

Epibenthic Community Statistics on functional ecology from 5 mile ledge

Jared Oviatt

2023

Contents

Initialize	1
Models - take out sal and waves	13

Initialize

Load libraries

```
my_ggsave <- function(filename, plot,
                        device = 'png',
                        path = './Figs/',
                        dpi = 300,
                        units = 'in',
                        width = 8,
                        height = 8,
                        bg = 'white')
{
  ggsave(filename, plot, device = device, path = path, dpi = dpi, units = units, width = width, height = height, bg = bg)
}
```

define new functions

```
theme_set(theme_few())
#theme_set(darkmode(theme_few()))

ax_text_size <- 12
lab_text_size <- 14
```

Set gg theme

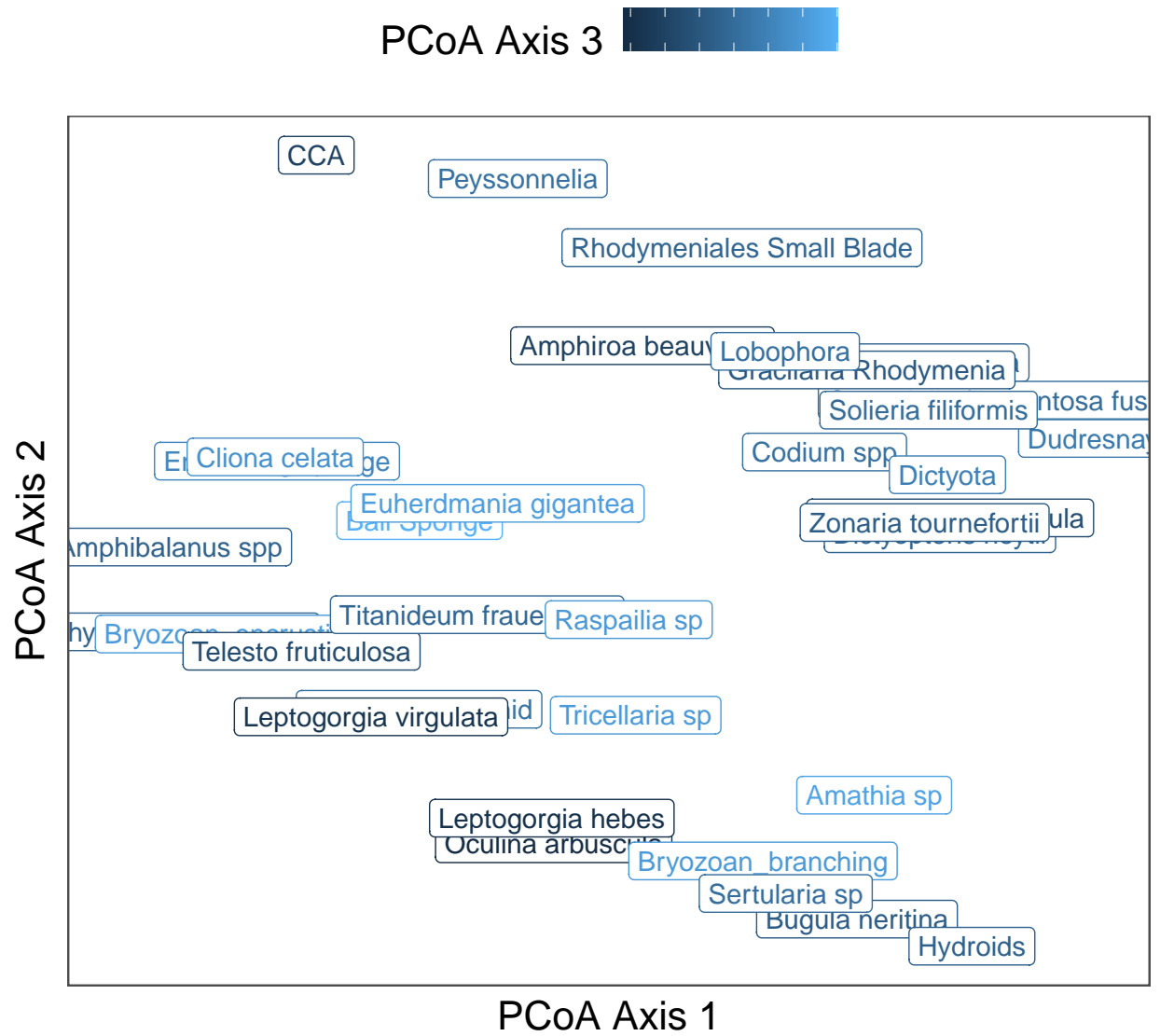
Load data

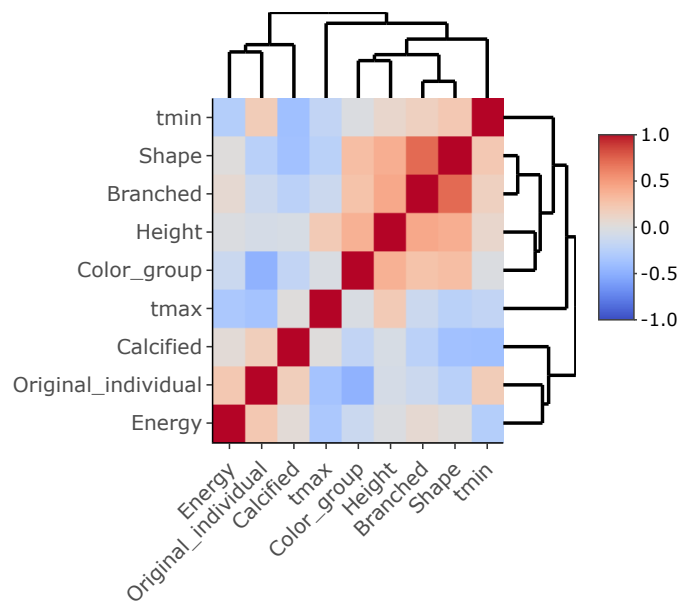
Make comm, tax, env, struc matrices

```
## Joining with 'by = join_by(Date)'  
## Joining with 'by = join_by(Date)'  
## Joining with 'by = join_by(Date)'  
## Joining with 'by = join_by(Date)'
```

Make trait space, trait comms

```
fig3 <- ggplot(trait_plot, aes(x = Axis.1, y = Axis.2, color = Axis.3)) +  
  geom_label(aes(label = rowname), show.legend = T) +  
  xlab('PCoA Axis 1') +  
  ylab('PCoA Axis 2') +  
  scale_color_continuous(name = 'PCoA Axis 3') +  
  theme(axis.text = element_blank(),  
        axis.ticks = element_blank(),  
        axis.title = element_text(size = lab_text_size + 2),  
        legend.text = element_blank(),  
        legend.title = element_text(size = lab_text_size + 2),  
        legend.position = 'top')  
fig3
```





Diversity

```
# Functional Diversity
fun_div <- dbFD(trait_space, fcomm, corr = 'cailliez', CWM.type = 'all') %>%
  as.data.frame() %>%
  rownames_to_column(var = 'rowname')
```

```
## Species x species distance matrix was not Euclidean. Cailliez correction was applied.
## FEve: Could not be calculated for communities with <3 functionally singular species.
## FDis: Equals 0 in communities with only one functionally singular species.
## FRic: To respect s > t, FRic could not be calculated for communities with <3 functionally singular species.
## FRic: Dimensionality reduction was required. The last 31 PCoA axes (out of 33 in total) were removed.
## FRic: Quality of the reduced-space representation (based on corrected distance matrix) = 0.3021053
## FDiv: Could not be calculated for communities with <3 functionally singular species.
```

