CO2-NDVI-effect_PETA_GAM_RLM

Import data

```
library(tidyverse)
library(data.table); #setDTthreads(threads = 0)
library(lubridate);
library(dtplyr);
library(RcppArmadillo)
library(sf); library(stars)
library(patchwork); library(colorspace)
library(zyp);
library(mgcv)
set.seed(333)
# IMPORT ##############################################################################
oz_poly <- sf::read_sf("../data_general/GADM/gadm36_AUS.gpkg",
                       layer="gadm36_AUS_1")
oz_poly <- st_as_sf(oz_poly)
oz_poly <- st_simplify(oz_poly, dTolerance = 0.05)

# vegetation index record
vi <- arrow::read_parquet("../data_general/MCD43/MCD43_AVHRR_NDVI_hybrid_2020-10-12.parquet") %>%
  as.data.table()
vi <- vi %>% lazy_dt() %>%
  mutate(ndvi_hyb_e1 = coalesce(ndvi_mcd_nm_pred, NA_real_),
         ndvi_hyb_e2 = coalesce(ndvi_mcd, NA_real_)) %>%
  mutate(ndvi_hyb = coalesce(ndvi_hyb_e2, ndvi_hyb_e1)) %>%
  mutate(ndvi_hyb = ifelse(between(ndvi_hyb,0,1),ndvi_hyb,NA_real_)) %>%
  as.data.table()
norms_vi <- vi[,`:=`(month=month(date))] %>%
  .[,.(ndvi_u = mean(ndvi_hyb,na.rm=TRUE),
       ndvi_sd = sd(ndvi_hyb,na.rm=TRUE)),keyby=.(x,y,month)]
vi <- norms_vi[vi,on=.(x,y,month)] %>%
  .[,`:=`(ndvi_anom = ndvi_hyb - ndvi_u)] %>%
  .[,`:=`(ndvi_anom_sd = ndvi_anom/ndvi_sd)]

lvi <- lazy_dt(vi)

lvi %>% group_by(x,y) %>% summarize(ndvi_u =mean(ndvi_hyb,na.rm=TRUE)) %>% show_query()
```

```
## `_DT2`[, .(ndvi_u = mean(ndvi_hyb, na.rm = TRUE)), keyby = .(x,
##     y)]
```

```
norms_vi <- vi[,`:=`(month=month(date))] %>%
  .[,.(ndvi_u = mean(ndvi_hyb,na.rm=TRUE),
       ndvi_sd = sd(ndvi_hyb,na.rm=TRUE)),keyby=.(x,y,month)]
```

```
vi <- norms_vi[vi,on=.(x,y,month)] %>%
  .[,`:=`(ndvi_anom = ndvi_hyb - ndvi_u)] %>%
  .[,`:=`(ndvi_anom_sd = ndvi_anom/ndvi_sd)]

vi <- vi %>% select(-season)

dat <- arrow::read_parquet("/home/sami/scratch/ARD_ndvi_aclim_anoms.parquet",
                           col_select = c(
                             "date", "hydro_year", "id",
                             "precip", "precip_12mo","precip_anom", "precip_anom_12mo","map",
                             "vpd15","vpd15_12mo","vpd15_anom", "vpd15_anom_12mo","mavpd15",
                             "pet","pet_anom","pet_anom_12mo","pet_12mo", "mapet",
                             "pe","pe_12mo","pe_anom_12mo",
                             'vc','veg_class',
                             'month',
                             "x", "y", "year")) %>%
  as.data.table()

norms_mape <- dat %>% lazy_dt() %>%
  filter(date>=ymd("1982-01-01") & date<=ymd("2011-12-31")) %>%
  group_by(x,y,hydro_year) %>%
  summarize(ppet_12mo = mean(precip_12mo/pet_12mo,na.rm=TRUE)) %>%
  ungroup() %>%
  group_by(x,y) %>%
  summarize(mappet = mean(ppet_12mo,na.rm=TRUE)) %>%
  ungroup() %>%
  as.data.table()
dat <- merge(dat,norms_mape,by=c("x","y"))

dat <- dat[,`:=`(pe_12mo = precip_12mo/pet_12mo)]
dat <- merge(dat,
             vi,
             by=c("x","y","date","month","year"),
             all=TRUE,allow.cartesian=TRUE)
dat <- dat[order(x,y,date)][, ndvi_3mo := frollmean(ndvi_hyb,n = 3,fill = NA,align='center',na.rm=TRUE)]

rm(vi); gc(full=TRUE)
```

```
##              used    (Mb) gc trigger     (Mb)    max used     (Mb)
## Ncells    2392989   127.8    3921038   209.5    3921038   209.5
## Vcells 1201452011 9166.4 3262010088 24887.2 3250881675 24802.3
```

```
# Attach CO2 from Mauna Loa
mlo <- readr::read_table("../data_general/CO2_growth_rate/co2_mm_mlo_20200405.txt",
                         skip = 72, col_names = F) %>%
  set_names(
    c("year","month","ddate","co2_avg","co2_int","co2_trend","ndays")
  ) %>%
  mutate(date = ymd(paste(year,month,1))) %>%
  select(date,co2_int,co2_trend) %>%
  as.data.table()
dat <- merge(mlo,dat,by="date")
center_co2 <- mean(dat$co2_int)
```

```
dat <- dat[,`:=`(cco2=co2_int-center_co2)]
gc()
```

```
##             used   (Mb) gc trigger    (Mb)   max used    (Mb)
## Ncells    2442619  130.5    3921038   209.5    3921038   209.5
## Vcells 1247957046 9521.2 3262010088 24887.2 3250881675 24802.3
```

```
dat <- dat[is.na(vc)==F]
dat <- dat[str_detect(vc,"Forests") |
             str_detect(vc, "Eucalypt") |
             str_detect(vc, "Rainforests")]
dat <- dat[x>= 140] # FILTER TO LON >= 140
dat[,`:=`(epoch = ifelse(date<ymd("2000-12-31"),'avhrr','modis'))]
dat[,`:=`(year=year(date),month=month(date))] %>%
  .[,`:=`(season = case_when(month%in%c(3:5)~'MAM',
                             month%in%c(6:8)~'JJA',
                             month%in%c(9:11)~'SON',
                             month%in%c(12,1,2)~'DJF'))]

# Add seasons
dat <- dat %>% mutate(epoch = as_factor(epoch),
                      season = factor(season, levels=c("SON","DJF","MAM","JJA")))

# Add Koppen climate zones
kop <- arrow::read_parquet("../data_general/Koppen_climate/BOM_Koppen_simplified7.parquet")
kop <- setDT(kop)
kop <- kop[,.(x,y,zone)]
dat <- merge(dat, kop, by=c("x","y"),all = T)

# Load LAI climatology 2002-2019 (before BS fires)
lai <- stars::read_stars("../data_general/Oz_misc_data/MCD15A3H_meanAnnualLAI_20020801_20190801.tif") %>%
  set_names(c("lai")) %>%
  as.data.table()
#*****************************************************************************
```

