

Meta-Analysis

Laura Schaerer

2024-06-11

Meta-Analysis for Permafrost Review:

```
#load required R packages
library(tidyverse)
```

```
## — Attaching core tidyverse packages — tidyverse 2.0.0 —
## ✓ dplyr      1.1.4      ✓ readr      2.1.5
## ✓ forcats    1.0.0      ✓ stringr    1.5.1
## ✓ ggplot2    3.5.1      ✓ tibble     3.2.1
## ✓ lubridate  1.9.3      ✓ tidyr      1.3.1
## ✓ purrr      1.0.2
## — Conflicts — tidyverse_conflicts() —
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(metafor)
```

```
## Loading required package: Matrix
##
## Attaching package: 'Matrix'
##
## The following objects are masked from 'package:tidyr':
##
##     expand, pack, unpack
##
## Loading required package: metadat
## Loading required package: numDeriv
##
## Loading the 'metafor' package (version 4.6-0). For an
## introduction to the package please type: help(metafor)
```

```
library(Matrix)
```

For all analyses obtained mean and standard error from figures using metaDigitise and compiled into a single spreadsheet for each analysis.

CO2 emissions from Active Layer vs Permafrost by temperature

Including studies that incubated Active Layer and Permafrost samples at one or more temperatures and measured CO2 emissions.

Studies included from review: Doherty, Dutta, Gentsch, Treat, Roy Chowdhury, Bracho, Jiang, Song 2023, Moni, Muller, Mackelprang 2011.

```
#load data
co2_flux_by_layer_temp <- read_csv("/Users/lgschaer/Desktop/MetaA_Figs/co2_flux_by_layer_temp.csv")
```

```
## Rows: 247 Columns: 9
## — Column specification —————
## Delimiter: ","
## chr (4): Author, variable, group_id, Layer
## dbl (5): Temperature, mean, sd, n, se
##
## 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.
```

```
#wrap data into necessary format
metaA <- co2_flux_by_layer_temp %>%
  group_by(Author, Layer, Temperature) %>%
  summarize(
    mean = mean(mean),
    sd = sqrt(sum(((n-1)*(sd^2)), na.rm = TRUE)/(sum(n)-n())),
    n = sum(n)) %>%
  mutate(Layer1 = Layer,
    Layer2 = Layer) %>%
  ungroup() %>%
  pivot_wider(names_from = "Layer", values_from = "mean", names_prefix = "mean_") %>%
  pivot_wider(names_from = "Layer1", values_from = "sd", names_prefix = "sd_") %>%
  pivot_wider(names_from = "Layer2", values_from = "n", names_prefix = "n_") %>%
  group_by(Author, Temperature) %>%
  reframe( #this works because there is only one observation for each group, doing this
to get rid of NAs, not actually adding or subtracting anything here!
    mean_Active_Layer = sum(mean_Active_Layer, na.rm = TRUE),
    mean_Permafrost = sum(mean_Permafrost, na.rm = TRUE),
    sd_Active_Layer = sum(sd_Active_Layer, na.rm = TRUE),
    sd_Permafrost = sum(sd_Permafrost, na.rm = TRUE),
    n_Active_Layer = sum(n_Active_Layer, na.rm = TRUE),
    n_Permafrost = sum(n_Permafrost, na.rm = TRUE)
  )
```

```
## `summarise()` has grouped output by 'Author', 'Layer'. You can override using
## the `.groups` argument.
```

```
#use escalc function to calculate effect sizes
##NOTE: I also did this using measure = "SMDH" to account for heteroskedasticity in the
data, results were exactly the same
effects <- escalc(measure = "SMD", m1i = mean_Active_Layer, m2i = mean_Permafrost,
  sd1i = sd_Active_Layer, sd2i = sd_Permafrost,
  n1i = n_Active_Layer, n2i = n_Permafrost, data = metaA)
head(effects)
```

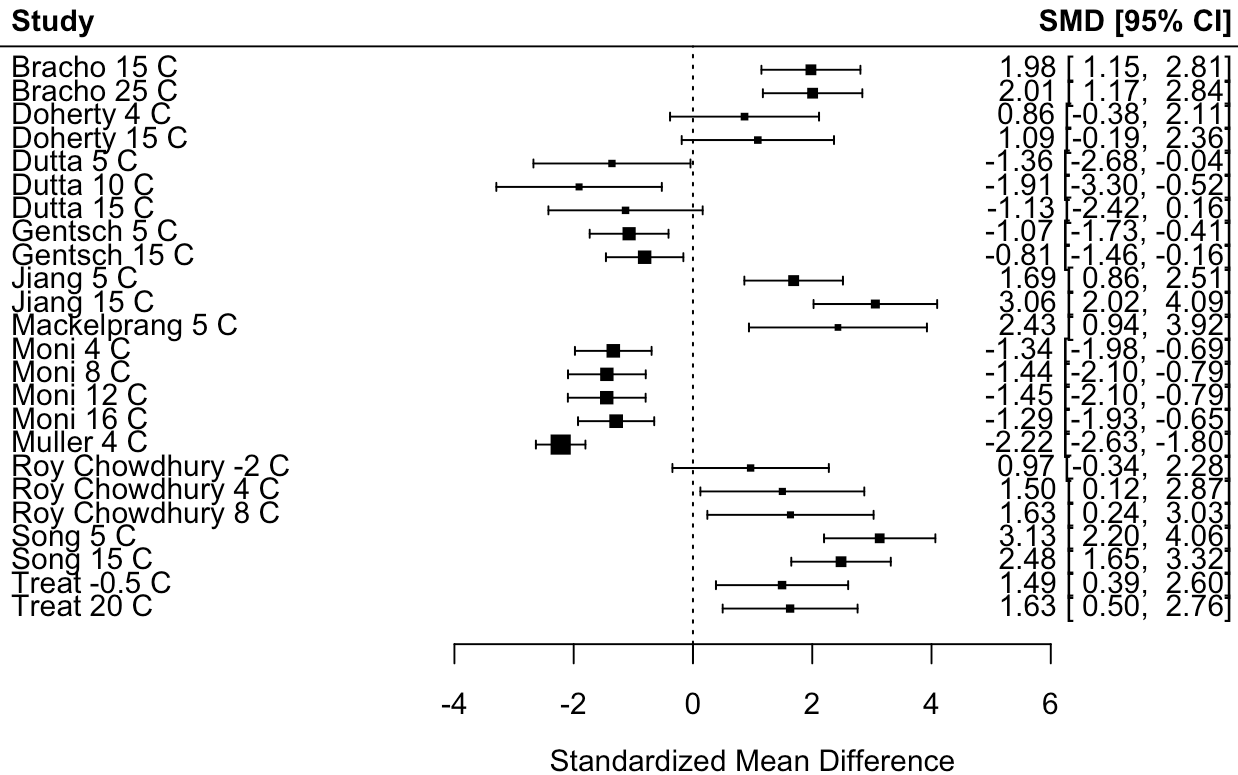
Author <chr>	Temperature <dbl>	mean_Active_Layer <dbl>	mean_Permafrost <dbl>	sd_Active_Layer <dbl>	sd_Permafrost <dbl>
1 Bracho	15	88.84943182	21.3068182	40.18355193	8.35053327
2 Bracho	25	160.86647727	49.8579545	64.74188111	16.83103777
3 Doherty	4	1.11809211	0.3355263	0.97676488	0.31211597
4 Doherty	15	2.35953947	0.6690789	1.65571167	0.67439125
5 Dutta	5	0.05876543	0.1189136	0.01988930	0.04444127
6 Dutta	10	0.09661202	0.1758689	0.01878837	0.04164862

