

# Test Analysis for Output file 1

Ranvir Singh Virk

2023-11-02

## Loading the data:

```
## Rows: 4218 Columns: 93
## -- Column specification -----
## Delimiter: "\t"
## chr  (1): state
## dbl (92): sim, soln, dist_x, time, step, pH, pe, N, Na, Ca, Mg, K, Al, Si, F...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

Total number of rows: 4218

Number of missing rows: 0

## Processing the data:

Cols needed:

For total CO2 capture (+basalt): Sol 1-5

For soil calcite: Sol 11, Calcite

For Effluent calcite: Sol 11, C(4)

Let's clean the data with choosing only 'transp' for state:

```
data.clean <- data |>
  dplyr::filter(state == "transp") |>
  dplyr::mutate(year = time/(3600*24*365)) |>
  dplyr::select(c('soln', 'Calcite', 'Sr', 'C(4)', 'year'))
```

Now that we have clean data, we need to convert the current mole/liter -> ton CO2/Ha and adjusting the average data for the plot

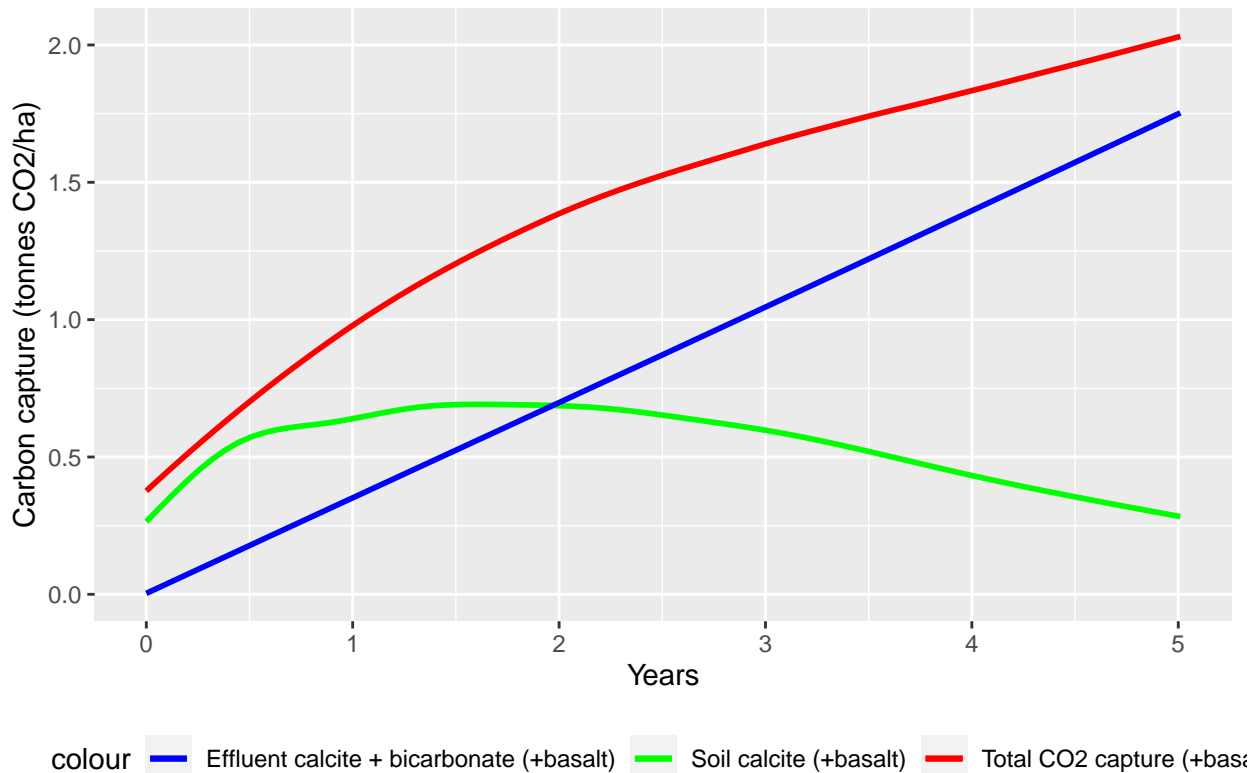
```
conversion_rate <- 500000*44/1000000
plot.data <- data.clean |>
  dplyr::group_by(year) |>
  dplyr::mutate(avg_calcite = ifelse(soln %in% 1:10, mean(Calcite[soln %in% 1:10]), Calcite)) |>
  dplyr::ungroup() |>
  dplyr::mutate(Calcite = avg_calcite) |>
  dplyr::select(-avg_calcite) |>
  dplyr::mutate(Calcite = Calcite * conversion_rate / ifelse(soln %in% 1:10, 1, 10),
    `C(4)` = `C(4)` * conversion_rate/10)
```

Plotting the data:

```
## Warning in full_join(calcite, soln_11_data, by = "year", suffix = c(".soil", : Detected an unexpected
## i Row 2421 of `x` matches multiple rows in `y`.
## i Row 1 of `y` matches multiple rows in `x`.
## i If a many-to-many relationship is expected, set `relationship =
##   "many-to-many"` to silence this warning.

## `geom_smooth()` using method = 'loess' and formula = 'y ~ x'
## `geom_smooth()` using method = 'gam' and formula = 'y ~ s(x, bs = "cs")'
## `geom_smooth()` using method = 'loess' and formula = 'y ~ x'
```

## Carbon Capture Analysis



Plot for Sr over the years:

```
sr_data <- plot.data |>
  filter(soln %in% c(5, 10, 11)) |>
  mutate(Sr = Sr*87.62*1000) |>
  dplyr::select(c("soln", "Sr", "year"))

# Plotting Sr over the years for soln 1, 5, and 11
ggplot(sr_data, aes(x = year, y = Sr, color = factor(soln))) +
  geom_smooth(se=F) +
  labs(title = "Strontium (Sr) Concentration Over Years",
       x = "Year",
       y = "Sr Concentration (mg/L)",
       color = "Solution") +
  scale_color_manual(values = c("red", "green", "blue"),
                    labels = c("Soln 5", "Soln 10", "Soln 11")) +
  theme_minimal() +
  theme(legend.position = "bottom")
```

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

### Strontium (Sr) Concentration Over Years