Only pixel locations with more than 6 observations

In practice this is robust to the number of observations (6 - 18 as a minimum nob)

```
coords_keep <- dat %>%
  lazy_dt() %>%
  group_by(x,y,hydro_year) %>%
  summarize(nobs = sum(is.na(ndvi_hyb)==F)) %>%
  ungroup() %>%
  as.data.table()
coords_keep <- coords_keep %>%
  filter(nobs >= 6) %>%
  group_by(x,y) %>%
  summarize(nobs_annual = n()) %>%
  ungroup()
```

Summarize NDVI and climate data to annual

```
dat_annual <- dat[ndvi_anom_sd >= -3.5 & ndvi_anom_sd <= 3.5] %>%
  .[date>= ymd("1981-11-01") & date<= ymd("2019-08-30")] %>%
  # .[,.(val = mean(ndvi_3mo, na.rm=TRUE)), by=.(x,y,season,hydro_year)] %>%
  .[,`:=`(epoch=ifelse(hydro_year < 2001,0,1))] %>%
  .[is.na(ndvi_hyb)==F] %>%
  .[is.na(pe_anom_12mo)==F] %>%
  merge(., coords_keep,by=c('x','y'))
dat_annual <- dat_annual[,`:=`(hydro_year_c = hydro_year-1982,
                               frac_p_anom = precip_anom_12mo/map,
                               frac_ppet_anom = pe_anom_12mo/mappet,
                               frac_pet_anom = pet_anom_12mo/mapet,
                               frac_vpd_anom = vpd15_anom_12mo/mavpd15)]
dat_annual <- dat_annual %>%
  select(x,y,date,ndvi_hyb,co2_trend,
         hydro_year,hydro_year_c,
         frac_p_anom,frac_ppet_anom,frac_pet_anom,frac_vpd_anom,
         map,mappet,mapet,mavpd15,
         epoch,nobs_annual)
dat_annual <- dat_annual[nobs_annual>=10]
dat_annual <- dat_annual[,.(ndvi_hyb = mean(ndvi_hyb,na.rm=TRUE),
                            co2 = mean(co2_trend,na.rm=TRUE),
                            frac_p_anom = mean(frac_p_anom,na.rm=TRUE),
                            frac_pet_anom = mean(frac_pet_anom,na.rm=TRUE),
                            frac_ppet_anom = mean(frac_ppet_anom,na.rm=TRUE),
                            frac_vpd_anom = mean(frac_vpd_anom,na.rm=TRUE),
                            epoch = mean(epoch, na.rm=TRUE),
                            map = mean(map,na.rm=TRUE),
                            mapet = mean(mapet, na.rm=TRUE),
                            mappet = mean(mappet, na.rm=TRUE),
                            mavpd15 = mean(mavpd15,na.rm=TRUE)),
                         by=.(x,y,hydro_year_c)]
dat_annual <- dat_annual[is.infinite(frac_ppet_anom)==F]
```

Regression: Long-term rate of NDVI increase per year

fraction of area experiencing greening:

```
sum(lt_ndvi_year$b1>0)/nrow(lt_ndvi_year)
```

```
## [1] 0.9046721
```

Fit GAM of CO2 effect and climate

```
dat_train <- dat_annual[sample(.N, floor(nrow(dat_annual)/2))]
dat_test <- fsetdiff(dat_annual,dat_train)
```

```
g3_eval <- bam(ndvi_hyb~
                 te(mappet,co2,k=5)+                       # MAP:MAPET x CO2, with ~5 knots
                 te(mavpd15,frac_vpd_anom,k=5,bs='cs')+  # VPD x VPD_anom ...
                 te(map,frac_p_anom,k=5,bs='cs')+         # MAP x P_anom ...
                 te(mapet,frac_pet_anom,k=5,bs='cs')+     # MAPET x PET_anom ...
                 epoch,
               data=dat_train,
               discrete = T, #
               select=TRUE, # turns on penalization term for splines
               method='fREML') # restricted maximum likelihood
```

Out of sample R2 of GAM

```
 yardstick::rsq_trad_vec(truth=dat_test$ndvi_hyb, estimate=predict(g3_eval,newdata=dat_test))
```

```
## [1] 0.9029181
```

RMSE of GAM

```
yardstick::rmse_vec(truth=dat_test$ndvi_hyb, estimate=predict(g3_eval,newdata=dat_test))
```

```
## [1] 0.05093744
```

Re-fit GAM with full data

```
g3_full <- bam(ndvi_hyb~
                 te(mappet,co2,k=5)+
                 te(mavpd15,frac_vpd_anom,k=5,bs='cs')+
                 te(map,frac_p_anom,k=5,bs='cs')+
                 te(mapet,frac_pet_anom,k=5,bs='cs')+
                 epoch,
               data=dat_annual,
               discrete = T,select=TRUE, method='fREML')
```

Regression: Robust linear model (RLM) CO2 + VPD + P + PET

Regression Thiel Sen: VPD

```
system.time(
  lt_v_sen <- dat[date>=ymd("1981-12-01")][date<=ymd("2019-11-30")] %>%
    .[,`:=`(hydro_year_c = hydro_year-1982)] %>%
    .[is.na(vpd15)==F] %>%
    .[,.(vpd15 = mean(vpd15,na.rm=TRUE)), by=.(x,y,hydro_year_c)] %>%
    .[,.(beta = list(coef(zyp.sen(vpd15~hydro_year_c)))),by=.(x,y)] %>%
```

```
    .[,`:=`(b0=unlist(beta)[1],
            b1=unlist(beta)[2]), by=.(x,y)]
)
```

```
##    user  system elapsed
##  28.300  15.309  12.131
```

Regression Thiel Sen: sqrt(VPD)

```
system.time(
  lt_vsq_sen <- dat[date>=ymd("1981-12-01")][date<=ymd("2019-11-30")] %>%
    .[,`:=`(hydro_year_c = hydro_year-1982)] %>%
    .[is.na(vpd15)==F] %>%
    .[,.(vpd15 = mean(sqrt(vpd15),na.rm=TRUE)), by=.(x,y,hydro_year_c)] %>%
    .[,.(beta = list(coef(zyp.sen(vpd15~hydro_year_c)))),by=.(x,y)] %>%
    .[,`:=`(b0=unlist(beta)[1],
            b1=unlist(beta)[2]), by=.(x,y)]
)
```

```
##    user  system elapsed
##  30.678  14.863  14.757
```

Construct relative increase of VPD in 2019

```
# mavpd15 is annual mean from 1982:2011
pred_vpd_e1 <- merge(unique(dat_annual[,.(x,y,mavpd15)]), lt_v_sen[,.(x,y,b0)], by=c("x","y")) %>%
  .[,`:=`(e='e1',
          frac_vpd_anom = (b0-mavpd15)/mavpd15)]

pred_vpd_e2 <- lt_v_sen[,`:=`(vpd15_0 = b0,
                              vpd15_1 = b0+16*b1,
                              vpd15_2 = b0+37*b1)] %>%
  merge(unique(dat_annual[,.(x,y,mavpd15)]), ., by=c("x","y")) %>%
  .[,`:=`(frac_vpd_anom = (vpd15_2-mavpd15)/mavpd15,
          e='e2')]
```