6 rows | 1-7 of 11 columns

SMD = 0 ~ There is no difference in CO2 emissions between Active Layer and Permafrost treatments

SMD < 0 ~ Mean of CO2 emissions from Active Layer is lower than mean of CO2 emissions from Permafrost

SMD > 0 ~ Mean of CO2 emissions from Permafrost is lower than mean of CO2 emissions from Active Layer



The figure shows variance in the responses, some studies/temperatures have higher CO₂ fluxes from the active layer and other studies/temperatures have higher CO₂ fluxes from the permafrost layer, there are no visual patterns.

```
#fit mixed-effects model with temperature as a moderator
### NOTE: super high p-value, seems temp doesn't have a statistically significant effect here
res <- rma(yi, vi, mods = ~ Temperature, slab=Author, data=effects)
res
```

```
##
## Mixed-Effects Model (k = 24; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):      2.9153 (SE = 0.9614)
## tau (square root of estimated tau^2 value):            1.7074
## I^2 (residual heterogeneity / unaccounted variability): 93.88%
## H^2 (unaccounted variability / sampling variability):   16.35
## R^2 (amount of heterogeneity accounted for):            0.00%
##
## Test for Residual Heterogeneity:
## QE(df = 22) = 376.9968, p-val < .0001
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 0.3487, p-val = 0.5549
##
## Model Results:
##
##              estimate      se    zval    pval    ci.lb    ci.ub
## intrcpt          0.1592  0.6423  0.2479  0.8042  -1.0996  1.4181
## Temperature      0.0327  0.0554  0.5905  0.5549  -0.0758  0.1412
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The model fit to temperature is insignificant, suggesting that temperature does not impact fluxes in a consistent way and there are likely other controlling factors that are not accounted for here.

CO2 emissions from Active Layer vs Permafrost by incubation time

Including Dutta, Jiang, Moni, Treat, Waldrop 2010, and Roy Chowdhury.

```
#load data
co2_flux_by_layer_time <- read_csv("/Users/lgschaer/Desktop/MetaA_Figs/co2_flux_by_layer_time.csv")
```

```
## Rows: 143 Columns: 8
## — Column specification —————
## Delimiter: ","
## chr (4): Article, Layer, variable, group_id
## dbl (4): Days, mean, sd, n
##
## 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.
```

```
#wrangle data into necessary format
metaA <- co2_flux_by_layer_time %>%
  group_by(Article, Layer, Days) %>%
  summarize(
    mean = mean(mean, na.rm = TRUE),
    sd = sqrt(sum(((n-1)*(sd^2)), na.rm = TRUE)/(sum(n)-n())),
    n = sum(n, na.rm = TRUE)) %>%
  mutate(Layer1 = Layer,
         Layer2 = Layer) %>%
  ungroup() %>%
  pivot_wider(names_from = "Layer", values_from = "mean", names_prefix = "mean_") %>%
  pivot_wider(names_from = "Layer1", values_from = "sd", names_prefix = "sd_") %>%
  pivot_wider(names_from = "Layer2", values_from = "n", names_prefix = "n_") %>%
  group_by(Article, Days) %>%
  reframe(
    mean_ActiveLayer = mean(mean_ActiveLayer, na.rm = TRUE),
    n_ActiveLayer = sum(n_ActiveLayer, na.rm = TRUE),
    sd_ActiveLayer = sqrt(sum(((n_ActiveLayer-1)*(sd_ActiveLayer^2)), na.rm = TRUE)/(sum(
n_ActiveLayer)-n()))),
    mean_Permafrost = mean(mean_Permafrost, na.rm = TRUE),
    n_Permafrost = sum(n_Permafrost, na.rm = TRUE),
    sd_Permafrost = sqrt(sum(((n_Permafrost-1)*(sd_Permafrost^2)), na.rm = TRUE)/(sum(n_
Permafrost)-n()))
  )
```

```
## `summarise()` has grouped output by 'Article', 'Layer'. You can override using
## the `.groups` argument.
```