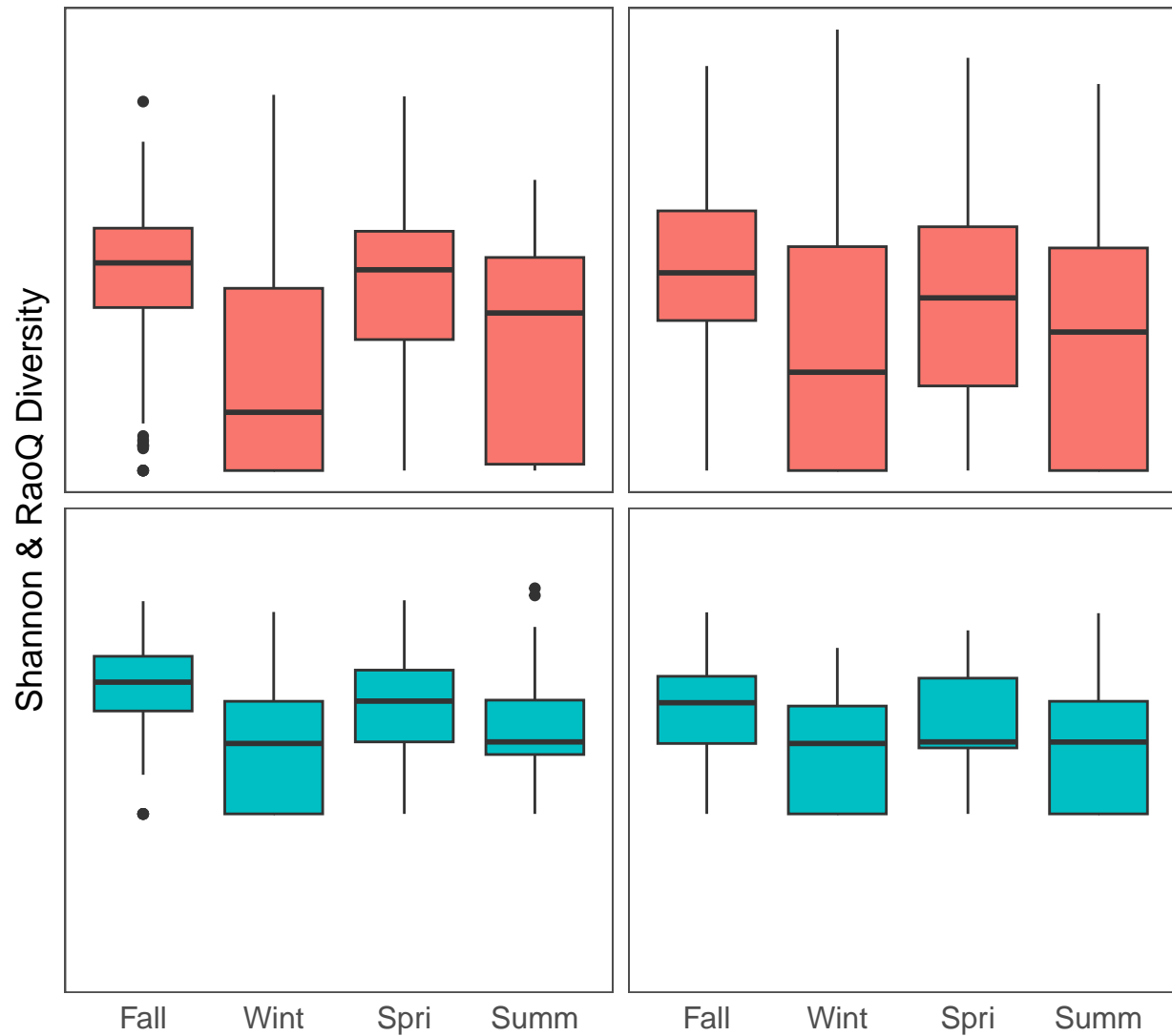
```
# Taxonomic Diversity (shannon)
tax_div <- diversity(fcomm, index = 'shannon', ) %>%
  as.data.frame() %>%
  rename('Shannon' = '.') %>%
  rownames_to_column(var = 'rowname')
```

```
# combo
comb_div <- left_join(fun_div, tax_div) %>%
  left_join(rownames_to_column(fmeta, var = 'rowname')) %>%
  select(rowname, Date, Site, Season, Quad, RaoQ, Shannon) %>%
  mutate(RaoQ = scale(RaoQ)[,1],
         Shanon = scale(Shannon)[,1]) %>%
  pivot_longer(cols = c('RaoQ', 'Shannon'), names_to = 'Index', values_to = 'Value')
```

```
## Joining with 'by = join_by(rowname)'
## Joining with 'by = join_by(rowname)'
```

```
fig5 <- ggplot(comb_div, aes(x = Season, y = Value, fill = Index)) +
  geom_boxplot(show.legend = F) +
  scale_x_discrete(labels = c('Fall', 'Wint', 'Spri', 'Summ')) +
  facet_wrap(Index ~ year(Date)) +
  xlab('') +
  ylab('Shannon & RaoQ Diversity') +
  theme(axis.text.y = element_blank(),
        axis.text.x = element_text(size = ax_text_size, hjust = 0.5),
        axis.ticks = element_blank(),
        strip.text = element_blank(),
        axis.title = element_text(size = lab_text_size),
        legend.text = element_blank(),
        legend.title = element_text(size = lab_text_size))
```