Percent increase of VPD, NDVI, Ca

```
lt_v_sen <- lt_v_sen[b0 > 0]
dVPD_VPD_sen <- mean(37*lt_v_sen$b1,na.rm=TRUE)/mean(lt_v_sen$b0,na.rm=TRUE)

# Actual percent increase in NDVI
lt_ndvi_year <- lt_ndvi_year[b0 > 0]
dNDVI_NDVI_rlm <- mean(37*lt_ndvi_year$b1,na.rm=TRUE)/mean(lt_ndvi_year$b0,na.rm=TRUE)

# Percent increase in Ca
dCa_Ca <-
```

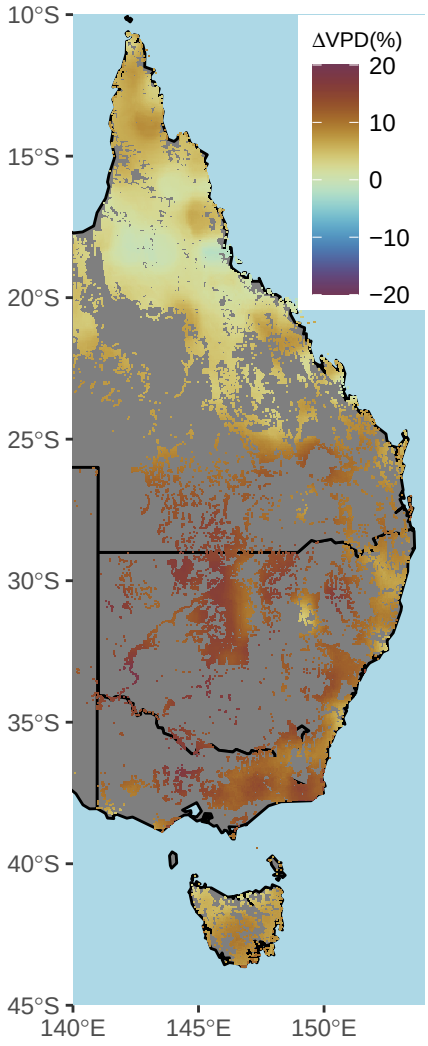
```
  diff(range(mlo[date>=ymd("1981-12-01") & date <= ymd("2019-08-30")]$co2_trend))/
  mean(mlo[date>=ymd("1981-12-01") & date <= ymd("1982-11-01")]$co2_trend)
dCa_Ca_e1 <-
  diff(range(mlo[date>=ymd("1981-12-01") & date <= ymd("2000-11-30")]$co2_trend))/
  mean(mlo[date>=ymd("1981-12-01") & date <= ymd("1982-11-01")]$co2_trend)
dCa_Ca_e2 <-
  diff(range(mlo[date>=ymd("2000-12-01") & date <= ymd("2019-08-30")]$co2_trend))/
  mean(mlo[date>=ymd("2000-12-01") & date <= ymd("2001-11-01")]$co2_trend)


# Expected WUE related increase using even split of WUE benefit (Donohue 2013)
0.5*(dCa_Ca - 0.5*dVPD_VPD_sen)
```

```
## [1] 0.08850517
```

Plotting —

Relative change in VPD from 1982-2019

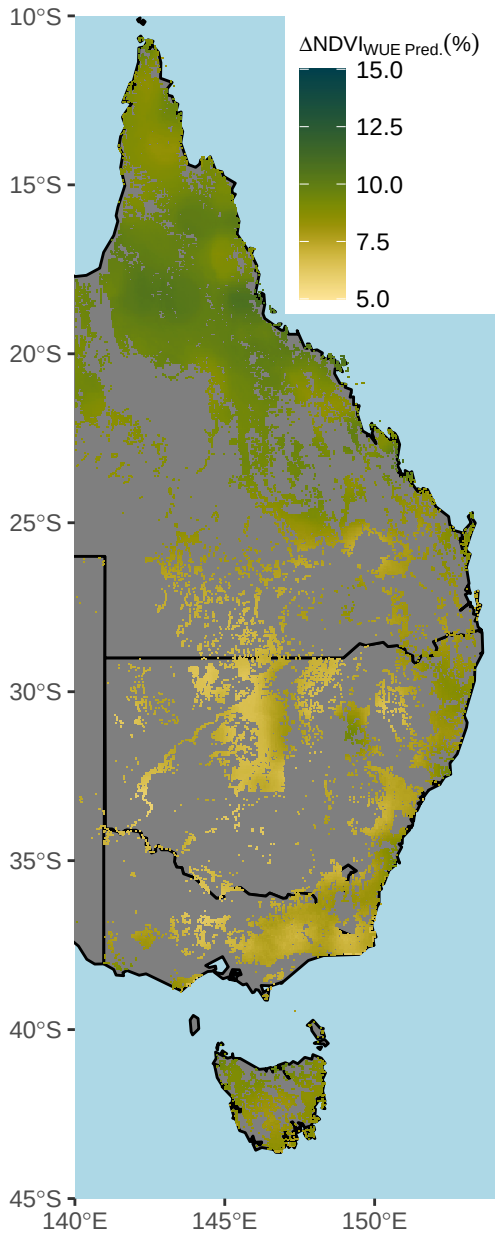
```
p_vpd_sen <- lt_v_sen %>%
  as_tibble() %>%
  filter(between(b1,-0.1,0.1)) %>%
  filter(b0 > 0) %>%
  ggplot(data=.,aes(x,y,fill=100*38*b1/b0))+
  geom_sf(data=oz_poly, inherit.aes = F, fill='gray50',color='black')+
  geom_tile()+
  scale_x_continuous(breaks=seq(140,154,by=5))+
  coord_sf(xlim = c(140,154),
           ylim = c(-45,-10), expand = FALSE)+
  labs(x=NULL,y=NULL)+
  scico::scale_fill_scico(expression(paste(Delta,"VPD(%)")),
                          palette ='roma0', direction=-1,
                          limits=c(-20,20),
                          oob=scales::squish)+
  # scale_fill_viridis_c(expression(paste(Delta,"VPD(%)")),
  #   option='A', limits=c(0,20), oob=scales::squish)+
  theme(panel.background = element_rect(fill='lightblue'),
        panel.grid = element_blank(),
        legend.title = element_text(size=8),
        legend.position = c(1,1),
        legend.justification = c(1,1)); p_vpd_sen
```



