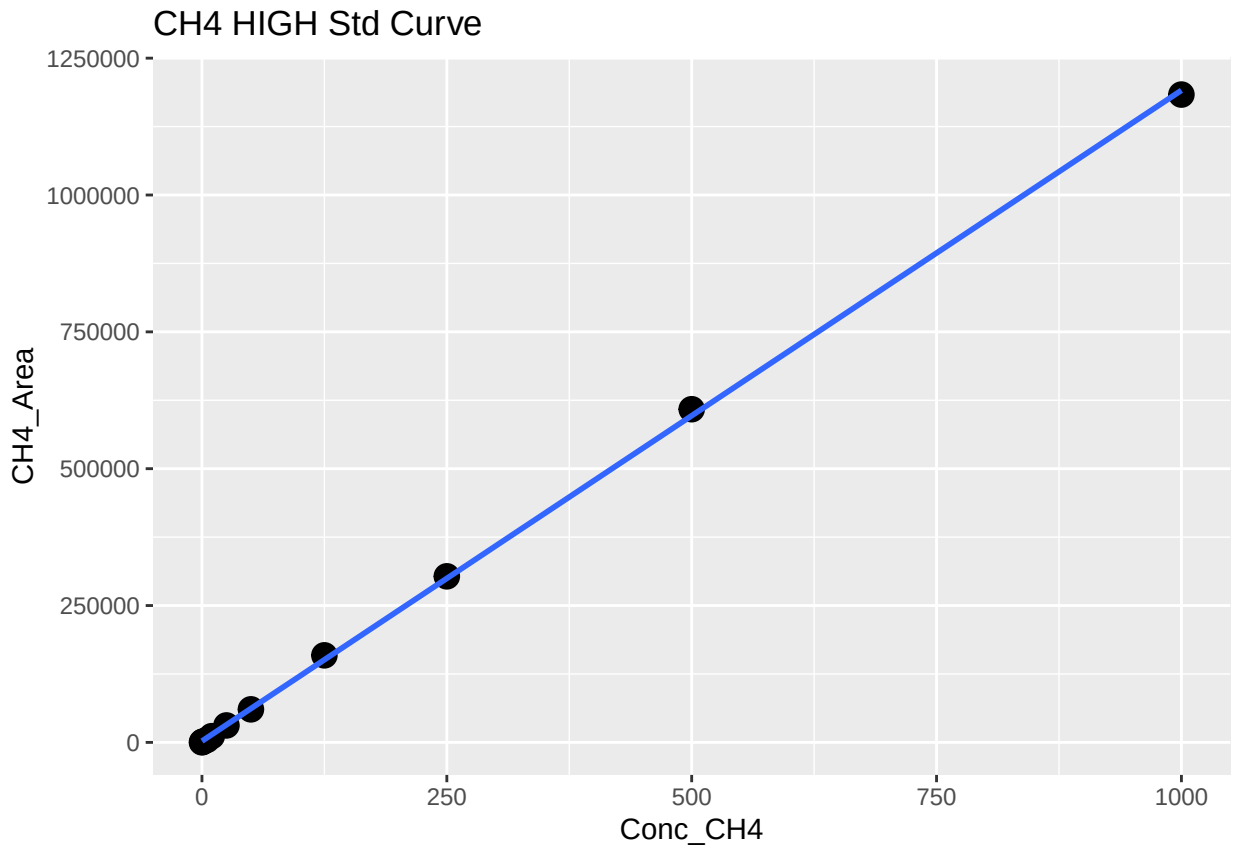
0.1 Read in first data file and assess standard curves

```
##       Machine       User Run_Date Sample_Year Sample_Month Sample_Day
## 1 Shimadzu GC B. Blakley 20251115          NA         <NA>         NA
## 2 Shimadzu GC B. Blakley 20251115          NA         <NA>         NA
## 3 Shimadzu GC B. Blakley 20251115          NA         <NA>         NA
## 4 Shimadzu GC B. Blakley 20251115          NA         <NA>         NA
## 5 Shimadzu GC B. Blakley 20251115          NA         <NA>         NA
## 6 Shimadzu GC B. Blakley 20251115          NA         <NA>         NA
##   Event_Stamp Sample_Time Time_Zone  G_W   Sample.ID Sample_Type    Conc_CO2
## 1        <NA>          NA      <NA> <NA>      LabAir     Lab air     unknown
## 2        <NA>          NA      <NA> <NA>      Blank1       Blank     unknown
## 3        <NA>          NA      <NA> <NA>      Blank2       Blank     unknown
## 