fig5



Bray distances over time

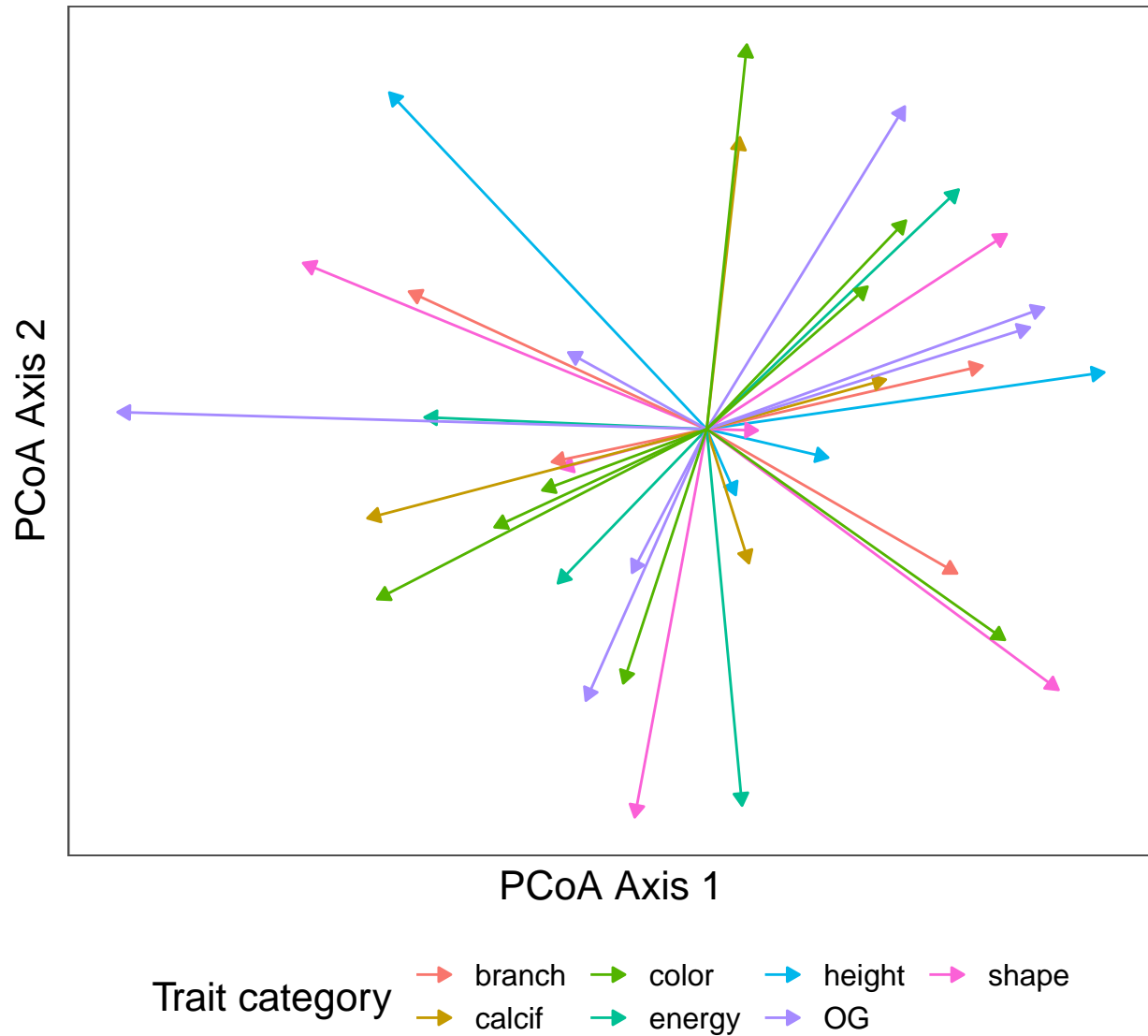
Plots

```
fig4 <- ggplot(trait_vectors, aes(x = Axis.1, y = Axis.2, color = Trait)) +
  geom_segment(aes(x = 0, y = 0, xend = Axis.1, yend = Axis.2), arrow = arrow(length = unit(2, 'mm')),
    scale_color_discrete(name = 'Trait category') +
    #geom_label(aes(label = rowname), show.legend = F) +
    xlab('PCoA Axis 1') +
    ylab('PCoA Axis 2') +
    theme(axis.text = element_blank(),
      axis.ticks = element_blank(),
      axis.title = element_text(size = lab_text_size + 2),
      legend.position = 'bottom',
      #legend.text.ali = 'top',
```

```

#legend.title.position = 'bottom',
legend.title = element_text(size = lab_text_size + 2),
legend.text = element_text(size = ax_text_size))
fig4

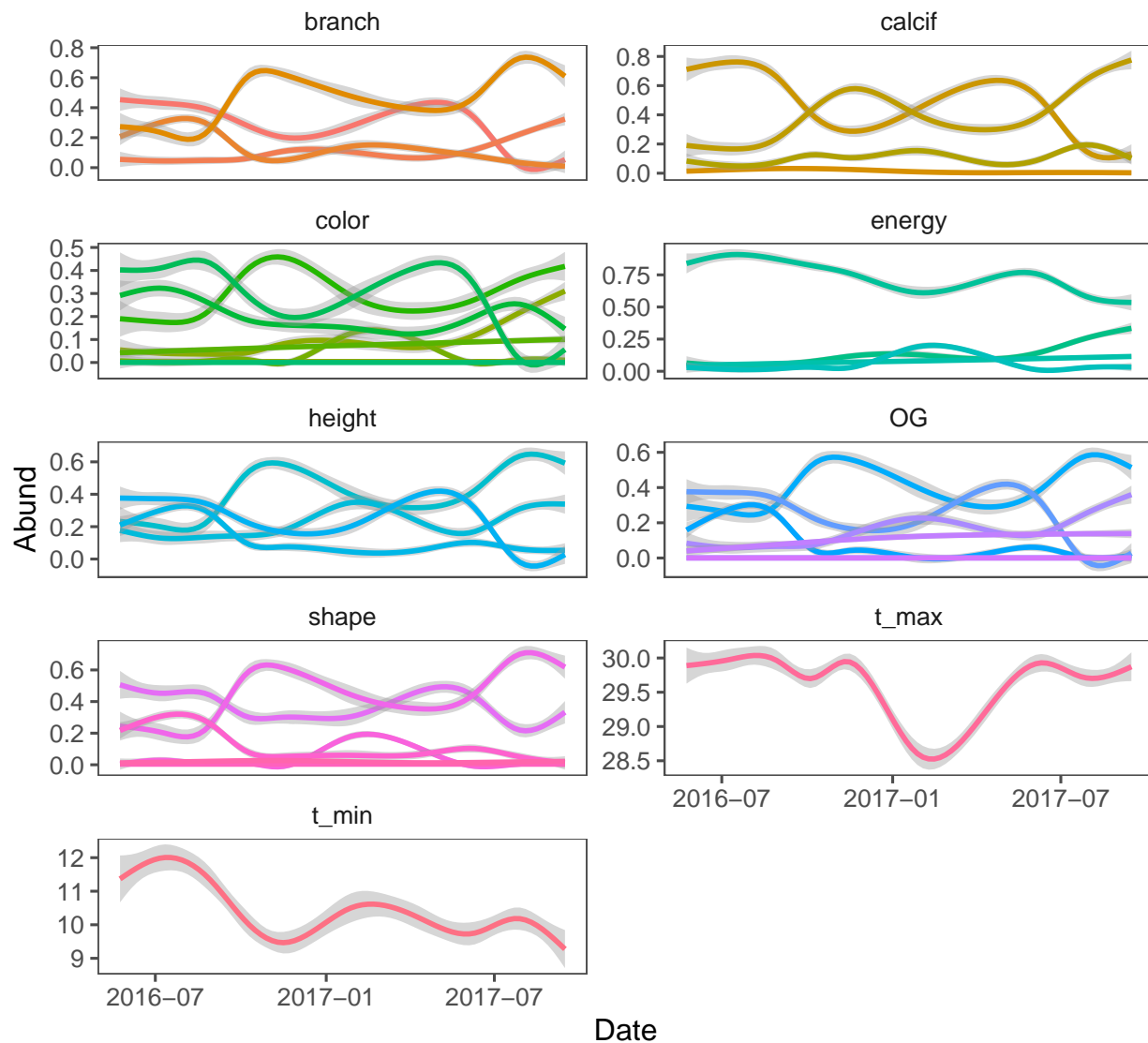
```



```

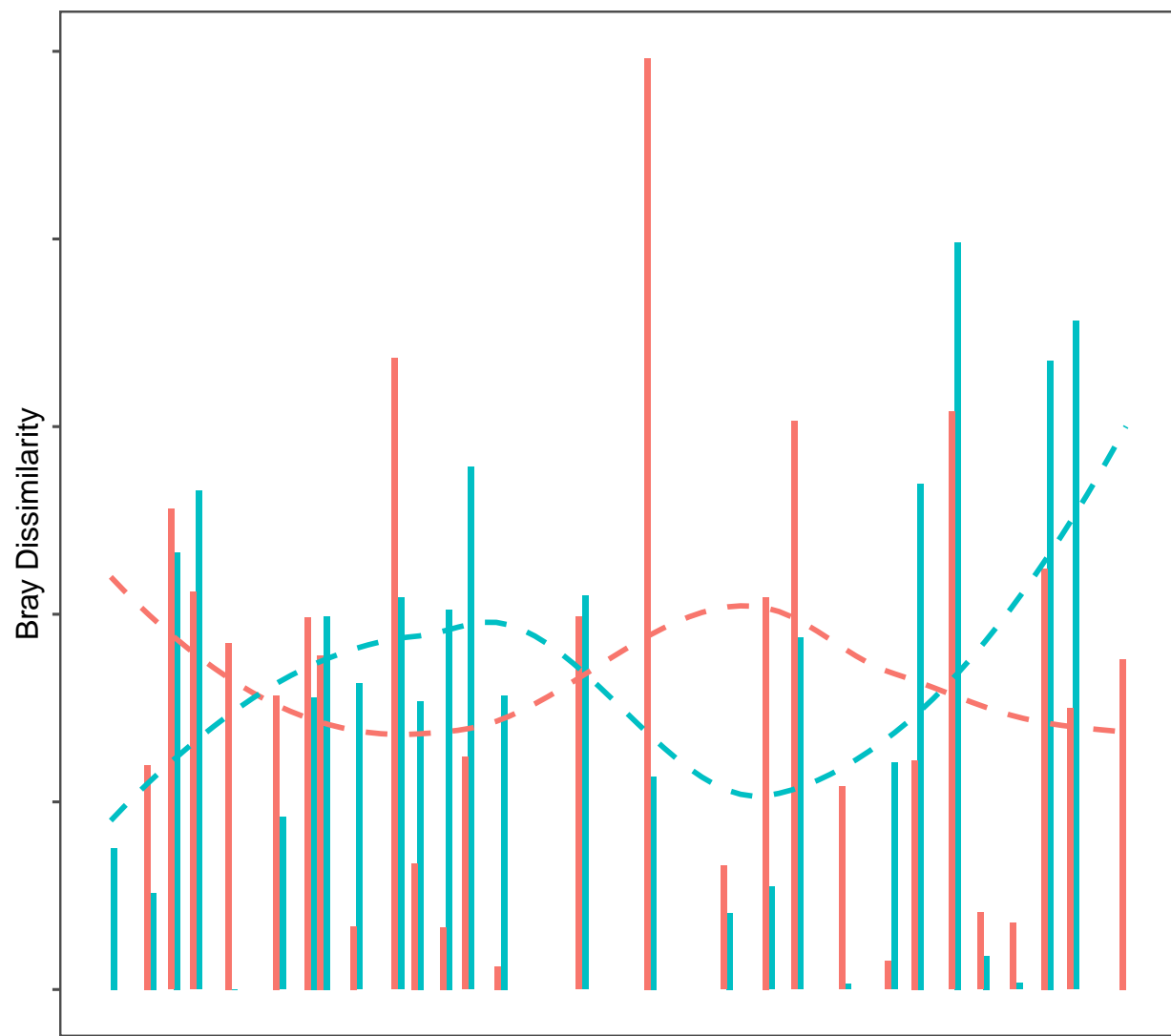
## Joining with 'by = join_by(Date, Bray, Type)'
## 'geom_smooth()' using method = 'gam' and formula = 'y ~ s(x, bs = "cs")'