```
#use escalc function to calculate effect sizes
##NOTE: I also did this using measure = "SMDH" to account for heteroskedasticity in the
data, results were exactly the same
effects <- escalc(measure = "SMD", m1i = mean_ActiveLayer, m2i = mean_Permafrost,
                  sd1i = sd_ActiveLayer, sd2i = sd_Permafrost,
                  n1i = n_ActiveLayer, n2i = n_Permafrost, data = metaA)
head(effects)
```

Article <chr>	D... <dbl>	mean_ActiveLayer <dbl>	n_ActiveLayer <dbl>	sd_ActiveLayer <dbl>	mean_Permafrost <dbl>	n_Perma <
1 Dutta	30	0.09263968	18	0	0.1844579	
2 Dutta	40	0.14148321	3	0	0.2972987	
3 Dutta	45	0.08859951	15	0	0.1670475	
4 Dutta	50	0.17013894	3	0	0.2811799	
5 Dutta	55	0.10306306	24	0	0.1962735	
6 Dutta	60	0.20595860	3	0	0.3169995	

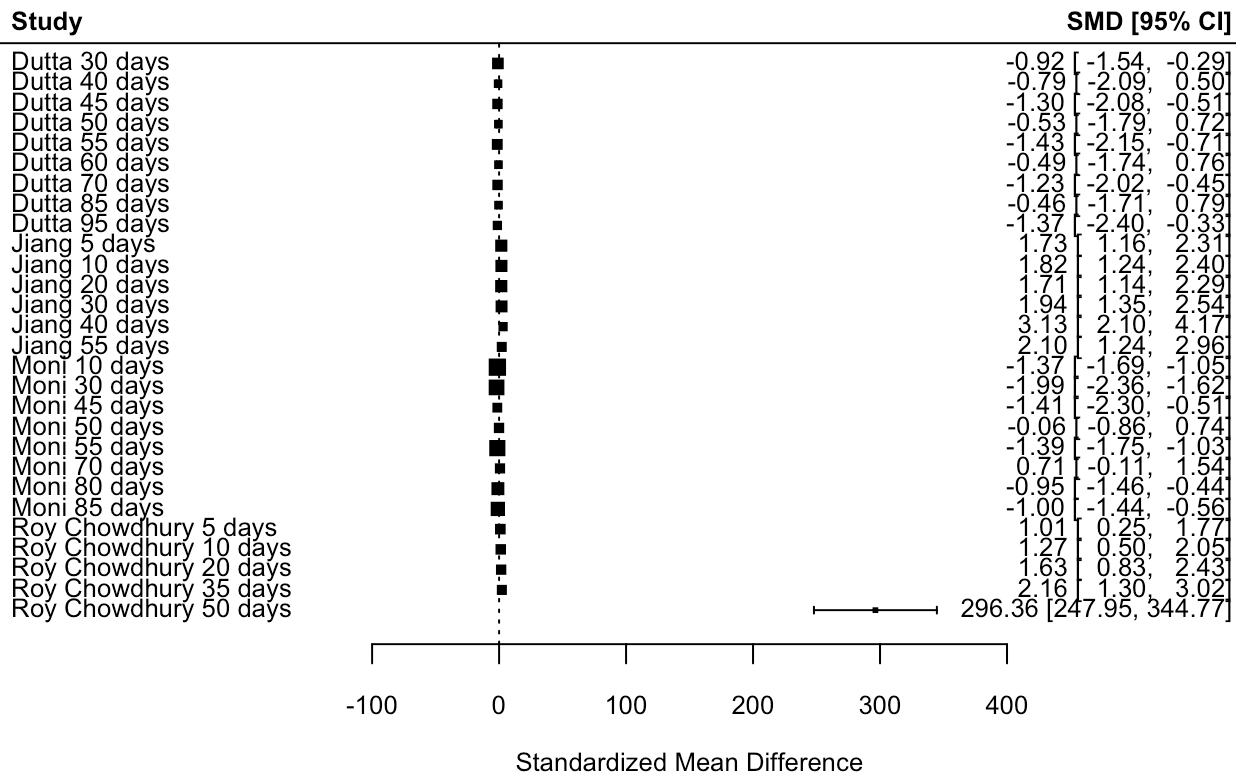
6 rows | 1-8 of 11 columns

SMD = 0 ~ There is no difference in CO2 emissions between Active Layer and Permafrost treatments

SMD < 0 ~ Mean of CO2 emissions from Active Layer is lower than mean of CO2 emissions from Permafrost

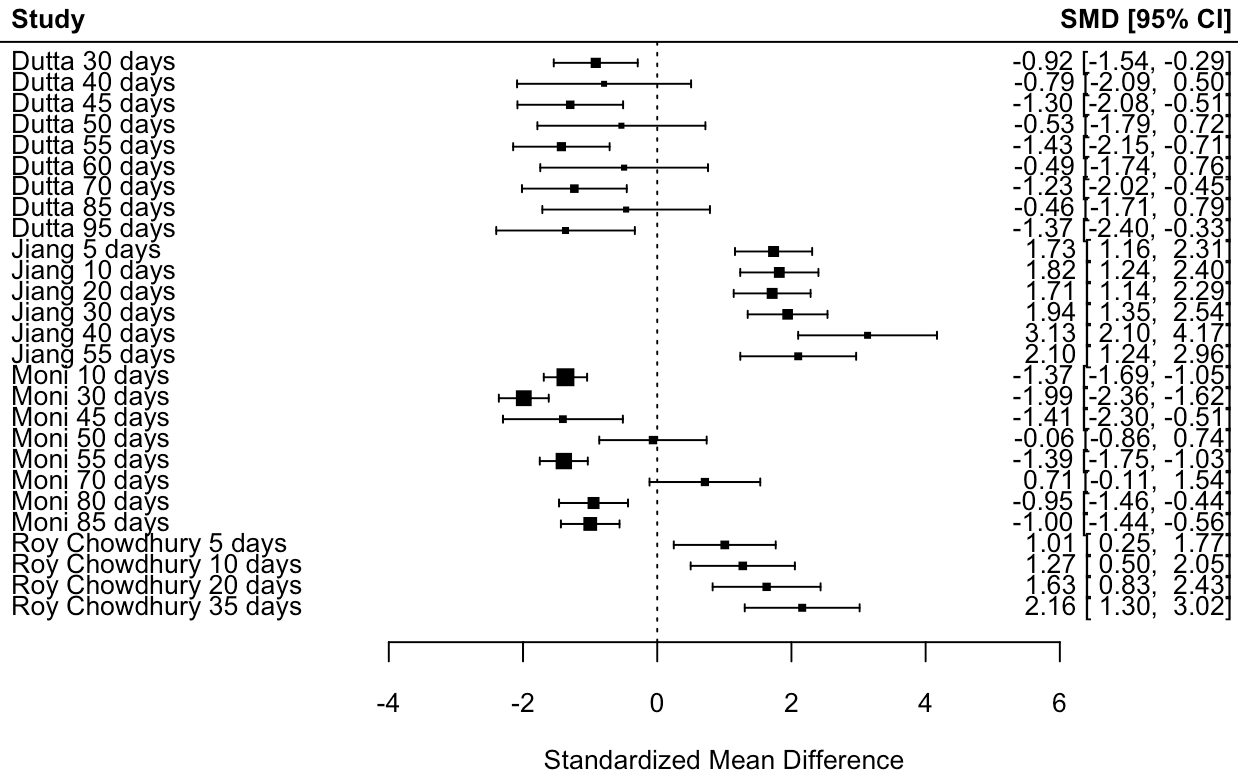
SMD > 0 ~ Mean of CO2 emissions from Permafrost is lower than mean of CO2 emissions from Active Layer