4        <NA>          NA      <NA> <NA>        0ppm    Standard           0
## 5        <NA>          NA      <NA> <NA> 0.416ppmCH4    Standard 0.008333333
## 6        <NA>          NA      <NA> <NA>  1.25ppmCH4    Standard       0.025
##      Conc_CH4    Conc_N2O CO2_Area CH4_Area N2O_Area Notes
## 1     unknown     unknown    38195     1552     1859    NA
## 2     unknown     unknown     1550      111      131    NA
## 3     unknown     unknown     1246      164      100    NA
## 4           0           0     1426      161       21    NA
## 5 0.416666667 0.208333333     2548      833     2001    NA
## 6        1.25       0.625    26845     1595     4069    NA
```

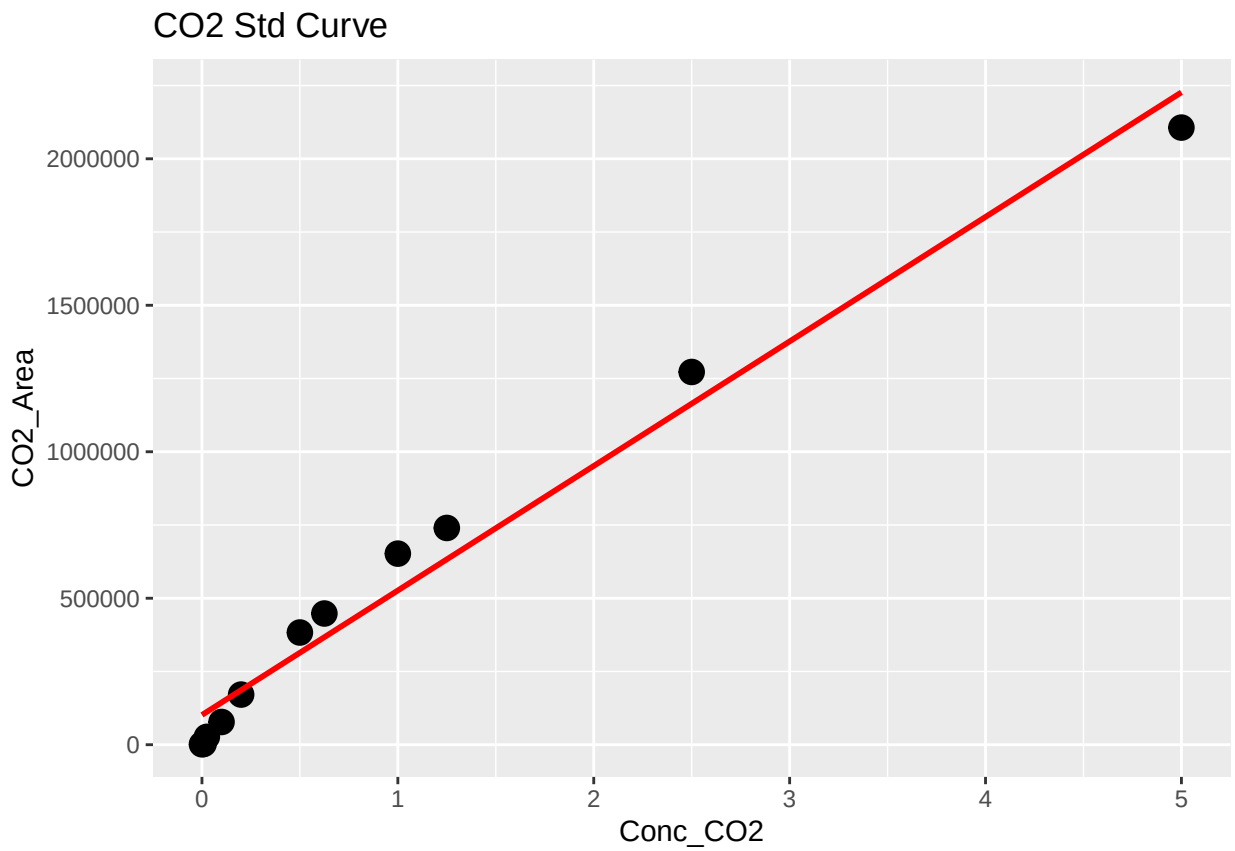
```
## `geom_smooth()` using formula = 'y ~ x'
```




```
##
## Call:
## lm(formula = stds_ch4_low$CH4_Area ~ stds_ch4_low$Conc_CH4)
##
## Residuals:
##       1       2       3       4       5       6       7
##   407.7   