```

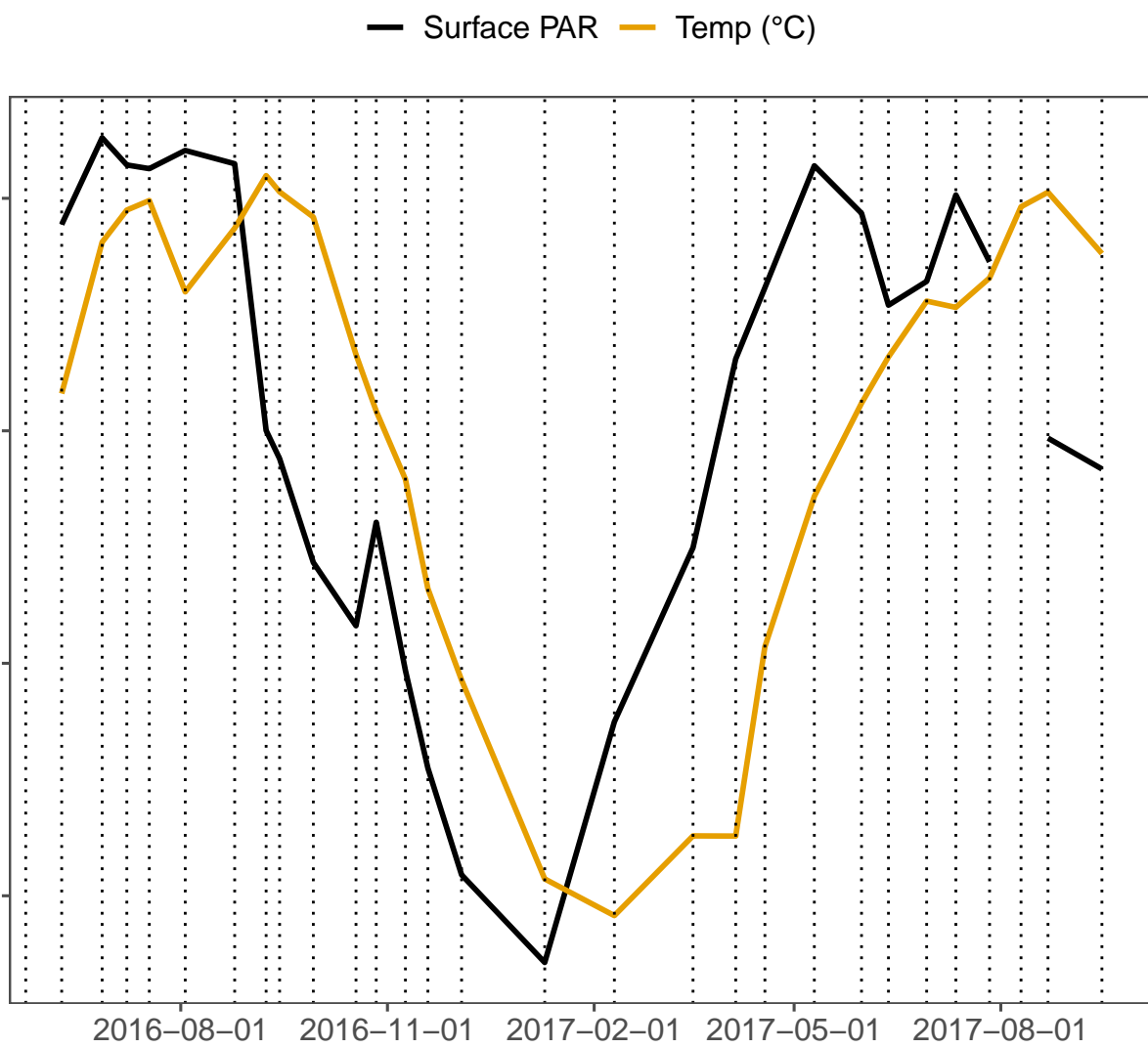



```
## 'geom_smooth()' using method = 'loess' and formula = 'y ~ x'
```

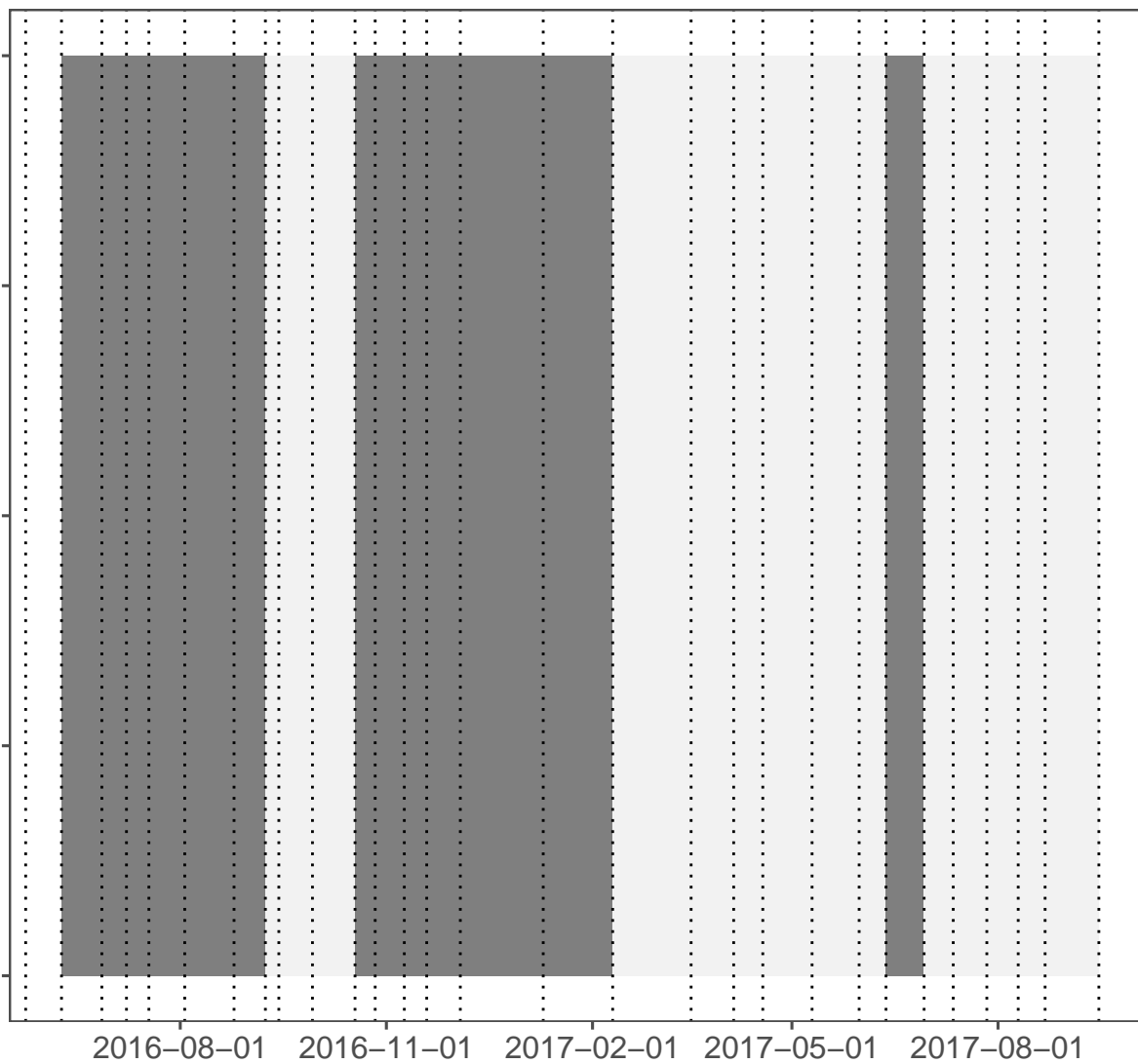
```
## Warning: Removed 2 rows containing missing values ('geom_col()').
```



Warning: Removed 60 rows containing missing values ('geom_line()').



Warning: Removed 2 rows containing missing values ('geom_rect()').



```
fig1 <- ggarrange(eplot, cplot, mplot, ncol = 1, heights = c(3,3,1))
```