```
forest(effects$yi, effects$vi, slab = paste0(effects$Article, " ", effects$Days, " days"), header=TRUE, top=2)
```



Since there is an outlier, visualize with those points removed

```
effects2 <- effects %>% filter(Article != "Roy Chowdhury" | Days != 50)
forest(effects2$yi, effects2$vi, slab = paste0(effects2$Article, " ", effects2$Days, " days"), header=TRUE, top=2)
```



Based on the figure, the effect size seems to vary more based on each study than by incubation time.

Check for statistically significant relationships with time.

```
res1 <- rma(yi, vi, mods = ~ Days, slab=Article, data=effects)
res1
```



```
##
## Mixed-Effects Model (k = 28; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):      1932.4312 (SE = 540.3028)
## tau (square root of estimated tau^2 value):             43.9594
## I^2 (residual heterogeneity / unaccounted variability): 99.99%
## H^2 (unaccounted variability / sampling variability):    17262.14
## R^2 (amount of heterogeneity accounted for):             0.00%
##
## Test for Residual Heterogeneity:
## QE(df = 26) = 658.2431, p-val < .0001
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 0.0226, p-val = 0.8806
##
## Model Results:
##
##           estimate      se    zval    pval    ci.lb    ci.ub
## intrcpt      6.0196  16.7811  0.3587  0.7198  -26.8708  38.9100
## Days          0.0496   0.3305  0.1502  0.8806   -0.5981   0.6974
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
res2 <- rma(yi, vi, mods = ~ Days, slab=Article, data=effects2)
res2 #significant when the outlier is removed
```

```
##
## Mixed-Effects Model (k = 27; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):      1.7342 (SE = 0.5377)
## tau (square root of estimated tau^2 value):             1.3169
## I^2 (residual heterogeneity / unaccounted variability): 94.16%
## H^2 (unaccounted variability / sampling variability):    17.11
## R^2 (amount of heterogeneity accounted for):             17.91%
##
## Test for Residual Heterogeneity:
## QE(df = 25) = 513.7678, p-val < .0001
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 6.2918, p-val = 0.0121
##
## Model Results:
##
##           estimate      se    zval    pval    ci.lb    ci.ub
## intrcpt      1.2173   0.5216  2.3337  0.0196   0.1949   2.2397  *
## Days        -0.0259   0.0103 -2.5083  0.0121  -0.0462  -0.0057  *
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The relationship with time is statistically significant if the outlier points (from the last time point in Roy Chowdhury et al) are removed.

CO2 emissions from Aerobic vs Anaerobic incubations

Including studies that incubated samples under both aerobic and anaerobic conditions and measured CO2 emissions.

Studies included: Capek 2015, Ernakovich 2017, Kane 2013, Knoblauch 2013, Knoblauch 2018, and Lee 2012