WUE pred using 0.5 allocation to foliar

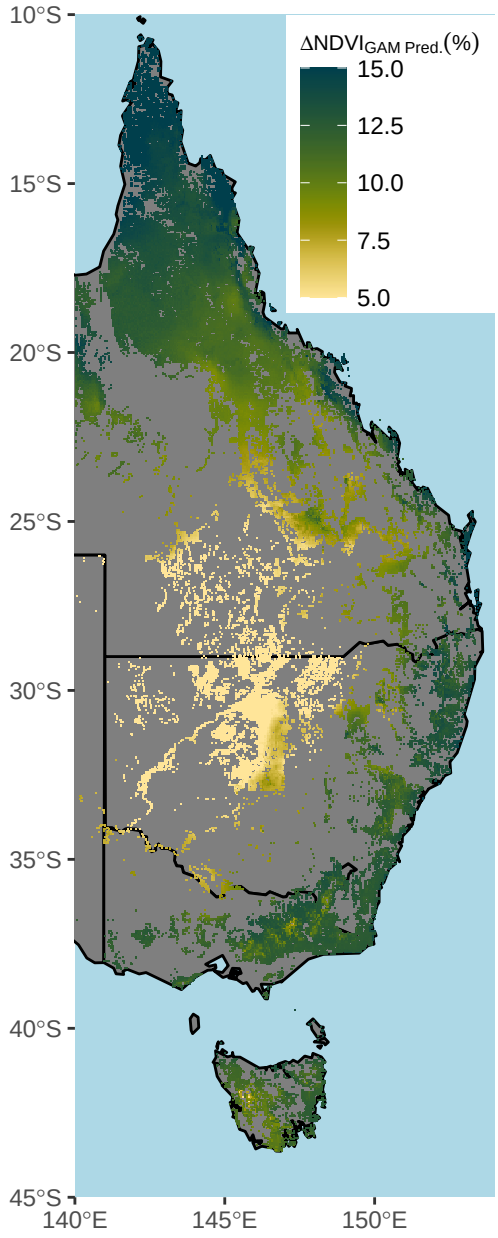
```
p_wue_sen <- lt_v_sen %>%
  as_tibble() %>%
  filter(b0 > 0) %>% # remove bad fits (don't really exist)
  filter(between(b1,-0.05,0.05)) %>% # remove bad fits (don't really exist)
  mutate(dVPD_VPD = b1*37/b0) %>% # relative fractional change in VPD
  mutate(expectation = 0.5*(dCa_Ca - 0.5*dVPD_VPD)) %>% # 0.5 alloc WUE pred
  # pull(expectation) %>% quantile(., c(0.01,0.99))
  ggplot(data=.,aes(x,y,fill=expectation*100))+
  geom_sf(data=oz_poly, inherit.aes = F, fill='gray50',color='black')+
  geom_tile()+
  scale_x_continuous(breaks=seq(140,154,by=5))+
  coord_sf(xlim = c(140,154),
           ylim = c(-45,-10), expand = FALSE)+
  labs(x=NULL,y=NULL)+
  scico::scale_fill_scico(expression(paste(Delta*NDVI[WUE~Pred.]("%"))),
                          palette = 'bamako',
```

```
                          direction = -1,
                          limits=c(5,15), #na.value = 'red',
                          oob=scales::squish
)+
theme(panel.background = element_rect(fill='lightblue'),
      panel.grid = element_blank(),
      legend.title = element_text(size=8),
      legend.position = c(1,1),
      legend.justification = c(1,1)); p_wue_sen
```



GAM estimated CO[2] contribution to NDVI where the effect of changes in P and PET are removed and the effect of increasing VPD is retained.

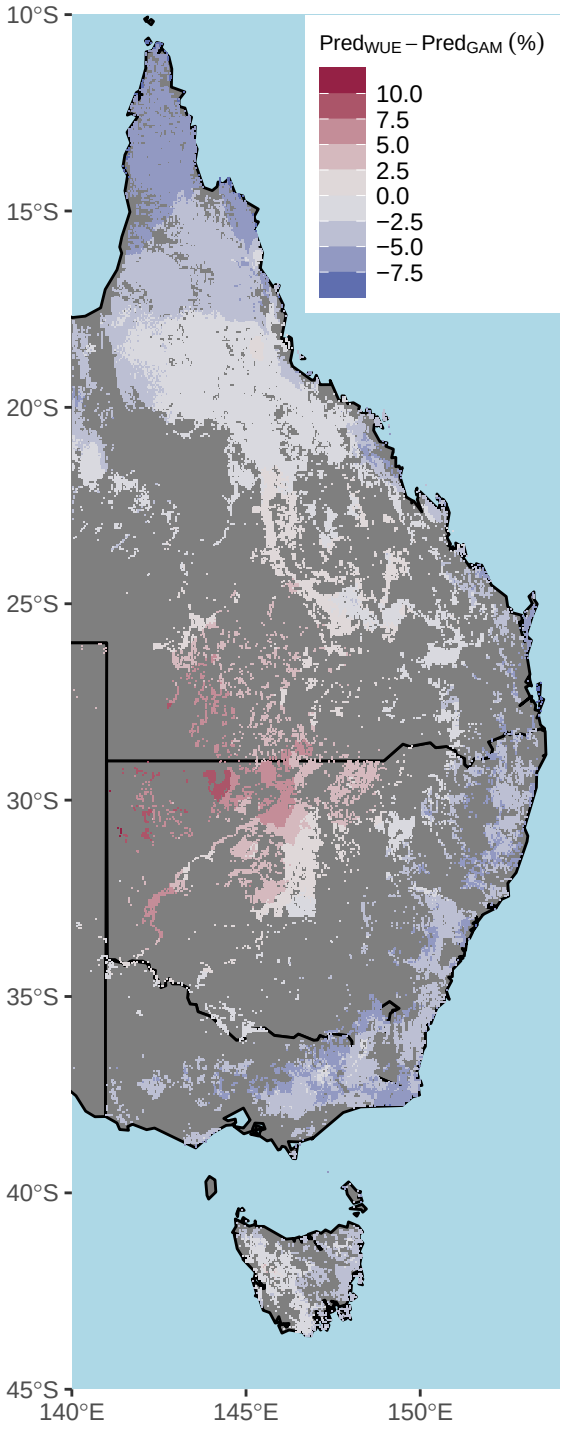
```
p_gam_pred <- bind_rows(tibble(pred_vpd_e1) %>% select(x,y,e,frac_vpd_anom),
                        tibble(pred_vpd_e2) %>% select(x,y,e,frac_vpd_anom)) %>%
  inner_join(., tibble(co2=c(340.8,411.9),e=c('e1','e2')),
             by=c('e')) %>%
  mutate(frac_p_anom=0,
         frac_ppet_anom=0,
         frac_pet_anom=0,
         # frac_vpd_anom=0,
         epoch=0) %>%
  inner_join(., tibble(unique(dat_annual[,.(x,y,map,mapet,mappet,mavpd15)])), by=c('x','y')) %>%
  inner_join(., kop,by=c('x','y')) %>%
  mutate(pred = predict(g3_full,newdata=.,type='response')) %>%
  filter(is.na(pred)==F) %>%
  select(x,y,e,frac_vpd_anom,pred,co2) %>%
  pivot_wider(., names_from=c('e'), values_from=c('pred','co2','frac_vpd_anom')) %>%
  mutate(dNDVI = 100*((pred_e2-pred_e1)/pred_e1)) %>%
  mutate(d_vpd = frac_vpd_anom_e2 - frac_vpd_anom_e1) %>%
  ggplot(data=.,aes(x,y,fill=dNDVI))+
  geom_sf(data=oz_poly, inherit.aes = F, fill='gray50',color='black')+
  geom_tile()+
  scale_x_continuous(breaks=seq(140,154,by=5))+
  coord_sf(xlim = c(140,154),
           ylim = c(-45,-10), expand = FALSE)+
  labs(x=NULL,y=NULL)+
  scico::scale_fill_scico(expression(paste(Delta*NDVI[GAM~Pred.]("%"))),
                          palette = 'bamako',
                          direction = -1,
                          limits=c(5,15), #na.value = 'red',
                          oob=scales::squish
  )+
  theme(panel.background = element_rect(fill='lightblue'),
        panel.grid = element_blank(),
        legend.title = element_text(size=8),
        legend.position = c(1,1),
        legend.justification = c(1,1)); p_gam_pred
```



