

```

filename <- "Biodeposition_experiment_summary - Carbonate chem_calc.csv"
df <- read.csv(paste(wd_Biodep,filename, sep = ""), stringsAsFactors = F, skip = 1)
df$Month <- factor(df$Month, levels=c("September", "April", "June"))
levels(df$Month) <- c("Sept", "April", "June")
df$Site <- factor(df$Site, levels=c("Eel Pond", "Provincetown"))
levels(df$Site) <- c("Eel", "Prov")

```

To estimate pH uncertainty I used this error propagation description on this website: [Error Propagation tutorial \(foothill.edu\)](#), page 5. Essentially, I calculate the SD and mean hydrogen ion concentration such that  $SD(pH) = SD([H+]) / (2.303 * \text{mean}([H+]))$ .

```

df$hydrogen_ion_conc <- 10^(-df$pH.out)

df_summary <- df %>%
  group_by(Month, Site) %>%
  summarise(
    Temp = round(mean(Temp.of.in.situ.seawater, na.rm = TRUE),1),
    Temp_SE = round(sd(Temp.of.in.situ.seawater, na.rm = TRUE) /
      sqrt(sum(!is.na(Temp.of.in.situ.seawater))),1),
    Sal = round(mean(Sal, na.rm = TRUE),1),
    Sal_SE = round(sd(Sal, na.rm = TRUE) / sqrt(sum(!is.na(Sal))),1),
    pH = round(-log10(mean(hydrogen_ion_conc, na.rm = TRUE)), 3),
    pH_SE = round((sd(hydrogen_ion_conc, na.rm = TRUE) /
      (2.303*mean(hydrogen_ion_conc, na.rm = TRUE)))/
      sqrt(sum(!is.na(hydrogen_ion_conc))),3),
    TC02 = round(mean(TC02..mm.kgS.W., na.rm = TRUE),1),
    TC02_SE = round(sd(TC02..mm.kgS.W., na.rm = TRUE) /
      sqrt(sum(!is.na(TC02..mm.kgS.W.))), 1),
    TA_meas = round(mean(TA.corrected..mm.kg.SW., na.rm = TRUE),1),
    TA_meas_SE = round(sd(TA.corrected..mm.kg.SW., na.rm = TRUE) /
      sqrt(sum(!is.na(TA.corrected..mm.kg.SW.))), 1),
    pCO2 = round(mean(pCO2.out..matm., na.rm = TRUE),1),
    pCO2_SE = round(sd(pCO2.out..matm., na.rm = TRUE) /
      sqrt(sum(!is.na(pCO2.out..matm.))), 1),
    arag.sat = round(mean(Omega.aragonite, na.rm = TRUE),2),
    arag_sat_SE = round(sd(Omega.aragonite, na.rm = TRUE) /
      sqrt(sum(!is.na(Omega.aragonite))), 2)
  )

```

`summarise()` has grouped output by 'Month'. You can override using the `.groups` argument.

```
kable(df_summary)
```

Month	Site	Temp	Temp_SE	Sal	Sal_SE	pH	pH_SE	TCO2	TCO2_SE	TA_meas	TA
Sept	Eel	22.9	0.0	33.2	NA	7.830	0.006	1877.0	0.6	1980.3	TA
Sept	Prov	17.4	0.2	33.4	NA	7.935	0.015	2016.5	10.4	2154.3	
April	Eel	11.9	0.2	31.0	NA	7.983	0.009	2004.2	7.2	2204.2	
April	Prov	11.8	0.3	31.7	NA	8.070	NA	1970.1	14.7	2184.8	
June	Eel	18.7	0.1	31.6	NA	7.839	0.030	1904.9	13.5	2065.8	
June	Prov	17.8	0.3	31.1	NA	8.193	0.045	1814.2	18.7	2067.9	