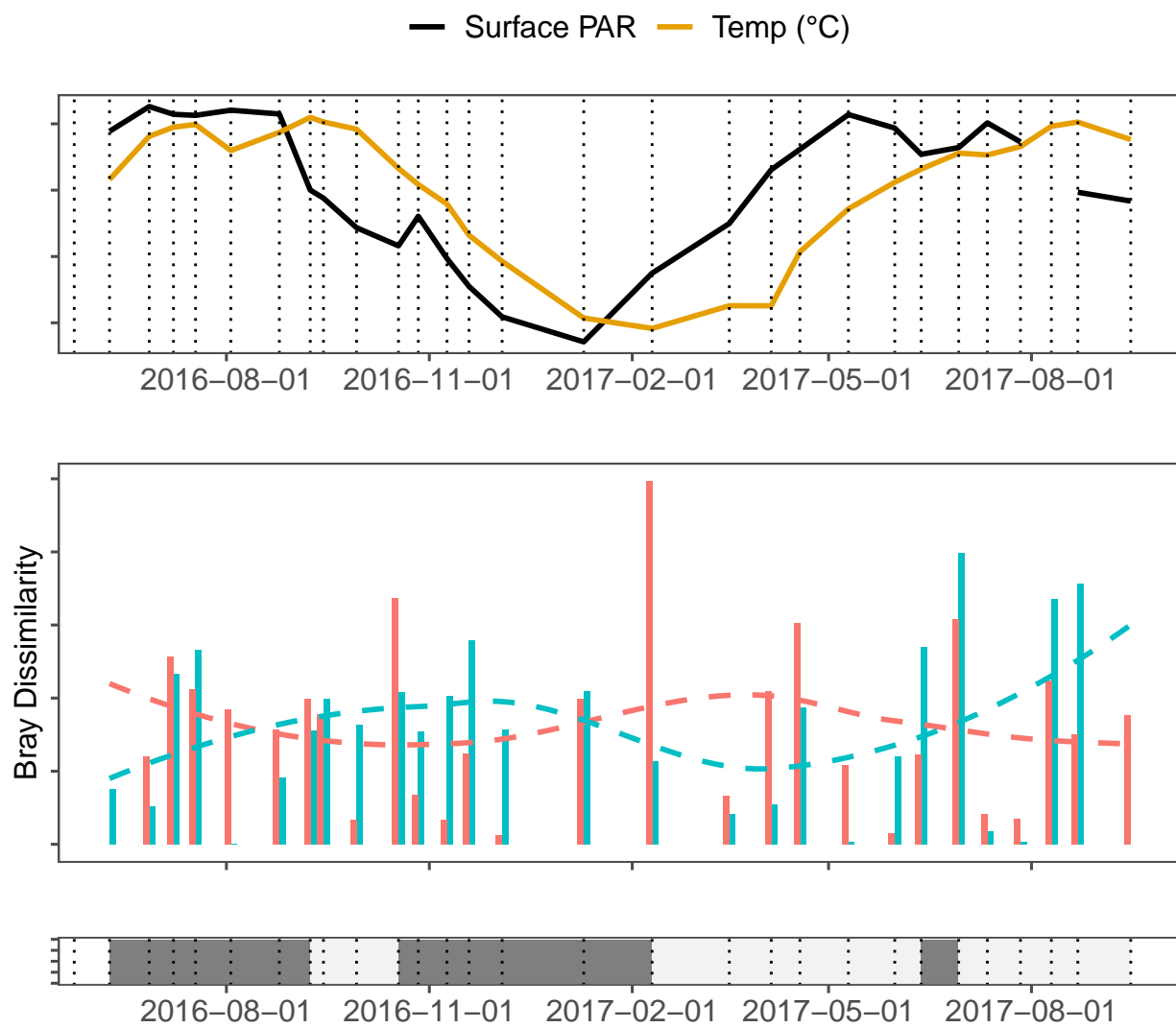
```
## Warning: Removed 60 rows containing missing values ('geom_line()').
```

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

```
## Warning: Removed 2 rows containing missing values ('geom_col()').
```

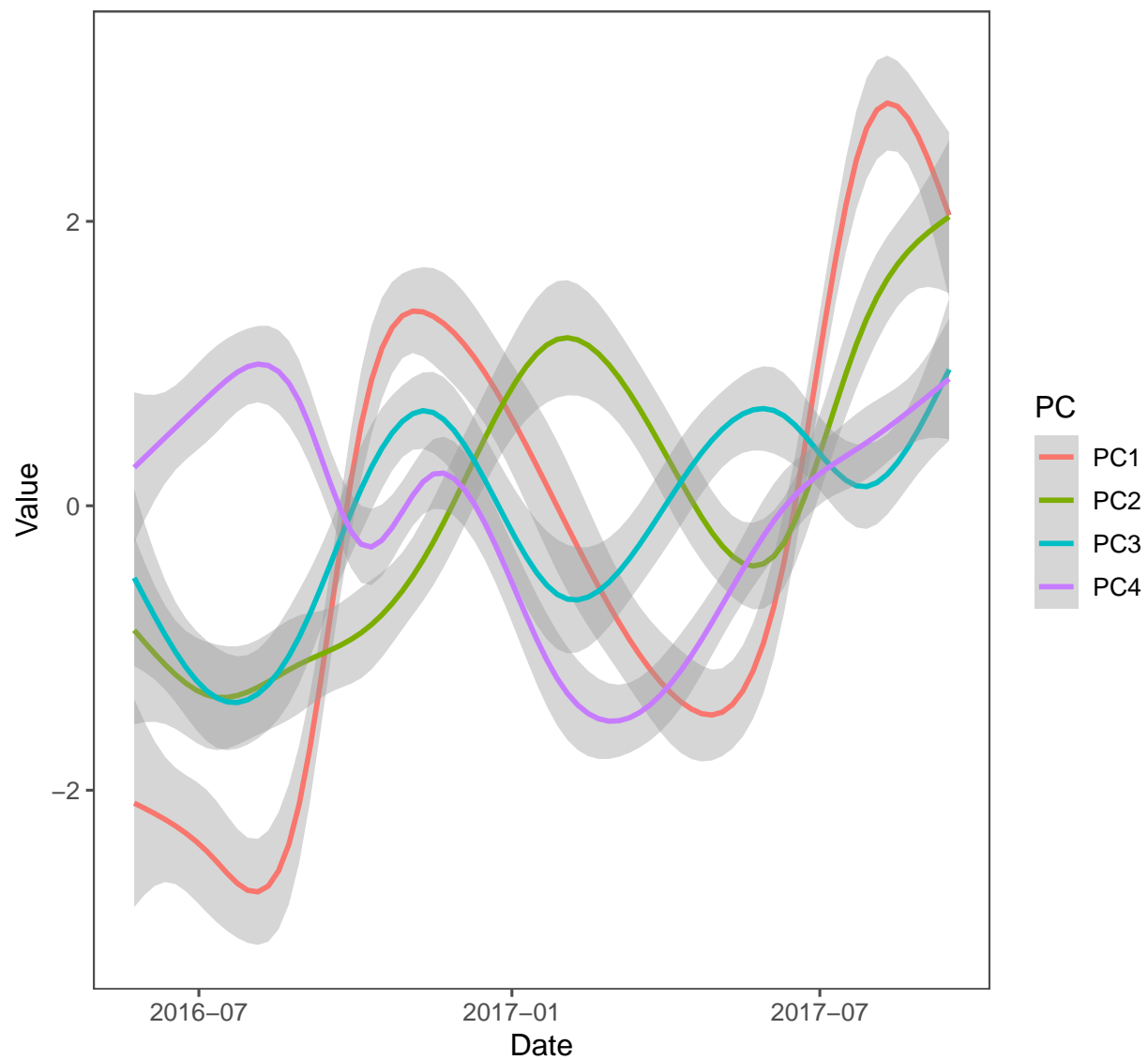
```
## Warning: Removed 2 rows containing missing values ('geom_rect()').
```

```
fig1
```