Plot residuals between 0.5*WUE and GAM

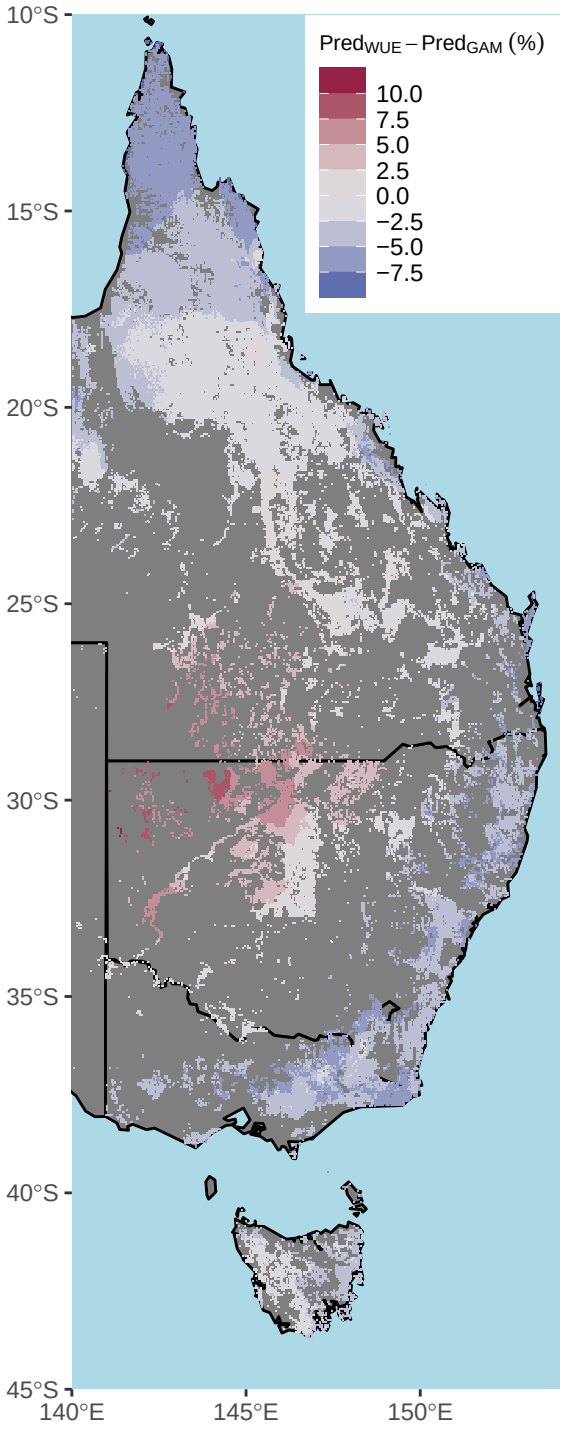
```
p_res <- inner_join({lt_v_sen %>%
    as_tibble() %>%
    filter(b0 > 0) %>%
    filter(between(b1,-0.05,0.05)) %>%
    mutate(dVPD_VPD = b1*37/b0) %>%
    mutate(expectation = 100*0.5*(dCa_Ca - 0.5*dVPD_VPD)) %>%
    select(x,y,expectation)},
    {bind_rows(tibble(pred_vpd_e1) %>% select(x,y,e,frac_vpd_anom),
               tibble(pred_vpd_e2) %>% select(x,y,e,frac_vpd_anom)) %>%
        inner_join(., tibble(co2=c(340.8,411.9),e=c('e1','e2')),
```

```
                  by=c('e')) %>%
      mutate(frac_p_anom=0,
             frac_ppet_anom=0,
             frac_pet_anom=0,
             # frac_vpd_anom=0,
             epoch=0) %>%
      inner_join(., tibble(unique(dat_annual[,.(x,y,map,mapet,mappet,mavpd15)])), by=c('x','y')) %>%
      inner_join(., kop,by=c('x','y')) %>%
      mutate(pred = predict(g3_full,newdata=.,type='response')) %>%
      filter(is.na(pred)==F) %>%
      select(x,y,e,frac_vpd_anom,pred,co2) %>%
      pivot_wider(., names_from=c('e'), values_from=c('pred','co2','frac_vpd_anom')) %>%
      mutate(dNDVI = 100*((pred_e2-pred_e1)/pred_e1)) %>%
      mutate(d_vpd = frac_vpd_anom_e2 - frac_vpd_anom_e1) %>%
      select(x,y,dNDVI)}, by=c('x','y')) %>%
ggplot(data=.,aes(x,y,fill=expectation-dNDVI))+
geom_sf(data=oz_poly, inherit.aes = F, fill='gray50',color='black')+
geom_tile()+
scale_x_continuous(breaks=seq(140,154,by=5))+
coord_sf(xlim = c(140,154),
         ylim = c(-45,-10), expand = FALSE)+
labs(x=NULL,y=NULL,fill=expression(paste(Pred[WUE]-Pred[GAM]~('%'))))+
scale_fill_binned_diverging(palette='Blue-Red',rev=F,n.breaks=10,
                            # cmax=90,
                            p1=1.5,p2=1.5)+
# scico::scale_fill_scico(expression(paste(Delta*NDVI[WUE~Pred.]("%")-Delta*NDVI[GAM~Pred.]("%"))),
#                         palette = 'roma',
#                         direction = -1,
#                         limits=c(-15,15), #na.value = 'red',
#                         oob=scales::squish
# )+
guides(fill=guide_colorbar())+
theme(panel.background = element_rect(fill='lightblue'),
      panel.grid = element_blank(),
      legend.title = element_text(size=8),
      legend.position = c(1,1),
      legend.justification = c(1,1)); p_res
```




```
p_res <- inner_join({lt_v_sen %>%
    as_tibble() %>%
    filter(b0 > 0) %>%
    filter(between(b1,-0.05,0.05)) %>%
    mutate(dVPD_VPD = b1*37/b0) %>%
    mutate(expectation = 100*0.5*(dCa_Ca - 0.5*dVPD_VPD)) %>%
    select(x,y,expectation)},
    {bind_rows(tibble(pred_vpd_e1) %>% select(x,y,e,frac_vpd_anom),
```

```
            tibble(pred_vpd_e2) %>% select(x,y,e,frac_vpd_anom)) %>%
      inner_join(., tibble(co2=c(340.8,411.9),e=c('e1','e2')),
                 by=c('e')) %>%
      mutate(frac_p_anom=0,
             frac_ppet_anom=0,
             frac_pet_anom=0,
             # frac_vpd_anom=0,
             epoch=0) %>%
      inner_join(., tibble(unique(dat_annual[,.(x,y,map,mapet,mappet,mavpd15)])), by=c('x','y')) %>%
      inner_join(., kop,by=c('x','y')) %>%
      mutate(pred = predict(g3_full,newdata=.,type='response')) %>%
      filter(is.na(pred)==F) %>%
      select(x,y,e,frac_vpd_anom,pred,co2) %>%
      pivot_wider(., names_from=c('e'), values_from=c('pred','co2','frac_vpd_anom')) %>%
      mutate(dNDVI = 100*((pred_e2-pred_e1)/pred_e1)) %>%
      mutate(d_vpd = frac_vpd_anom_e2 - frac_vpd_anom_e1) %>%
      select(x,y,dNDVI)}, by=c('x','y')) %>%
ggplot(data=.,aes(x,y,fill=expectation-dNDVI))+
geom_sf(data=oz_poly, inherit.aes = F, fill='gray50',color='black')+
geom_tile()+
scale_x_continuous(breaks=seq(140,154,by=5))+
coord_sf(xlim = c(140,154),
         ylim = c(-45,-10), expand = FALSE)+
labs(x=NULL,y=NULL,fill=expression(paste(Pred[WUE]-Pred[GAM]~('%'))))+
scale_fill_binned_diverging(palette='Blue-Red',rev=F,n.breaks=10,
                            # cmax=90,
                            p1=1.5,p2=1.5)+
# scico::scale_fill_scico(expression(paste(Delta*NDVI[WUE~Pred.]("%")-Delta*NDVI[GAM~Pred.]("%"))),
#                         palette = 'roma',
#                         direction = -1,
#                         limits=c(-15,15), #na.value = 'red',
#                         oob=scales::squish
# )+
guides(fill=guide_colorbar())+
theme(panel.background = element_rect(fill='lightblue'),
      panel.grid = element_blank(),
      legend.title = element_text(size=8),
      legend.position = c(1,1),
      legend.justification = c(1,1)); p_res
```