Obtained mean and standard error from figures using metaDigitise and compiled into a single spreadsheet.

```
#load data
co2_flux_by_headspace_temp <- read_csv("/Users/lgschaer/Desktop/MetaA_Figs/co2_flux_by_headspace_temp.csv")
```

```
## Rows: 64 Columns: 9
## — Column specification —————
## Delimiter: ","
## chr (4): Article, variable, group_id, Headspace
## dbl (5): Temperature, mean, sd, n, se
##
## 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.
```

```

#wrap data into necessary format
metaA <- co2_flux_by_headspace_temp %>%
  group_by(Article, Headspace, Temperature) %>%
  summarize(
    mean = mean(mean, na.rm = TRUE),
    n = sum(n, na.rm = TRUE),
    sd = sqrt(sum(((n-1)*(sd^2)), na.rm = TRUE)/(sum(n)-n())) %>%
  mutate(Headspace1 = Headspace,
    Headspace2 = Headspace,
    manual_sd = case_when(Article == "Ernakovich 2017" & Headspace == "Aerobic" & Temperature == 1 ~ 121.9947,
      Article == "Ernakovich 2017" & Headspace == "Aerobic" & Temperature == 15 ~ 319.2053,
      Article == "Ernakovich 2017" & Headspace == "Anaerobic" & Temperature == 1 ~ 61.6669,
      Article == "Ernakovich 2017" & Headspace == "Anaerobic" & Temperature == 15 ~ 169.8947,
      Article == "Knoblauch 2018" & Headspace == "Aerobic" ~ 61.3674,
      Article == "Knoblauch 2018" & Headspace == "Anaerobic" ~ 11.6623),
    sd = ifelse(is.na(sd), manual_sd, sd)
  ) %>%
  ungroup() %>%
  pivot_wider(names_from = "Headspace", values_from = "mean", names_prefix = "mean_") %>%
  pivot_wider(names_from = "Headspace1", values_from = "sd", names_prefix = "sd_") %>%
  pivot_wider(names_from = "Headspace2", values_from = "n", names_prefix = "n_") %>%
  group_by(Article, Temperature) %>%
  reframe( #this works because there is only one observation for each group, doing this to get rid of NAs, not actually adding anything here!
    mean_Aerobic = sum(mean_Aerobic, na.rm = TRUE),
    mean_Anaerobic = sum(mean_Anaerobic, na.rm = TRUE),
    sd_Aerobic = sum(sd_Aerobic, na.rm = TRUE),
    sd_Anaerobic = sum(sd_Anaerobic, na.rm = TRUE),
    n_Aerobic = sum(n_Aerobic, na.rm = TRUE),
    n_Anaerobic = sum(n_Anaerobic, na.rm = TRUE)
  )

```

```

## `summarise()` has grouped output by 'Article', 'Headspace'. You can override
## using the `.groups` argument.

```

```

#use escalc function to calculate effect sizes
##NOTE: I also did this using measure = "SMDH" to account for heteroskedasticity in the
data, results were exactly the same
effects <- escalc(measure = "SMD", m1i = mean_Aerobic, m2i = mean_Anaerobic,
  sd1i = sd_Aerobic, sd2i = sd_Anaerobic,
  n1i = n_Aerobic, n2i = n_Anaerobic, data = metaA)
head(effects)

```

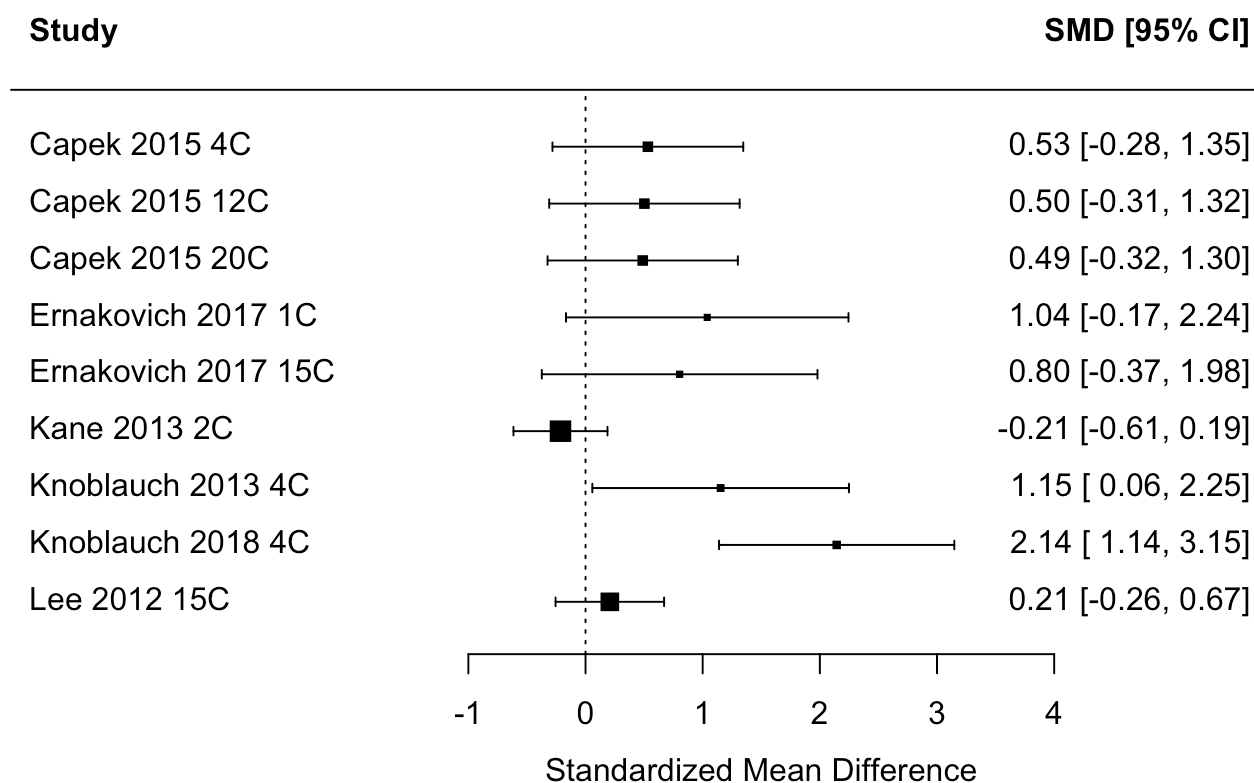
Article <chr>	Temperature <dbl>	mean_Aerobic <dbl>	mean_Anaerobic <dbl>	sd_Aerobic <dbl>	sd_Anaerobic <dbl>	n_
1 Capek 2015	4	1.026790	0.4868459	1.323231	0.4108475	
2 Capek 2015	12	1.926066	1.0045481	2.268351	1.0564518	
3 Capek 2015	20	2.876592	2.0648874	1.730524	1.4700778	
4 Ernakovich 2017	1	228.767681	119.9487961	121.994678	61.6669061	
5 Ernakovich 2017	15	533.013877	310.4134955	319.205314	169.8947265	
6 Kane 2013	2	7.970305	9.9845116	10.964455	7.3895414	