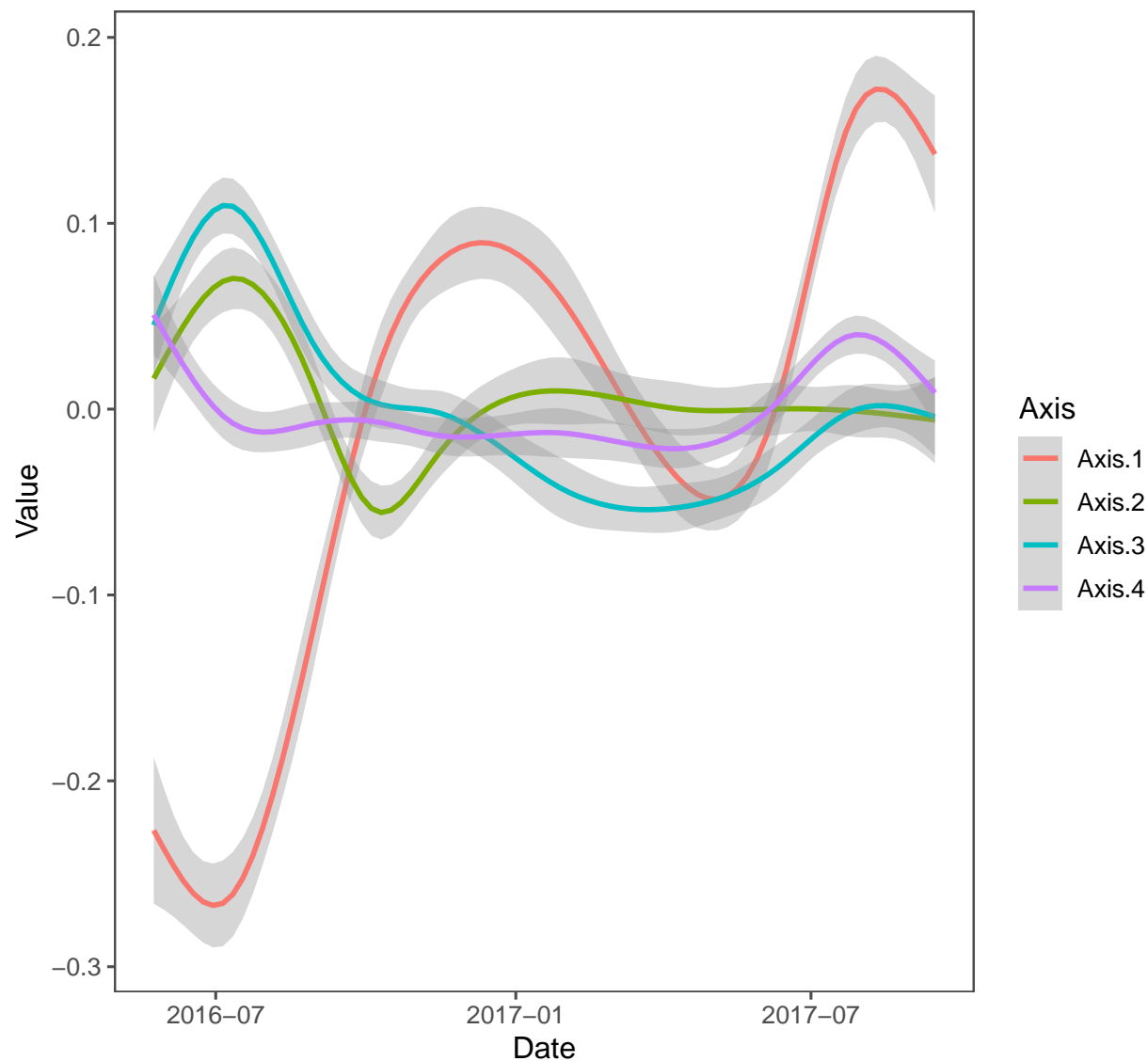


Models - take out sal and waves

```
## Joining with 'by = join_by(rowname)'
## Joining with 'by = join_by(rowname)'
## 'geom_smooth()' using method = 'gam' and formula = 'y ~ s(x, bs = "cs")'
```



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



```
# compare models before doing it for a bunch
lmer(scale(comm_pcoa$Axis.1) ~ PAR_mov * Temp_mov + Salinity_ppt_mov + Wave_ht_ft_mov + (1|Site) + (1|Q

## Linear mixed model fit by REML ['lmerMod']
## Formula: scale(comm_pcoa$Axis.1) ~ PAR_mov * Temp_mov + Salinity_ppt_mov +
##   Wave_ht_ft_mov + (1 | Site) + (1 | Quad) + (1 | Date)
##   Data: scale_env
## REML criterion at convergence: 2463.092
## Random effects:
##   Groups   Name                Std.Dev.
##   Date      (Intercept) 0.5943
##   Quad      (Intercept) 0.3750
##   Site      (Intercept) 0.1185
##   Residual                    0.5883
## Number of obs: 1288, groups:  Date, 25; Quad, 20; Site, 3
## Fixed Effects:
```

```
##      (Intercept)          PAR_mov          Temp_mov  Salinity_ppt_mov
##      0.12645          -0.46948          0.10365          0.12056
## Wave_ht_ft_mov  PAR_mov:Temp_mov
##      -0.00629          -0.12016
```

```
lmer(scale(cwm_pca$PC1) ~ PAR_mov * Temp_mov + Salinity_ppt_mov + Wave_ht_ft_mov + (1|Site) + (1|Quad)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: scale(cwm_pca$PC1) ~ PAR_mov * Temp_mov + Salinity_ppt_mov +
##      Wave_ht_ft_mov + (1 | Site) + (1 | Quad) + (1 | Date)
##      Data: scale_env
## REML criterion at convergence: 3042.304
## Random effects:
## Groups   Name                Std.Dev.
## Date     (Intercept) 0.4971
## Quad     (Intercept) 0.3196
## Site     (Intercept) 0.2245
## Residual                    0.7455
## Number of obs: 1288, groups: Date, 25; Quad, 20; Site, 3
## Fixed Effects:
##      (Intercept)          PAR_mov          Temp_mov  Salinity_ppt_mov
##      0.06429          -0.37064          0.28175          0.02242
## Wave_ht_ft_mov  PAR_mov:Temp_mov
##      0.07732          -0.02583
```

```
lmer(scale(comm_pcoa$Axis.1) ~ PAR_mov + Temp_mov + (1|Site) + (1|Quad) + (1|Date), data = scale_env)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: scale(comm_pcoa$Axis.1) ~ PAR_mov + Temp_mov + (1 | Site) + (1 |
##      Quad) + (1 | Date)
##      Data: scale_env
## REML criterion at convergence: 2790.211
## Random effects:
## Groups   Name                Std.Dev.
## Date     (Intercept) 0.6169
## Quad     (Intercept) 0.3824
## Site     (Intercept) 0.1157
## Residual                    0.6128
## Number of obs: 1403, groups: Date, 28; Quad, 20; Site, 3
## Fixed Effects:
## (Intercept)          PAR_mov          Temp_mov
## -0.03835          -0.41205          0.07608
```

```
lmer(scale(cwm_pca$PC1) ~ PAR_mov + Temp_mov + (1|Site) + (1|Quad) + (1|Date), data = scale_env)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: scale(cwm_pca$PC1) ~ PAR_mov + Temp_mov + (1 | Site) + (1 | Quad) +
##      (1 | Date)
##      Data: scale_env
## REML criterion at convergence: 3268.42
## Random effects:
## Groups   Name                Std.Dev.
```



```
## Date      (Intercept) 0.4919
## Quad      (Intercept) 0.3105
## Site      (Intercept) 0.2083
## Residual                0.7354
## Number of obs: 1403, groups: Date, 28; Quad, 20; Site, 3
## Fixed Effects:
## (Intercept)      PAR_mov      Temp_mov
##      -0.02055      -0.37355      0.22146
```

```
# scaling helped, added Date as rf
```

```
# may need to average communities on each date.
```

```
mod1c <- lmer(scale(comm_pcoa$Axis.1) ~ PAR_mov + Temp_mov + (1|Site) + (1|Quad) + (1|Date), data = scale_env)
mod2c <- lmer(scale(comm_pcoa$Axis.2) ~ PAR_mov + Temp_mov + (1|Site) + (1|Quad) + (1|Date), data = scale_env)
mod3c <- lmer(scale(comm_pcoa$Axis.3) ~ PAR_mov + Temp_mov + (1|Site) + (1|Quad) + (1|Date), data = scale_env)
mod4c <- lmer(scale(comm_pcoa$Axis.4) ~ PAR_mov + Temp_mov + (1|Site) + (1|Quad) + (1|Date), data = scale_env)
```

```
mod1t <- lmer(scale(cwm_pca$PC1) ~ PAR_mov + Temp_mov + (1|Site) + (1|Quad) + (1|Date), data = scale_env)
mod2t <- lmer(scale(cwm_pca$PC2) ~ PAR_mov + Temp_mov + (1|Site) + (1|Quad) + (1|Date), data = scale_env)
mod3t <- lmer(scale(cwm_pca$PC3) ~ PAR_mov + Temp_mov + (1|Site) + (1|Quad) + (1|Date), data = scale_env)
mod4t <- lmer(scale(cwm_pca$PC4) ~ PAR_mov + Temp_mov + (1|Site) + (1|Quad) + (1|Date), data = scale_env)
```

```
mod1c
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: scale(comm_pcoa$Axis.1) ~ PAR_mov + Temp_mov + (1 | Site) + (1 |
##      Quad) + (1 | Date)
##      Data: scale_env
## REML criterion at convergence: 2790.211
## Random effects:
## Groups      Name          Std.Dev.
## Date      (Intercept) 0.6169
## Quad      (Intercept) 0.3824
## Site      (Intercept) 0.1157
## Residual                0.6128
## Number of obs: 1403, groups: Date, 28; Quad, 20; Site, 3
## Fixed Effects:
## (Intercept)      PAR_mov      Temp_mov
##      -0.03835      -0.41205      0.07608
```

```
mod2c
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: scale(comm_pcoa$Axis.2) ~ PAR_mov + Temp_mov + (1 | Site) + (1 |
##      Quad) + (1 | Date)
##      Data: scale_env
## REML criterion at convergence: 3713.586
## Random effects:
## Groups      Name          Std.Dev.
## Date      (Intercept) 0.3067
## Quad      (Intercept) 0.1857
## Site      (Intercept) 0.4121
## Residual                0.8777
## Number of obs: 1403, groups: Date, 28; Quad, 20; Site, 3
```

```
## Fixed Effects:
## (Intercept)      PAR_mov      Temp_mov
##      0.05302      0.24016      -0.09299
```

mod3c

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: scale(comm_pcoa$Axis.3) ~ PAR_mov + Temp_mov + (1 | Site) + (1 |
##      Quad) + (1 | Date)
##      Data: scale_env
## REML criterion at convergence: 3646.938
## Random effects:
## Groups      Name          Std.Dev.
## Date        (Intercept)  0.4220
## Quad         (Intercept)  0.1778
## Site         (Intercept)  0.2840
## Residual                    0.8528
## Number of obs: 1403, groups:  Date, 28; Quad, 20; Site, 3
## Fixed Effects:
## (Intercept)      PAR_mov      Temp_mov
##      0.06036      0.08375      0.26488
```

mod4c

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: scale(comm_pcoa$Axis.4) ~ PAR_mov + Temp_mov + (1 | Site) + (1 |
##      Quad) + (1 | Date)
##      Data: scale_env
## REML criterion at convergence: 3706.504
## Random effects:
## Groups      Name          Std.Dev.
## Date        (Intercept)  0.3360
## Quad         (Intercept)  0.3105
## Site         (Intercept)  0.2092
## Residual                    0.8698
## Number of obs: 1403, groups:  Date, 28; Quad, 20; Site, 3
## Fixed Effects:
## (Intercept)      PAR_mov      Temp_mov
##      -0.027251      -0.003781      0.099973
```

mod1t

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: scale(cwm_pca$PC1) ~ PAR_mov + Temp_mov + (1 | Site) + (1 | Quad) +
##      (1 | Date)
##      Data: scale_env
## REML criterion at convergence: 3268.42
## Random effects:
## Groups      Name          Std.Dev.
## Date        (Intercept)  0.4919
## Quad         (Intercept)  0.3105
## Site         (Intercept)  0.2083
## Residual                    0.7354
```