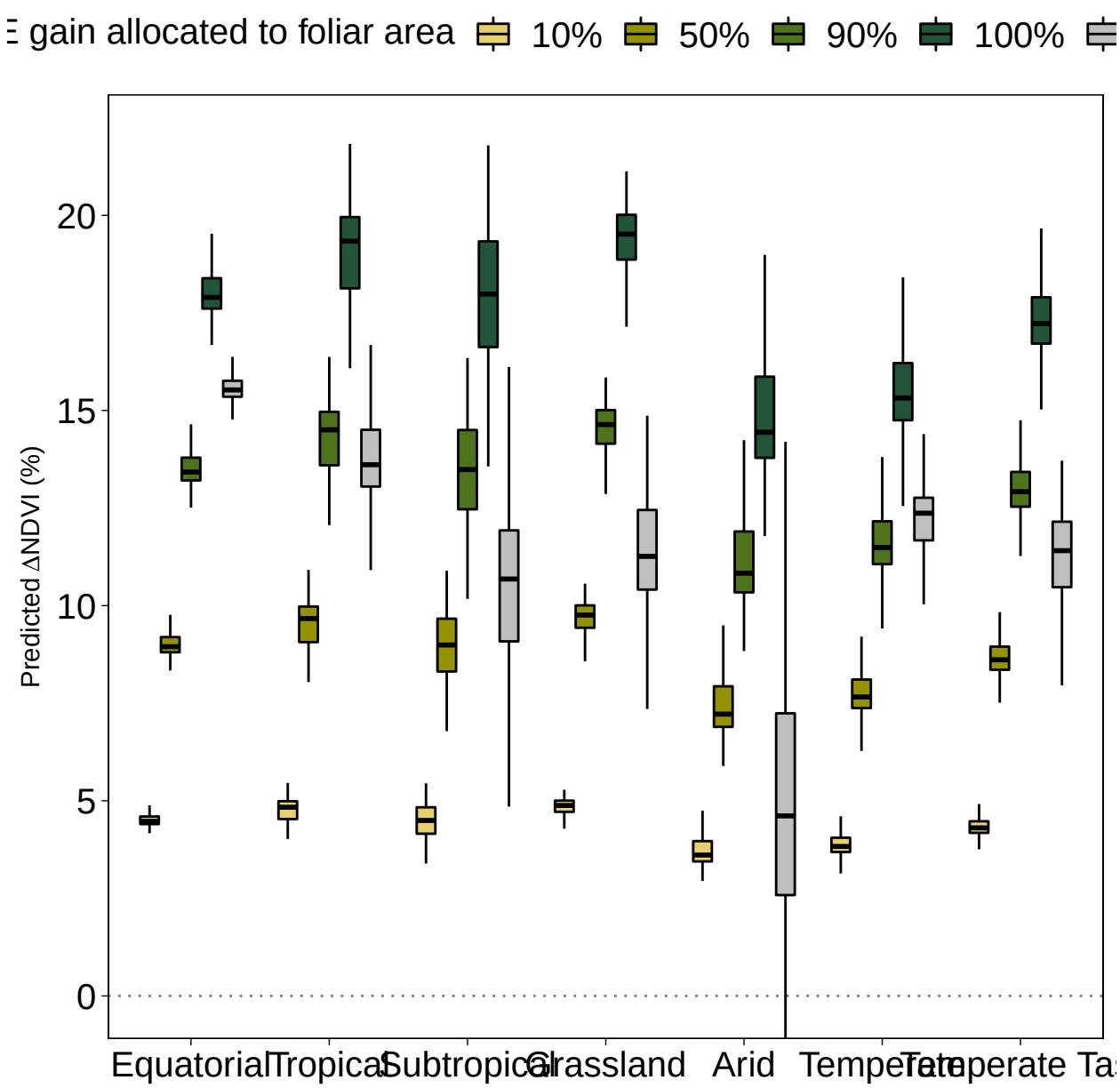


Boxplots WUE preds. (0.25,0.5,0.75,1 & GAM estimate

```
vec_cols <- scico::scico(n=4, palette = 'bamako',
                         direction = -1, end=0.9, begin=0.2)
p_box <- bind_rows(tibble(pred_vpd_e1) %>% select(x,y,e,frac_vpd_anom),
                   tibble(pred_vpd_e2) %>% select(x,y,e,frac_vpd_anom)) %>%
```