573.0   321.7 -1400.2  -554.0   827.5  -175.7
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)           -246.68     402.73  -0.613    0.567
## stds_ch4_low$Conc_CH4  1215.97      18.69  65.076 1.62e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 846.3 on 5 degrees of freedom
## Multiple R-squared:  0.9988, Adjusted R-squared:  0.9986
## F-statistic:  4235 on 1 and 5 DF,  p-value: 1.622e-08

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

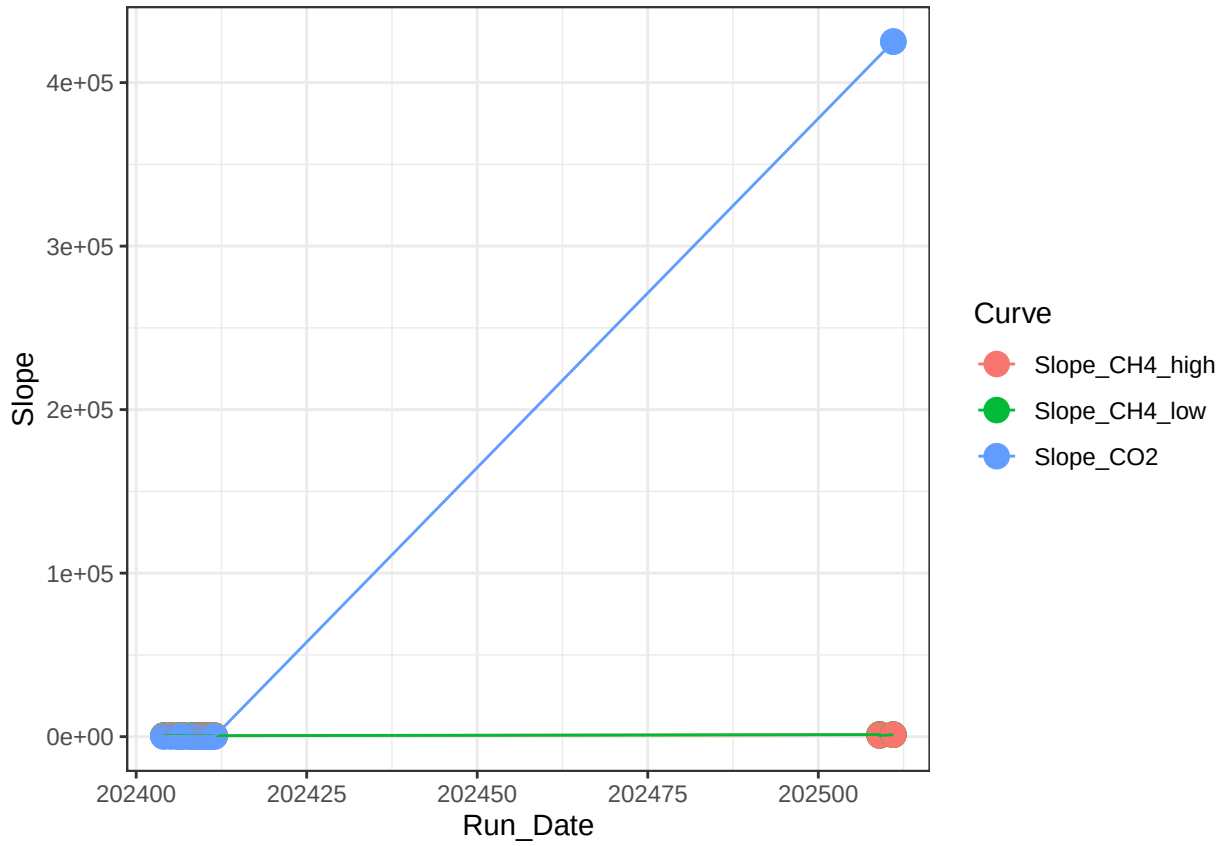



```
##
## Call:
## lm(formula = stds_ch4$CH4_Area ~ stds_ch4$Conc_CH4)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##  -7725  -2597  -2049   1328  11856
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       2387.034   2097.508   1.138    0.284
## stds_ch4$Conc_CH4 1188.835      6.029 197.184   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5968 on 9 degrees of freedom
## Multiple R-squared:  0.9998, Adjusted