```
## Number of obs: 1403, groups: Date, 28; Quad, 20; Site, 3
## Fixed Effects:
## (Intercept)      PAR_mov      Temp_mov
##      -0.02055      -0.37355      0.22146
```

mod2t

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: scale(cwm_pca$PC2) ~ PAR_mov + Temp_mov + (1 | Site) + (1 | Quad) +
##      (1 | Date)
##      Data: scale_env
## REML criterion at convergence: 3617.544
## Random effects:
## Groups      Name      Std.Dev.
## Date      (Intercept) 0.3883
## Quad      (Intercept) 0.1480
## Site      (Intercept) 0.3125
## Residual              0.8463
## Number of obs: 1403, groups: Date, 28; Quad, 20; Site, 3
## Fixed Effects:
## (Intercept)      PAR_mov      Temp_mov
##      -0.04667      -0.10786      -0.09537
```

mod3t

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: scale(cwm_pca$PC3) ~ PAR_mov + Temp_mov + (1 | Site) + (1 | Quad) +
##      (1 | Date)
##      Data: scale_env
## REML criterion at convergence: 3880.135
## Random effects:
## Groups      Name      Std.Dev.
## Date      (Intercept) 0.3354
## Quad      (Intercept) 0.1545
## Site      (Intercept) 0.1800
## Residual              0.9344
## Number of obs: 1403, groups: Date, 28; Quad, 20; Site, 3
## Fixed Effects:
## (Intercept)      PAR_mov      Temp_mov
##      -0.05010      -0.13881      0.02942
```

mod4t

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: scale(cwm_pca$PC4) ~ PAR_mov + Temp_mov + (1 | Site) + (1 | Quad) +
##      (1 | Date)
##      Data: scale_env
## REML criterion at convergence: 3803.069
## Random effects:
## Groups      Name      Std.Dev.
## Date      (Intercept) 0.23387
## Quad      (Intercept) 0.12258
## Site      (Intercept) 0.08959
```

```
## Residual                0.91594
## Number of obs: 1403, groups: Date, 28; Quad, 20; Site, 3
## Fixed Effects:
## (Intercept)      PAR_mov      Temp_mov
##      0.00776      -0.04456      0.32962
```

```
results <- as.data.frame(matrix(nrow = 6, ncol = 0))

r1c <- summary(mod1c)$coefficients %>%
  as.data.frame() %>%
  rownames_to_column(var = 'Var') %>%
  mutate(PC = factor(1),
         Type = 'tax')
r2c <- summary(mod2c)$coefficients %>%
  as.data.frame() %>%
  rownames_to_column(var = 'Var') %>%
  mutate(PC = factor(2),
         Type = 'tax')
r3c <- summary(mod3c)$coefficients %>%
  as.data.frame() %>%
  rownames_to_column(var = 'Var') %>%
  mutate(PC = factor(3),
         Type = 'tax')
r4c <- summary(mod4c)$coefficients %>%
  as.data.frame() %>%
  rownames_to_column(var = 'Var') %>%
  mutate(PC = factor(4),
         Type = 'tax')
r1f <- summary(mod1t)$coefficients %>%
  as.data.frame() %>%
  rownames_to_column(var = 'Var') %>%
  mutate(PC = factor(1),
         Type = 'fun')
r2f <- summary(mod2t)$coefficients %>%
  as.data.frame() %>%
  rownames_to_column(var = 'Var') %>%
  mutate(PC = factor(2),
         Type = 'fun')
r3f <- summary(mod3t)$coefficients %>%
  as.data.frame() %>%
  rownames_to_column(var = 'Var') %>%
  mutate(PC = factor(3),
         Type = 'fun')
r4f <- summary(mod4t)$coefficients %>%
  as.data.frame() %>%
  rownames_to_column(var = 'Var') %>%
  mutate(PC = factor(4),
         Type = 'fun')
results <- bind_rows(r1c, r2c, r3c, r4c, r1f, r2f, r3f, r4f) %>%
  group_by(Var) %>%
  mutate(Estimate = abs(Estimate),
         Std_error = `Std. Error`,
         t_value = `t value`) %>%
  ungroup() %>%
```

```

mutate(PC = factor(PC, levels = c(4,3,2,1)),
       Type = factor(Type)) %>%
select(-`Std. Error`, -`t value`) %>%
filter(Var != '(Intercept)') %>%
filter(PC %in% c(1, 2, 3, 4)) %>%
mutate(Valid_estimate = Estimate * (Estimate > Std_error),
       Var = if_else(Var == 'PAR_mov', 'Surface PAR', 'Temp (\u00B0C)')) %>%
group_by(Var, Type) %>%
mutate(Total_valid_estimate = sum(Valid_estimate)) %>%
ungroup()

```

results

```

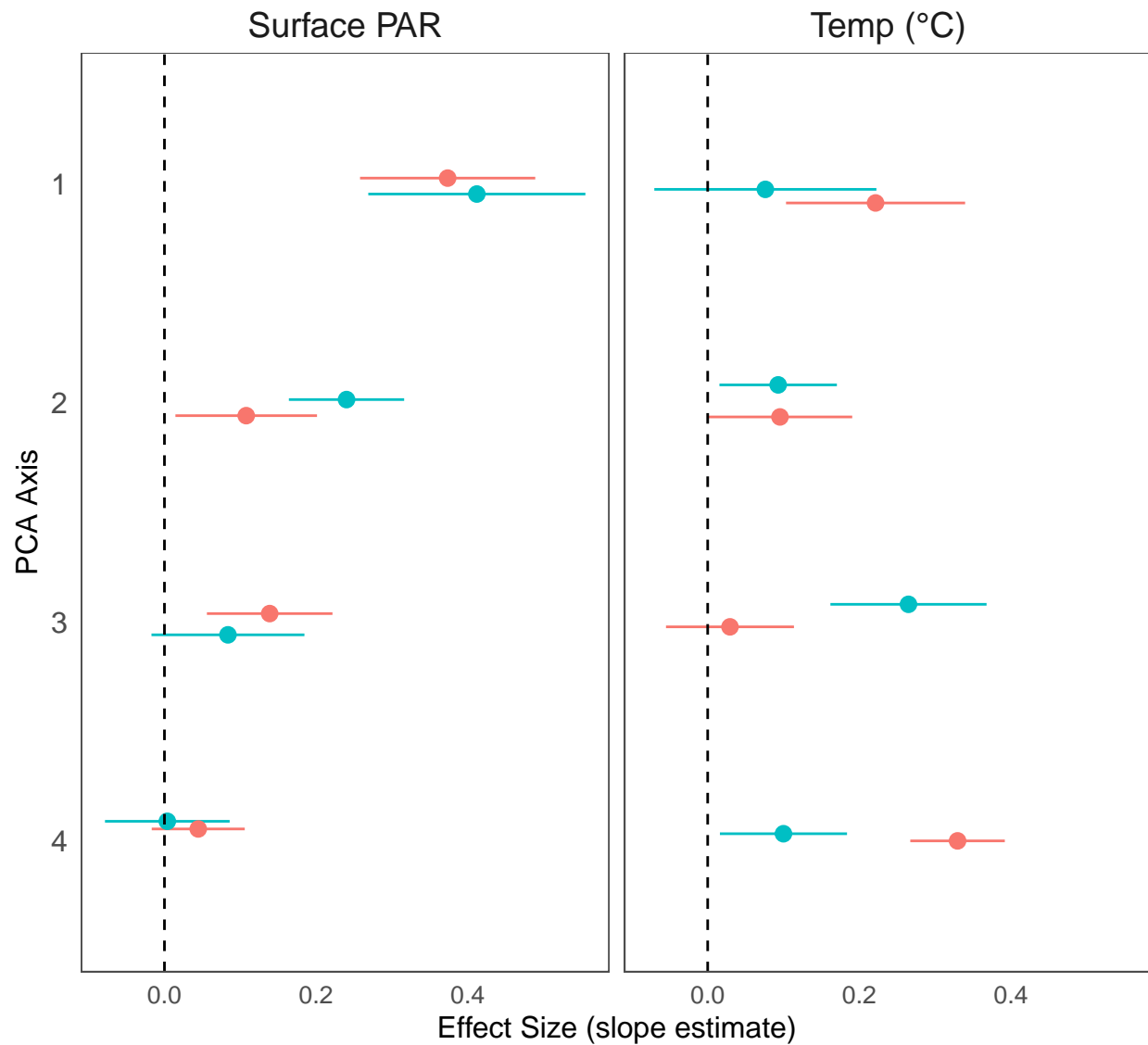
## # A tibble: 16 x 8
##   Var      Estimate PC   Type Std_error t_value Valid_estimate
##   <chr>      <dbl> <fct> <fct>    <dbl>    <dbl>        <dbl>
## 1 Surface PAR  0.412   1    tax    0.143   -2.88         0.412
## 2 Temp (°C)   0.0761  1    tax    0.146    0.519         0
## 3 Surface PAR  0.240   2    tax    0.0760   3.16         0.240
## 4 Temp (°C)   0.0930  2    tax    0.0775  -1.20         0.0930
## 5 Surface PAR  0.0838  3    tax    0.101    0.830         0
## 6 Temp (°C)   0.265   3    tax    0.103    2.57         0.265
## 7 Surface PAR  0.00378  4    tax    0.0822  -0.0460        0
## 8 Temp (°C)   0.100   4    tax    0.0838   1.19         0.100
## 9 Surface PAR  0.374   1    fun    0.116   -3.23         0.374
## 10 Temp (°C)   0.221   1    fun    0.118    1.87         0.221
## 11 Surface PAR  0.108   2    fun    0.0934  -1.16         0.108
## 12 Temp (°C)   0.0954  2    fun    0.0953  -1.00         0.0954
## 13 Surface PAR  0.139   3    fun    0.0828  -1.68         0.139
## 14 Temp (°C)   0.0294  3    fun    0.0844   0.349         0
## 15 Surface PAR  0.0446  4    fun    0.0614  -0.726         0
## 16 Temp (°C)   0.330   4    fun    0.0623   5.29         0.330
## # i 1 more variable: Total_valid_estimate <dbl>