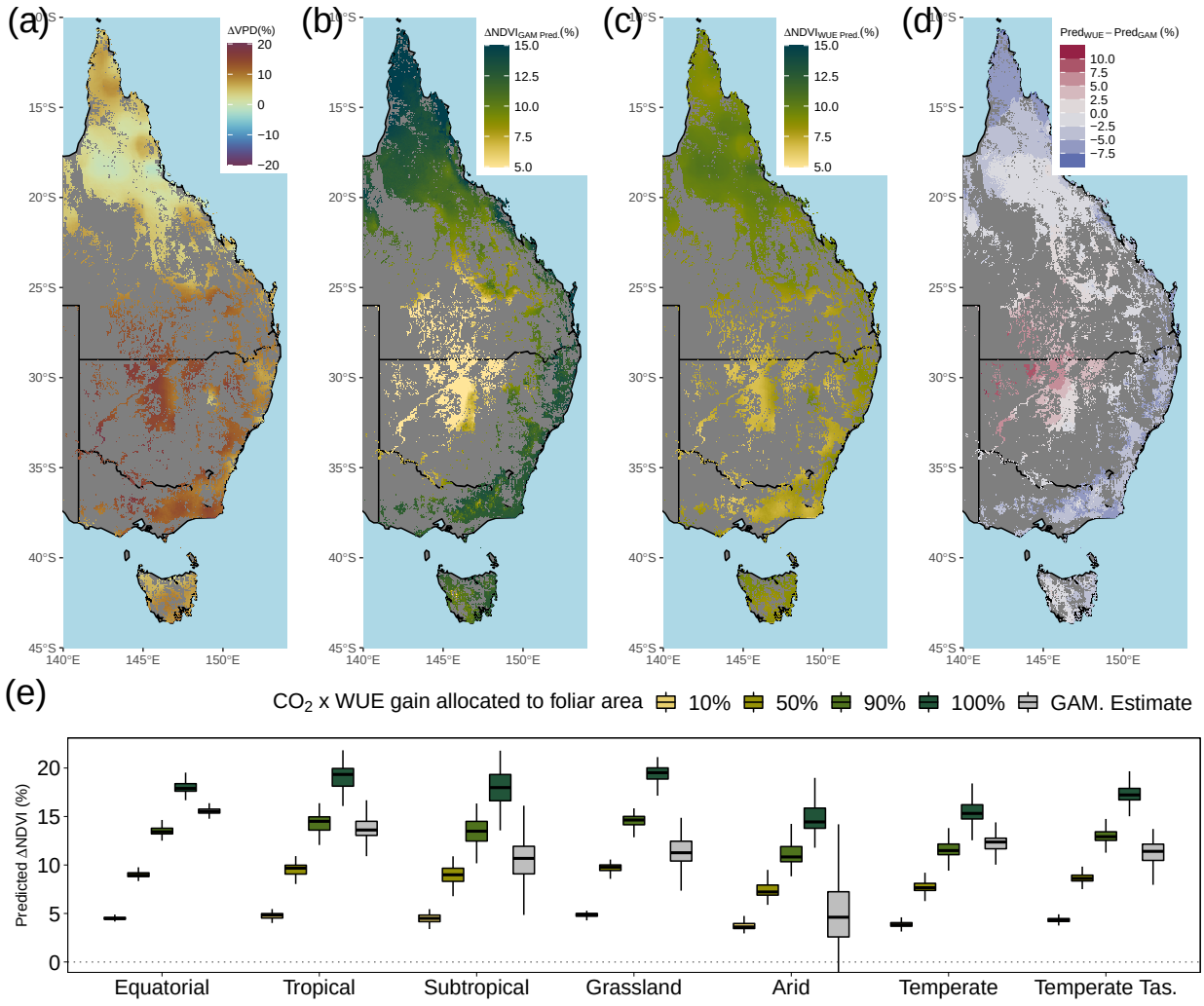
```
inner_join(., tibble(co2=c(340.8,411.9),e=c('e1','e2')),
           by=c('e')) %>%
mutate(frac_p_anom=0,
       frac_ppet_anom=0,
       frac_pet_anom=0,
       epoch=0) %>%
inner_join(., tibble(unique(dat_annual[,.(x,y,map,mapet,mappet,mavpd15)])), by=c('x','y')) %>%
inner_join(., kop,by=c('x','y')) %>%
mutate(pred = predict(g3_full,newdata=.,type='response')) %>%
filter(is.na(pred)==F) %>%
select(x,y,e,frac_vpd_anom,pred,co2) %>%
pivot_wider(., names_from=c('e'), values_from=c('pred','co2','frac_vpd_anom')) %>%
mutate(dNDVI = 100*((pred_e2-pred_e1)/pred_e1)) %>%
inner_join(., {
  lt_v_sen %>%
    as_tibble() %>%
    filter(b0 > 0) %>%
    filter(between(b1,-0.05,0.05)) %>%
    mutate(dVPD_VPD = b1*37/b0) %>%
    mutate(wue25 = 100*(0.25*(dCa_Ca - 0.5*dVPD_VPD)),
           wue50 = 100*(0.5*(dCa_Ca - 0.5*dVPD_VPD)),
           wue75 = 100*(0.75*(dCa_Ca - 0.5*dVPD_VPD)),
           wue100 = 100*(1*(dCa_Ca - 0.5*dVPD_VPD))) %>%
    select(x,y,wue25,wue50,wue75,wue100)
},by=c("x","y")) %>%
inner_join(., kop, by=c('x','y')) %>%
mutate(res25 = dNDVI-wue25,
       res50 = dNDVI-wue50,
       res75 = dNDVI-wue75
) %>%
select(zone,
       wue25,
       wue50,
       wue75,
       wue100,
       dNDVI
) %>%
gather(-zone,key='allocation',
       value='res') %>% #ggplot(data=.,aes(value,y=zone,fill=key))+geom_boxplot()
mutate(allocation = factor(allocation,
                           levels=c("wue25","wue50","wue75","wue100","dNDVI"),
                           labels = c('10%','50%','90%','100%','GAM. Estimate'),
                           ordered = TRUE)) %>%
ggplot(data=.,aes(y=res,
                  x=zone,
                  fill=allocation))+
geom_hline(aes(yintercept=0),color='grey50',lty=3)+
geom_boxplot(#draw_quantiles = c(0.25,0.5,0.75),
  # trim=TRUE,
  outlier.colour = NA,
  color='black')+
# scico::scale_fill_scico_d(
#   palette='bamako',direction = -1, begin = 0.2,end=0.95,
```

```
  #   guide=guide_legend(title.position = 'top'))+
  scale_fill_manual(#"% WUE benefit allocated to foliar area:",
    values=c(vec_cols[1], vec_cols[2], vec_cols[3], vec_cols[4],"grey"))+
  labs(x=NULL,
       y=expression(paste(Predicted~Delta*NDVI," (%)")),
       fill=expression(paste(CO[2]*' x '*WUE,' gain allocated to foliar area')))+
  scale_x_discrete(limits=(structure(c(1L,2L,3L,4L,5L,6L,7L),# c(5L, 4L, 6L, 2L, 1L, 3L, 7L),
                                     .Label = c("Equatorial",
                                                "Tropical",
                                                "Subtropical",
                                                "Grassland",
                                                "Arid",
                                                "Temperate",
                                                "Temperate Tas."),
                                     class = c("ordered", "factor"))))+
  coord_cartesian(ylim = c(0,22))+
  # coord_flip()+
  theme_linedraw()+
  guides(fill=guide_legend(title.position='left'))+
  theme(legend.position = 'top',
        legend.justification = c(1,1),
        legend.text = element_text(size=15),
        legend.title = element_text(size=15),
        # legend.key.height = unit(0.2,'cm'),
        axis.text = element_text(size=15),
        panel.grid = element_blank());
p_box
```