R-squared:  0.9997
## F-statistic: 3.888e+04 on 1 and 9 DF,  p-value: < 2.2e-16

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




```
##
## Call:
## lm(formula = stds_co2$CO2_Area ~ stds_co2$Conc_CO2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -120368  -92743  -15428   93822  125607
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         101566      38299   2.652   0.0264 *
## stds_co2$Conc_CO2   425054      21624  19.656 1.06e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 103900 on 9 degrees of freedom
## Multiple R-squared:  0.9772, Adjusted R-squared:  0.9747
## F-statistic: 386.4 on 1 and 9 DF,  p-value: 1.058e-08

##       Curve       R2    Slope Intercept
## 1 Slope_CO2 0.974707 425053.7  101565.6

##   X          Curve        R2    Slope  Intercept Run_Date
## 1 1  Slope_CH4_low 0.9996407 633.0314  -320.6965   202404
## 2 2 Slope_CH4_high 0.9993443 601.5512 14817.1191   202404
```

```
## 3 3      Slope_CO2 0.9999907 168.3200 10075.2183   202404
## 4 4  Slope_CH4_low 0.9996407 633.0314  -320.6965   202404
## 5 5 Slope_CH4_high 0.9993443 601.5512 14817.1191   202404
## 6 6      Slope_CO2 0.9999907 168.3200 10075.2183   202404
```

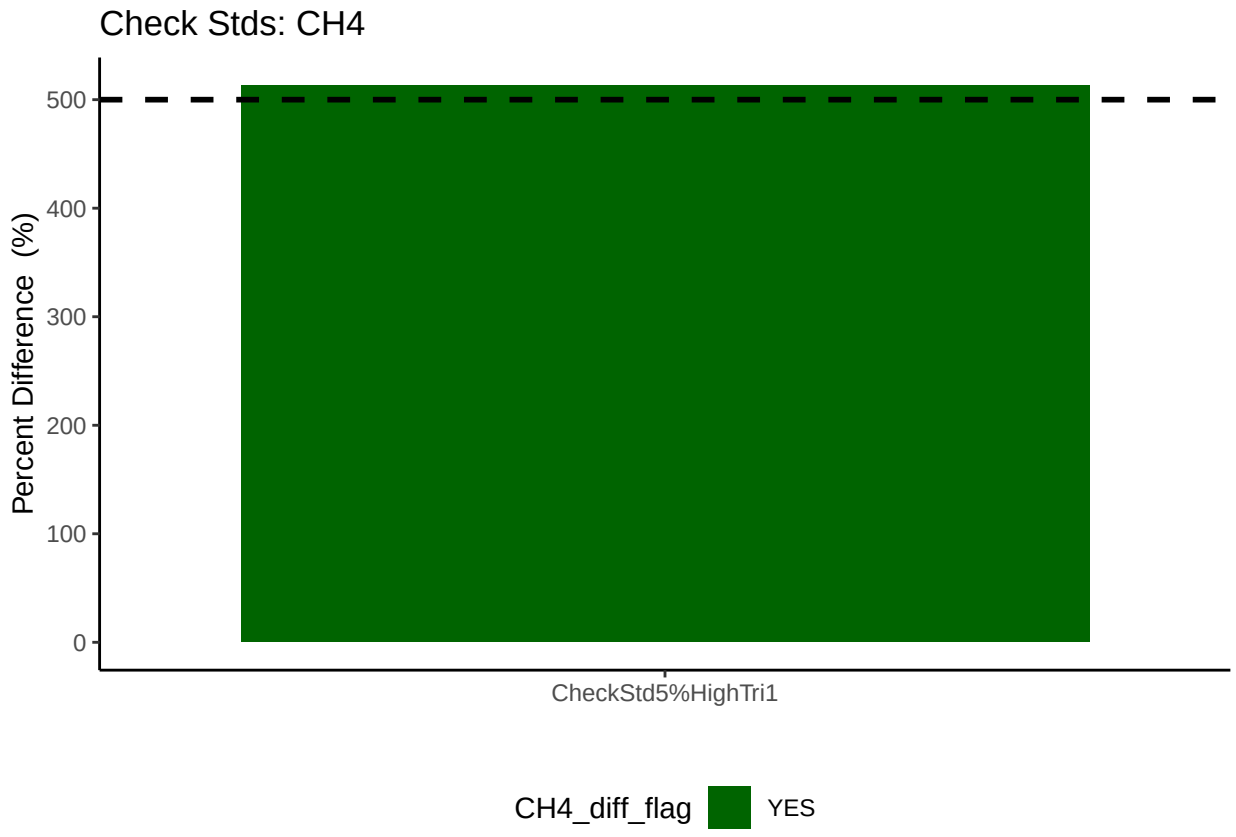


0.2 Now calculate the CH4 & CO2 concentrations in ppm

```
#head(raw)

#pull out methane standards
Samples <-  raw %>%
  filter(!str_detect(Sample_Type, "Standard")) %>%
  filter(!str_detect(Sample_Type, "STD_CO2")) %>%
  filter(!str_detect(Sample_Type, "Blank")) %>%
  filter(!str_detect(Sample_Type, "Lab air")) %>%
   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)
```

```r
#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)
```

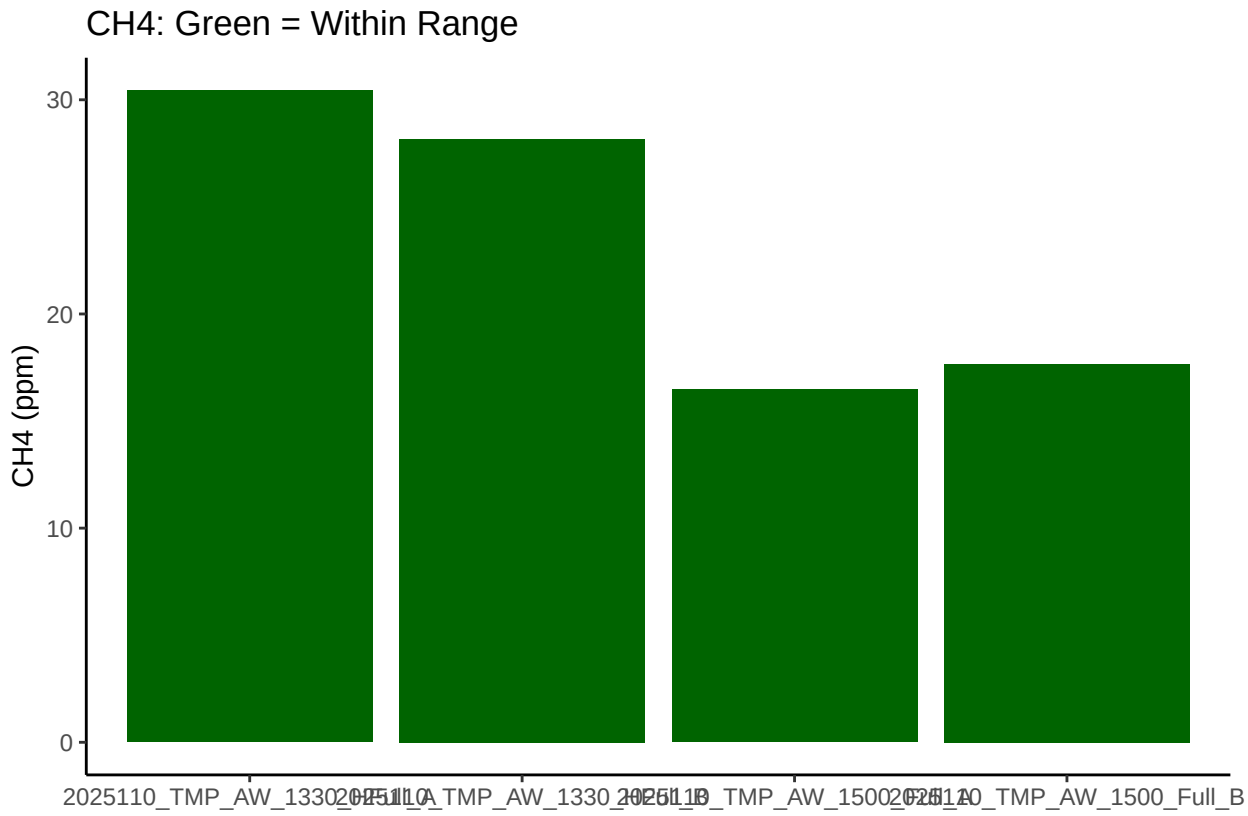
0.3 Check the Check Standards



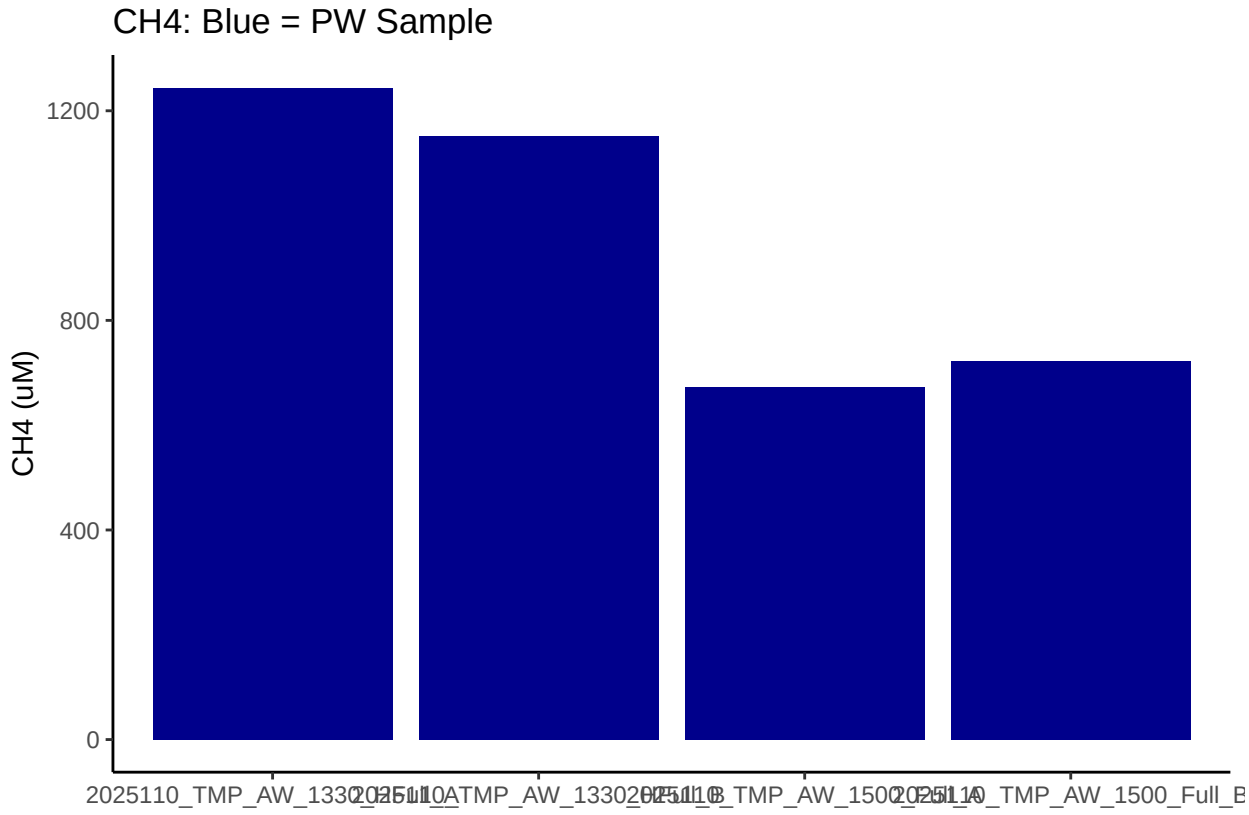
0.4 Dilution correct samples

```r
Samples$Dilution_Factor <- 1
```

```r