6 rows | 1-8 of 11 columns

SMD = 0 ~ There is no difference in CO2 emissions between Aerobic and Anaerobic treatments

SMD < 0 ~ Mean CO2 emissions from Aerobic Treatments are lower than mean of CO2 emissions from Anaerobic Treatments

SMD > 0 ~ Mean CO2 emissions from Anaerobic Treatments are lower than mean of CO2 emissions from Aerobic Treatments



Most studies have positive effect sizes for all temperatures (except for Kane et al 2013), suggesting that higher CO2 emissions come from Aerobic treatments on average.

```
#fit mixed-effects model with temperature as a moderator
### NOTE: super high p-value, seems temp doesn't have a statistically significant effect here
res <- rma(yi, vi, mods = ~ Temperature, slab=Article, data=effects)
res
```

```
##
## Mixed-Effects Model (k = 9; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):      0.3593 (SE = 0.2932)
## tau (square root of estimated tau^2 value):            0.5994
## I^2 (residual heterogeneity / unaccounted variability): 68.91%
## H^2 (unaccounted variability / sampling variability):   3.22
## R^2 (amount of heterogeneity accounted for):            0.00%
##
## Test for Residual Heterogeneity:
## QE(df = 7) = 24.2317, p-val = 0.0010
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 0.2416, p-val = 0.6230
##
## Model Results:
##
##              estimate      se      zval      pval      ci.lb      ci.ub
## intrcpt          0.8031  0.4103   1.9572  0.0503   -0.0011   1.6073
## Temperature     -0.0184  0.0374  -0.4916  0.6230   -0.0918   0.0550
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Temperature has a statistically non-significant relationship with effect sizes.

CH4 emissions from Active Layer vs Permafrost incubations by temperature

```
#load data
ch4_flux_by_layer_temp <- read_csv("/Users/lgschaer/Desktop/MetaA_Figs/ch4_flux_by_layer_temp.csv")
```

```
## Rows: 58 Columns: 9
## — Column specification —————
## Delimiter: ","
## chr (4): Article, Layer, variable, group_id
## dbl (5): Time, Temperature, mean, n, sd
##
## 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.
```

```
#wrap data into necessary format
metaA <- ch4_flux_by_layer_temp %>%
  mutate(Layer2 = Layer,
         Layer3 = Layer) %>%
  pivot_wider(names_from = "Layer", values_from = "mean", names_prefix = "mean_") %>%
  pivot_wider(names_from = "Layer2", values_from = "sd", names_prefix = "sd_") %>%
  pivot_wider(names_from = "Layer3", values_from = "n", names_prefix = "n_") %>%
  group_by(Article, Temperature) %>%
  summarize(
    mean_ActiveLayer = mean(mean_ActiveLayer, na.rm = TRUE),
    n_ActiveLayer = sum(n_ActiveLayer, na.rm = TRUE),
    sd_ActiveLayer = sqrt(sum(((n_ActiveLayer-1)*(sd_ActiveLayer^2))), na.rm = TRUE)/(sum(
    (n_ActiveLayer)-n()))),
    mean_Permafrost = mean(mean_Permafrost, na.rm = TRUE),
    n_Permafrost = sum(n_Permafrost, na.rm = TRUE),
    sd_Permafrost = sqrt(sum(((n_Permafrost-1)*(sd_Permafrost^2))), na.rm = TRUE)/(sum(n_
    Permafrost)-n()))
  )
```

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

```
#use escalc function to calculate effect sizes
##NOTE: I also did this using measure = "SMDH" to account for heteroskedasticity in the
data, results were exactly the same
effects <- escalc(measure = "SMD", m1i = mean_ActiveLayer, m2i = mean_Permafrost,
                  sd1i = sd_ActiveLayer, sd2i = sd_Permafrost,
                  n1i = n_ActiveLayer, n2i = n_Permafrost, data = metaA) %>%
  as_tibble()
head(effects)
```

	n_ActiveLayer <dbl>	sd_ActiveLayer <dbl>	mean_Permafrost <dbl>	n_Permafrost <dbl>	sd_Permafrost <dbl>	yi <dbl>
	156	2.936734	0.16538713	156	0.255714376	0.4355245
	160	73.924913	12.68778280	160	30.924987635	0.2721476
	6	6.333805	9.84848485	6	14.960922468	-0.6754423
	156	1.148269	0.01223802	44	0.005320722	0.6281277
	162	20.341541	0.01772757	44	0.006089705	0.4406429
	168	29.215939	0.01945983	40	0.007981429	0.4549234

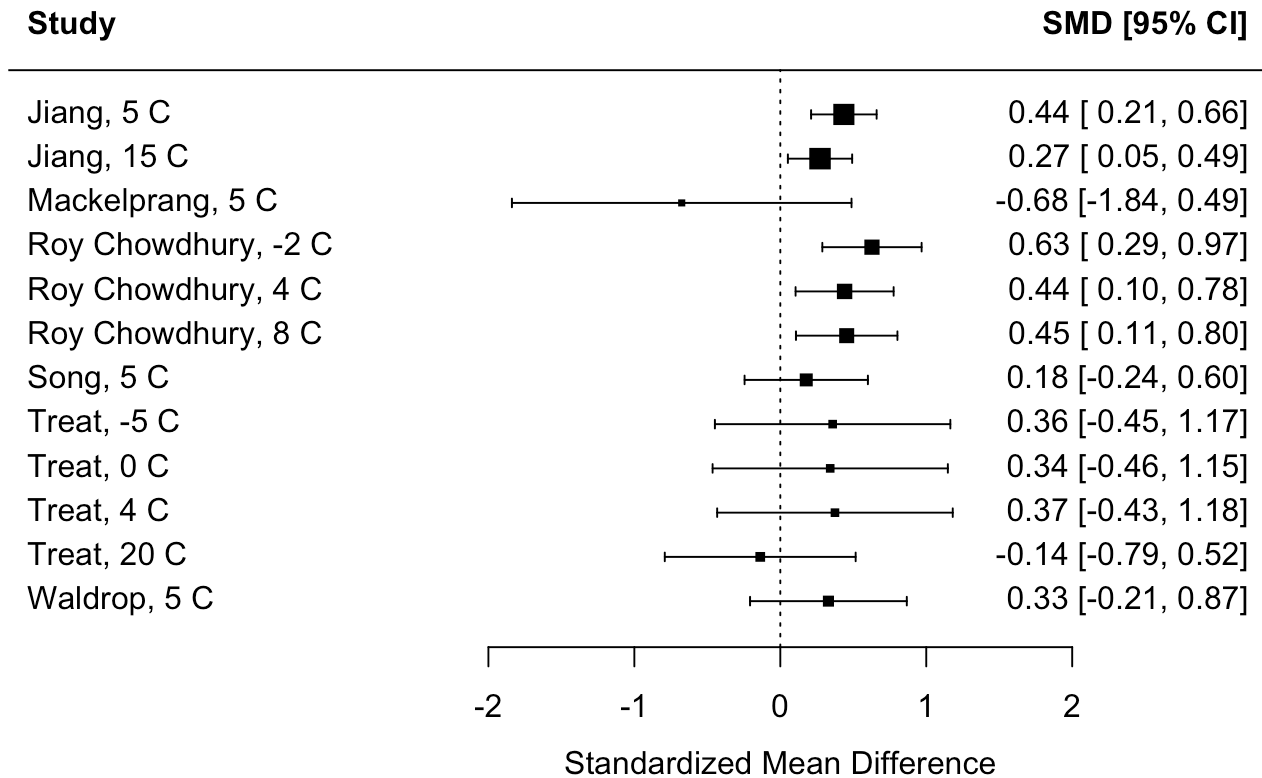
6 rows | 4-10 of 10 columns

SMD = 0 ~ There is no difference in CH₄ emissions between Active Layer and Permafrost treatments

SMD < 0 ~ Mean of CH₄ emissions from Active Layer is lower than mean of CH₄ emissions from Permafrost

SMD > 0 ~ Mean of CH₄ emissions from Permafrost is lower than mean of CH₄ emissions from Active Layer