I used the following to add grouped columns to the table [Create Awesome HTML Table with knitr::kable and kableExtra \(r-project.org\)](#) and [Replace column names in kable/R markdown - Stack Overflow](#)

```
#dim(df_summary)
# names(df_summary) <- c("Month",
#                         "Site",
#                         "Temp Mean",
#                         "Temp SE",
#                         "Sal Mean",
#                         "Sal SE",
#                         "pH Mean",
#                         "pH SE",
#                         "TCO2 Mean",
#                         "TCO2 SE",
#                         "TA Mean",
#                         "TA SE",
#                         "pCO2 Mean",
#                         "pCO2 SE",
#                         "Aragonite Sat Mean",
#                         "Aragonite Sat SE")
str(df_summary)
```

```
gropd_df [6 x 16] (S3: grouped_df/tbl_df/tbl/data.frame)
 $ Month      : Factor w/ 3 levels "Sept","April",...: 1 1 2 2 3 3
 $ Site       : Factor w/ 2 levels "Eel","Prov": 1 2 1 2 1 2
 $ Temp       : num [1:6] 22.9 17.4 11.9 11.8 18.7 17.8
 $ Temp_SE    : num [1:6] 0 0.2 0.2 0.3 0.1 0.3
 $ Sal        : num [1:6] 33.2 33.4 31 31.7 31.6 31.1
 $ Sal_SE     : num [1:6] NA NA NA NA NA NA
 $ pH         : num [1:6] 7.83 7.93 7.98 8.07 7.84 ...
 $ pH_SE      : num [1:6] 0.006 0.015 0.009 NA 0.03 0.045
```

```

$ TC02      : num [1:6] 1877 2016 2004 1970 1905 ...
$ TC02_SE   : num [1:6] 0.6 10.4 7.2 14.7 13.5 18.7
$ TA_meas   : num [1:6] 1980 2154 2204 2185 2066 ...
$ TA_meas_SE : num [1:6] 3.7 1.5 NA NA 3 2
$ pCO2      : num [1:6] 621 508 444 344 615 ...
$ pCO2_SE   : num [1:6] 9.9 24.4 10.8 NA 47.4 30.4
$ arag.sat   : num [1:6] 1.84 2.01 1.7 2.02 1.56 3.08
$ arag.sat_SE: num [1:6] 0.03 0.09 0.04 NA 0.1 0.17
- attr(*, "groups")= tibble [3 x 2] (S3: tbl_df/tbl/data.frame)
..$ Month: Factor w/ 3 levels "Sept","April",...: 1 2 3
..$ .rows: list<int> [1:3]
.. ..$ : int [1:2] 1 2
.. ..$ : int [1:2] 3 4
.. ..$ : int [1:2] 5 6
.. ..@ ptype: int(0)
..- attr(*, ".drop")= logi TRUE

```

```

kbl(df_summary[,2:16], col.names = c(
  "Site",
  "Mean",
  "SE",
  "Mean",
  "SE",
  "Mean",
  "SE",
  "Mean",
  "SE",
  "Mean",
  "SE",
  "Mean",
  "SE",
  "Mean",
  "SE")
) %>%
kable_minimal() %>%
add_header_above(c(" " = 1,
  "Temp" = 2,
  "Sal" = 2,
  "pH" = 2,
  "TC02" = 2,
  "TA" = 2,

```

Site	Temp		Sal		pH		TCO2		TA		pCO2		$\Omega$ ar	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
<b>September</b>														
Eel	22.9	0.0	33.2	NA	7.830	0.006	1877.0	0.6	1980.3	3.7	621.3	9.9	1.84	0.03
Prov	17.4	0.2	33.4	NA	7.935	0.015	2016.5	10.4	2154.3	1.5	508.3	24.4	2.01	0.09
<b>April</b>														
Eel	11.9	0.2	31.0	NA	7.983	0.009	2004.2	7.2	2204.2	NA	443.7	10.8	1.70	0.04
Prov	11.8	0.3	31.7	NA	8.070	NA	1970.1	14.7	2184.8	NA	343.9	NA	2.02	NA
<b>June</b>														
Eel	18.7	0.1	31.6	NA	7.839	0.030	1904.9	13.5	2065.8	3.0	614.8	47.4	1.56	0.10
Prov	17.8	0.3	31.1	NA	8.193	0.045	1814.2	18.7	2067.9	2.0	244.8	30.4	3.08	0.17

```

                                "pCO2" = 2,
                                "$\\Omega$ ar" = 2)) %>%
kable_paper( full_width = F) %>%
kableExtra::pack_rows("September", 1, 2) %>%
pack_rows("April", 3, 4)%>%
pack_rows("June", 5, 6)%>%
kable_styling(latex_options="scale_down")

write.csv(df_summary, paste(wd_Biodep,"Output/SW_cond_averages.csv", sep = ""))

```