#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("darkgreen","red"))+
        #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)


print(ch4_samples)
```


CH4: Green = Within Range
CH4 (ppm)
30
20
10
0
2025110_TMP_AW_1330_Full_B
2025110_TMP_AW_1330_Full_A
2025110_TMP_AW_1500_Full_A
2025110_TMP_AW_1500_Full_B

0.5 If samples are water calculate gas in water



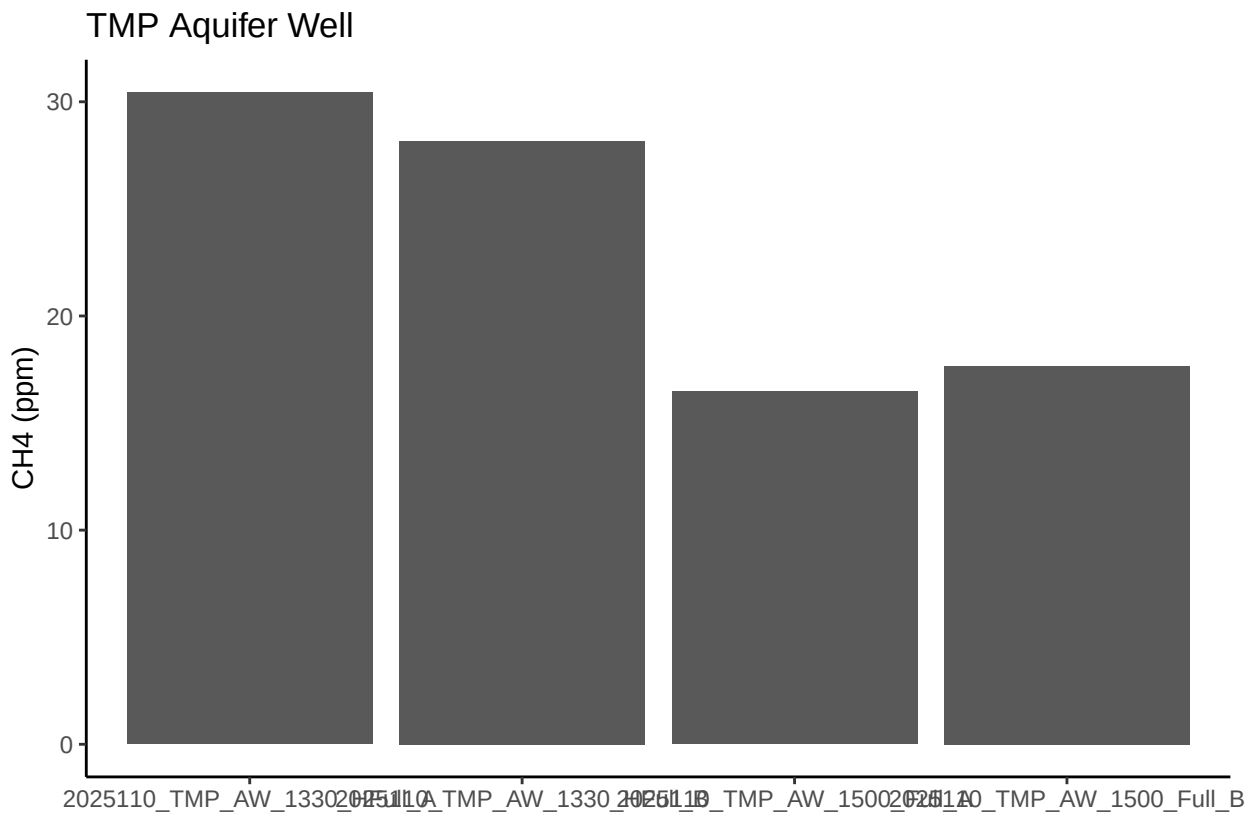
0.6 Write out processed data & slopes

```
#check results
#head(Samples)

#pull out what we need
#Samples1 <- Samples[ ,c(1:3,6:17,21:24)]
#head(Samples1)

Samples1 <-  Samples %>%
  filter(str_detect(Sample_Type, "Sample"))

final_data <- Samples1 %>%
  #select(Project, Plot, grid, sample_name, Vial_ID, date, ) %>%
    rename(Sample_ID = Sample.ID) %>%
  mutate(
    Project = "COMPASS",    # new column with same value on every row
    Experiment = "TEMPEST: Well Test",
    Sample_Date = "2025-11-10",