```

```

rplot <- results %>%
  ggplot(aes(y = PC, x = Estimate, color = Type)) +
  #geom_col(aes(y = Var, x = Total_valid_estimate, fill = Type), color = 'grey25', position = 'dodge')
  geom_pointrange(aes(xmin = Estimate - Std_error, xmax = Estimate + Std_error), position = position_dodge()) +
  geom_vline(xintercept = 0, linetype = 2, color = 'black') +
  xlab('Effect Size (slope estimate)') +
  ylab('PCA Axis') +
  scale_color_discrete(name = 'Analysis', labels = c('Functional', 'Taxonomic')) +
  facet_grid(~ Var) +
  theme(axis.ticks = element_blank(),
        axis.text.y = element_text(size = ax_text_size),
        plot.margin = margin(0,0,8,8),
        strip.text = element_text(size = lab_text_size))
rplot

```



```
#
# legend_fig <- results %>%
#   ggplot(aes(y = PC, x = Estimate, fill = Type)) +
#   geom_col() +
#   scale_fill_discrete(name = 'Analysis ', labels = c('Functional', 'Taxonomic')) +
#   theme(legend.title = element_text(size = 36),
#         legend.text = element_text(size = 30),
#         legend.key.size = unit(2, 'cm'),
#         legend.position = 'top'#,
#         #legend.key.spacing.x = unit(.25, 'in')
#   )
# legend_fig

# total_plot <- results %>%
#   select(Var, Type, Total_valid_estimate) %>%
#   unique() %>%
#   ggplot(aes(x = Var, y = Total_valid_estimate, fill = Type)) +
```

```

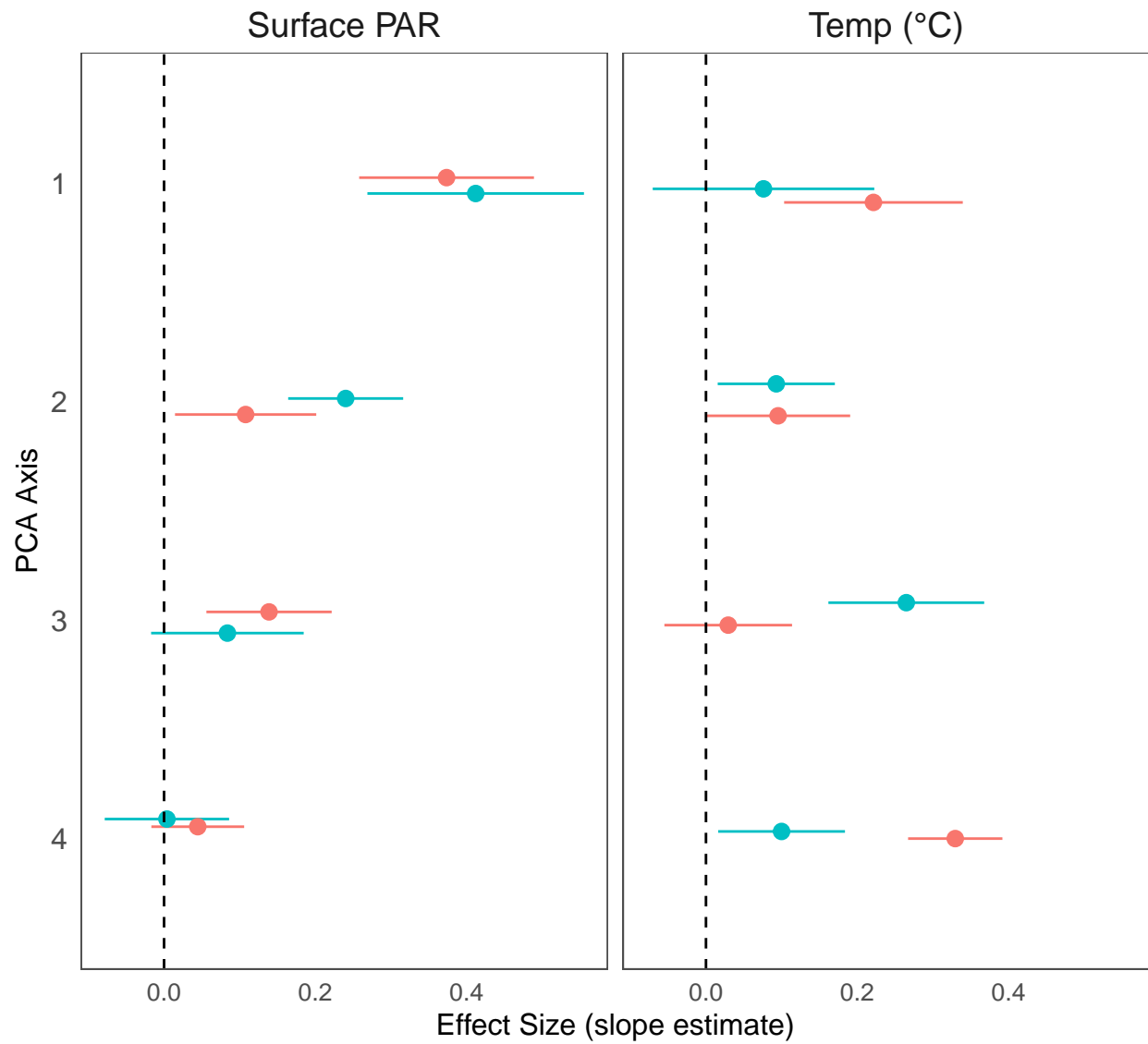
#   geom_col(position = position_dodge2()) +
#   xlab('') +
#   ylab('') +
#   scale_fill_discrete(name = 'Analysis', labels = c('Functional', 'Taxonomic')) +
#   facet_wrap(~ Var, scales = 'free_x') +
#   theme(axis.ticks = element_blank(),
#         axis.text = element_text(size = ax_text_size),
#         axis.text.x = element_blank(),
#         legend.title = element_text(size = lab_text_size),
#         legend.text = element_text(size = ax_text_size),
#         plot.margin = margin(0,0,20,8),
#         strip.text = element_text(size = lab_text_size))
# total_plot
#
# res_combined_plot <- ggarrange(rplot, total_plot, nrow = 1)
# ggarrange(eplot, cplot, res_combined_plot, ncol = 1)

```

```

fig2 <- rplot
fig2

```



```
#ggarrange(eplot, cplot, rplot, ncol = 1)
```