```
#plot it
forest(effects$yi, effects$vi, slab = paste0(effects$Article, ", ", effects$Temperature,
" C"), header=TRUE, top=2)
```



Most effect sizes are positive (excluding Mackelprang and Treat), suggesting that higher methane emissions are emitted from the active layer compared to the permafrost layer.

```
#fit mixed-effects model with temperature as a moderator, super high p-value, seems time
doesn't have a big effect here
res_temp <- rma(yi, vi, mods = ~ Temperature, slab=Article, data=effects)
res_temp
```

```
##
## Mixed-Effects Model (k = 12; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):      0.0000 (SE = 0.0156)
## tau (square root of estimated tau^2 value):            0.0018
## I^2 (residual heterogeneity / unaccounted variability): 0.01%
## H^2 (unaccounted variability / sampling variability):    1.00
## R^2 (amount of heterogeneity accounted for):            0.00%
##
## Test for Residual Heterogeneity:
## QE(df = 10) = 6.5299, p-val = 0.7690
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 3.3980, p-val = 0.0653
##
## Model Results:
##
##              estimate      se      zval      pval      ci.lb      ci.ub
## intrcpt          0.4882  0.0871   5.6071 <.0001    0.3176  0.6589 ***
## Temperature     -0.0174  0.0094  -1.8434  0.0653   -0.0358  0.0011 .
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

There does not appear to be a statistically significant relationship between temperature and effect size.

CH4 emissions from Active Layer vs Permafrost incubations by time

```
#load data
ch4_flux_by_layer_time <- read_csv("/Users/lgschaer/Desktop/MetaA_Figs/ch4_flux_by_layer_time.csv")
```

```
## Rows: 73 Columns: 8
## — Column specification —————
## Delimiter: ","
## chr (4): Article, Layer, variable, group_id
## dbl (4): Time, mean, n, sd
##
## 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.
```



```
#wrap data into necessary format
metaA <- ch4_flux_by_layer_time %>%
  mutate(Layer2 = Layer,
         Layer3 = Layer) %>%
  pivot_wider(names_from = "Layer", values_from = "mean", names_prefix = "mean_") %>%
  pivot_wider(names_from = "Layer2", values_from = "sd", names_prefix = "sd_") %>%
  pivot_wider(names_from = "Layer3", values_from = "n", names_prefix = "n_") %>%
  group_by(Article, Time) %>%
  summarize(
    mean_ActiveLayer = mean(mean_ActiveLayer, na.rm = TRUE),
    n_ActiveLayer = sum(n_ActiveLayer, na.rm = TRUE),
    sd_ActiveLayer = sqrt(sum(((n_ActiveLayer-1)*(sd_ActiveLayer^2)), na.rm = TRUE)/(sum(
n_ActiveLayer)-n()))),
    mean_Permafrost = mean(mean_Permafrost, na.rm = TRUE),
    n_Permafrost = sum(n_Permafrost, na.rm = TRUE),
    sd_Permafrost = sqrt(sum(((n_Permafrost-1)*(sd_Permafrost^2)), na.rm = TRUE)/(sum(n_
Permafrost)-n()))
  )
```

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

```
#use escalc function to calculate effect sizes
##NOTE: I also did this using measure = "SMDH" to account for heteroskedasticity in the
data, results were exactly the same
effects <- escalc(measure = "SMD", m1i = mean_ActiveLayer, m2i = mean_Permafrost,
                 sd1i = sd_ActiveLayer, sd2i = sd_Permafrost,
                 n1i = n_ActiveLayer, n2i = n_Permafrost, data = metaA) %>%
  as_tibble()
head(effects)
```