    Sample_Time = c("13:30", "13:30", "15:00", "15:00"),
    Tank_Status = c("half full", "half full", "full", "full"),
    Replicate = c("A", "B", "A", "B")#,
```

```
    # = run_notes     # new column with notes about the run
  )

#this needs altered to match the tempest metadata and clean up
final_data <- final_data %>%
  rename(
    sample_name = Sample_ID,
    CH4_ppm = CH4_Conc_ppm_dilcorr ,
    CH4_uM = CH4_Conc_umol,
    # add more rename pairs as needed
  ) %>%
  select(
    Project, Experiment, Sample_Date, Sample_Time, Replicate, sample_name,
    CH4_ppm, CH4_uM, CH4_Flag #, tdn_mgL, tdn_uM, tdn_flag, Analysis_runtime,
    #Run_notes
    # list columns in the order you want them
  )

head(final_data)
```

```
##   Project         Experiment Sample_Date Sample_Time Replicate
## 1 COMPASS TEMPEST: Well Test  2025-11-10       13:30         A
## 2 COMPASS TEMPEST: Well Test  2025-11-10       13:30         B
## 3 COMPASS TEMPEST: Well Test  2025-11-10       15:00         A
## 4 COMPASS TEMPEST: Well Test  2025-11-10       15:00         B
##                   sample_name  CH4_ppm    CH4_uM     CH4_Flag
## 1 2025110_TMP_AW_1330_HFull_A 30.44132 1.2442199 Within Range
## 2  2025110_TMP_AW_1500_Full_A 16.47796 0.6734993 Within Range
## 3 2025110_TMP_AW_1330_HFull_B 28.15836 1.1509093 Within Range
## 4  2025110_TMP_AW_1500_Full_B 17.66056 0.7218353 Within Range
```

```
ch4_shakeys <-  ggplot(data = final_data, aes(x = sample_name, y = CH4_ppm)) +
       geom_bar(stat = 'identity') +
        #scale_fill_manual(values=c("darkgrey","darkblue"))+
        #scale_fill_gradient2(low='red', mid='white', high='blue', space='Lab') +
        theme_classic() + labs(x= " ", y="CH4 (ppm)", title="TMP Aquifer Well") +
        theme(legend.position="none")
ch4_shakeys
```




```
write.csv(final_data, "Processed Data/TMP_202511_FW_WellTest_GHG_Processed.csv")
```

#end