Article <chr>	T... <dbl>	mean_ActiveLayer <dbl>	n_ActiveLayer <dbl>	sd_ActiveLayer <dbl>	mean_Permafrost <dbl>	n_
Jiang	5	2.901759	32	1.420574	3.04428619	
Jiang	12	2.900348	32	3.880380	2.97673568	
Jiang	35	11.142045	32	19.803746	5.35886390	
Jiang	55	26.070805	32	67.220720	18.34283255	
Roy Chowdhury	5	1.963542	45	5.375175	0.01079637	
Roy Chowdhury	12	5.746669	45	14.769774	0.01977823	

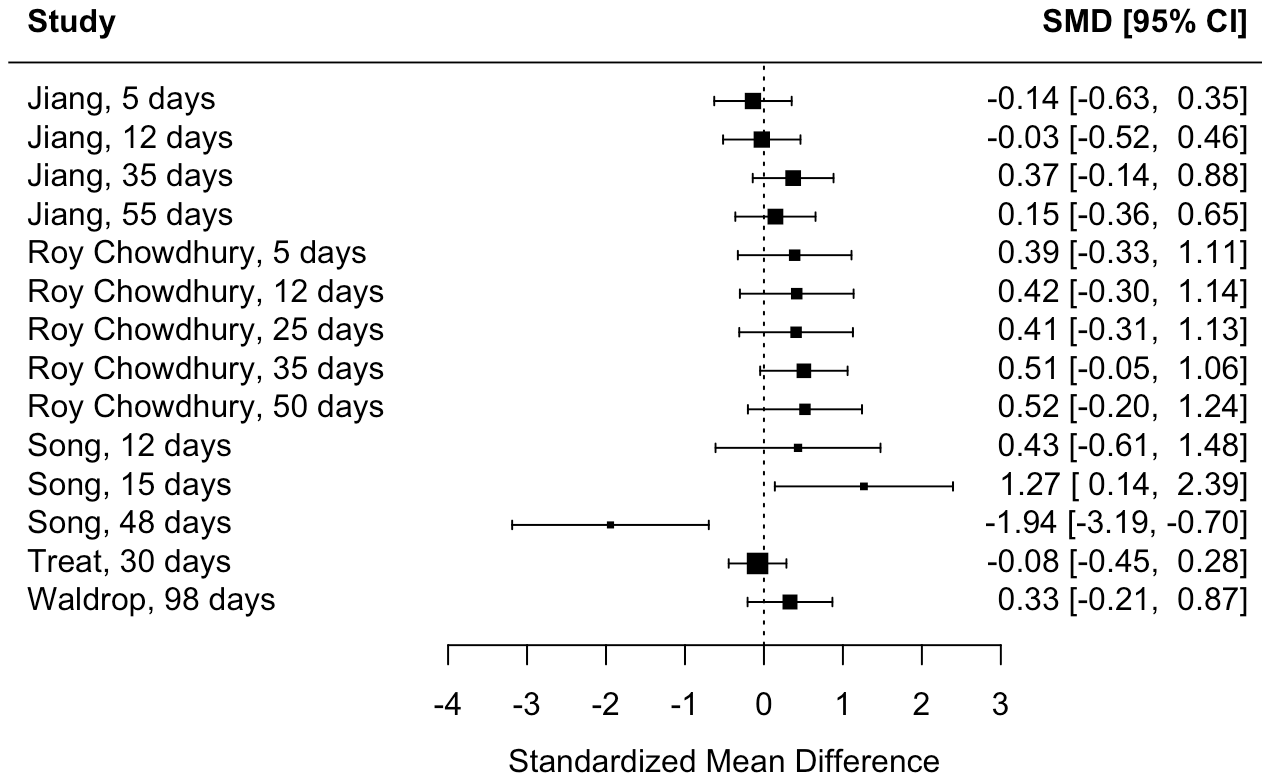
6 rows | 1-7 of 10 columns

SMD = 0 ~ There is no difference in CH₄ emissions between Active Layer and Permafrost treatments

SMD < 0 ~ Mean of CH₄ emissions from Active Layer is lower than mean of CH₄ emissions from Permafrost

SMD > 0 ~ Mean of CH₄ emissions from Permafrost is lower than mean of CH₄ emissions from Active Layer

```
#plot it
forest(effects$yi, effects$vi, slab = paste0(effects$Article, ", ", effects$Time, " day
s"), header=TRUE, top=2)
```



Again, most effect sizes are positive (excluding Song 2014 et al after 48 days), suggesting that higher methane emissions are emitted from the active layer compared to the permafrost layer.

```
#fit mixed-effects model with temperature as a moderator, super high p-value, seems time
doesn't have a big effect here
res <- rma(yi, vi, mods = ~ Time, slab=Article, data=effects)
res
```

```
##
## Mixed-Effects Model (k = 14; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):      0.0278 (SE = 0.0481)
## tau (square root of estimated tau^2 value):             0.1666
## I^2 (residual heterogeneity / unaccounted variability): 22.66%
## H^2 (unaccounted variability / sampling variability):    1.29
## R^2 (amount of heterogeneity accounted for):             0.00%
##
## Test for Residual Heterogeneity:
## QE(df = 12) = 23.0995, p-val = 0.0269
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 0.0655, p-val = 0.7980
##
## Model Results:
##
##           estimate      se    zval    pval    ci.lb    ci.ub
## intrcpt      0.1597  0.1546  1.0329  0.3016  -0.1433  0.4627
## Time          0.0010  0.0037  0.2559  0.7980  -0.0063  0.0082
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

There does not appear to be a statistically significant relationship between incubation time and effect size.

NOTE: Used the formula here to estimate average standard deviations throughout:

<https://www.statology.org/averaging-standard-deviations/> (<https://www.statology.org/averaging-standard-deviations/>)