merge figures

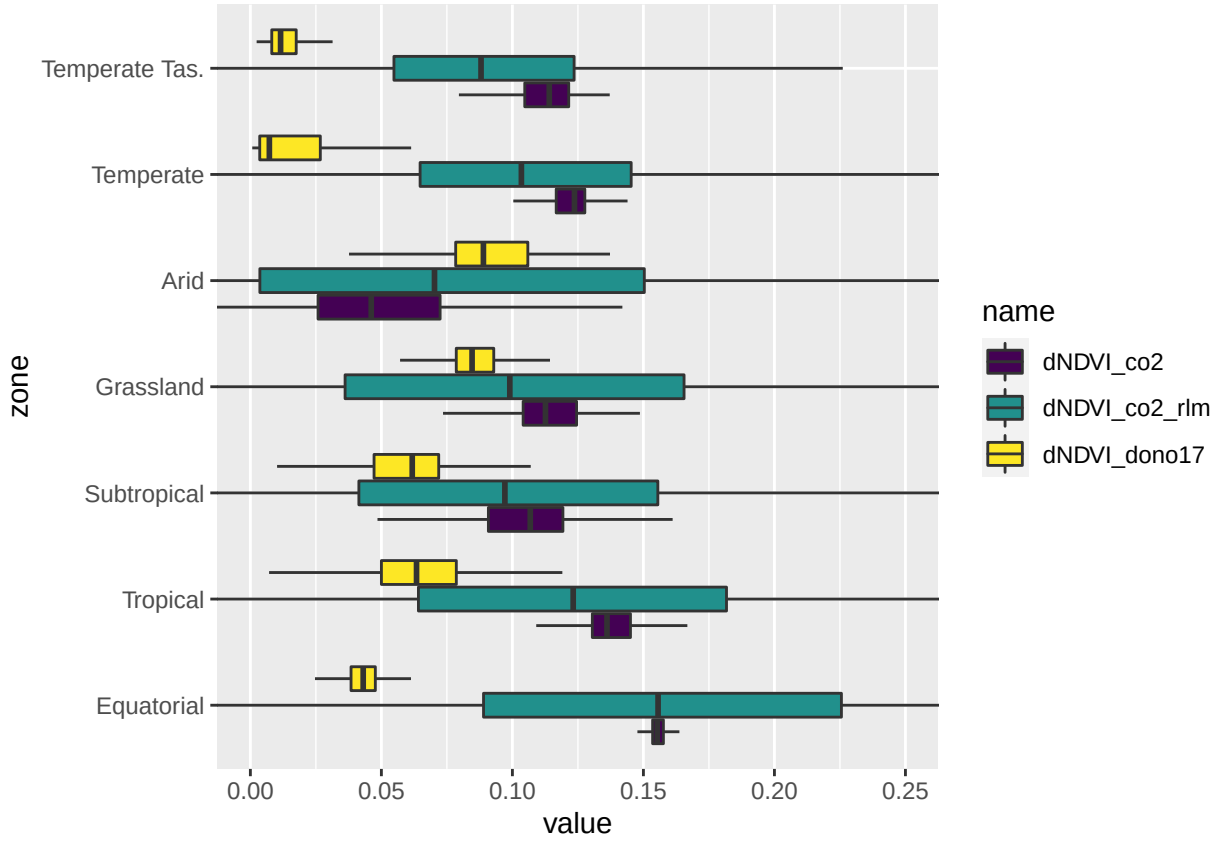
```
library(cowplot)
p_left <- ggdraw(p_vpd_sen)+draw_label(label='(a)', x=0.07,y=0.985,size = 25)
p_mid1 <- ggdraw(p_gam_pred)+draw_label(label='(b)',x=0.05,y=0.985,size=25)
p_mid2 <- ggdraw(p_wue_sen)+draw_label(label='(c)', x=0.05,y=0.985, size=25)
p_right <- ggdraw(p_res)+draw_label(label = '(d)', x=0.05,y=0.985,size=25)
p_bottom <- ggdraw(p_box)+draw_label(label = '(e)', x=0.015,y=0.95,size=25)
cp_r <- cowplot::plot_grid(p_left,p_mid1,p_mid2,p_right,
                           nrow = 1,
                           rel_widths = c(1,1,1,1))
cp_r/p_bottom+patchwork::plot_layout(heights = c(1,0.5))
```



Prepare PETA prediction comparison

```
# PETA prediction
dono17 <- {lt_vsq_sen %>%
    as_tibble() %>%
    rowwise() %>%
    filter(b0 > 0) %>%
    # filter(between(b1,-0.05,0.05)) %>%
    mutate(dVPD_VPD = b1*37/b0) %>%
    mutate(wue_expectation = 0.5*(dCa_Ca - 0.5*dVPD_VPD)) %>% # 0.5*WUE benefit
    mutate(peta = (1+dCa_Ca)/(1+dVPD_VPD) -1) %>% # eq2 (Donohue 2017)
    # mutate(peta = (  (1+(dCa_Ca/341.16))/
    #                  (1+(dVPD_VPD/b0))) - 1) %>%
    select(x,y,wue_expectation,peta)} %>%
  inner_join(., lai,by=c('x','y')) %>% # join LAI
  mutate(alpha = 1-exp(-0.5*lai)) %>%  # eq 6
  mutate(beta = 1/(1+peta) - 1) %>% # eq 12
  mutate(dL = peta*(1-alpha)**2) %>%  # eq 9
  mutate(dE = -beta*alpha + beta)
```

```r
# GAM estimate
gest <- bind_rows(tibble(pred_vpd_e1) %>% select(x,y,e,frac_vpd_anom),
                          tibble(pred_vpd_e2) %>% select(x,y,e,frac_vpd_anom)) %>%
    inner_join(., tibble(co2=c(340.8,411.9),e=c('e1','e2')),
               by=c('e')) %>%
    mutate(frac_p_anom=0,
           frac_ppet_anom=0,
           frac_pet_anom=0,
           # frac_vpd_anom=0,
           epoch=0) %>%
    inner_join(., tibble(unique(dat_annual[,.(x,y,map,mapet,mappet,mavpd15)])), by=c('x','y')) %>%
    inner_join(., kop,by=c('x','y')) %>%
    mutate(pred = predict(g3_full,newdata=.,type='response')) %>%
    filter(is.na(pred)==F) %>%
    select(x,y,e,frac_vpd_anom,pred,co2) %>%
    pivot_wider(., names_from=c('e'), values_from=c('pred','co2','frac_vpd_anom')) %>%
    mutate(dNDVI_co2 = ((pred_e2-pred_e1)/pred_e1)) %>%
    mutate(d_vpd = frac_vpd_anom_e2 - frac_vpd_anom_e1) %>%
    select(x,y,dNDVI_co2)

# RLM estimates
lest <- bind_rows(tibble(pred_vpd_e1) %>% select(x,y,e,frac_vpd_anom),
                  tibble(pred_vpd_e2) %>% select(x,y,e,frac_vpd_anom)) %>%
  inner_join(., tibble(co2=c(340.8,411.9),e=c('e1','e2')),
             by=c('e')) %>%
  mutate(frac_p_anom=0,
         frac_ppet_anom=0,
         frac_pet_anom=0,
         epoch=0) %>%
  inner_join(., tibble(unique(dat_annual[,.(x,y,map,mapet,mappet,mavpd15)])), by=c('x','y')) %>%
  inner_join(., kop,by=c('x','y')) %>%
  inner_join(., rlm_ndvi_annual_co2_ppet_epoch,by=c("x","y")) %>%
  mutate(pred = b0 + b1*(co2-min_co2)+b2*frac_vpd_anom) %>%
  filter(is.na(pred)==F) %>%
  select(x,y,e,frac_vpd_anom,pred,co2) %>%
  pivot_wider(., names_from=c('e'), values_from=c('pred','co2','frac_vpd_anom')) %>%
  mutate(dNDVI_co2_rlm = ((pred_e2-pred_e1)/pred_e1)) %>%
  mutate(d_vpd = frac_vpd_anom_e2 - frac_vpd_anom_e1) %>%
  select(x,y,dNDVI_co2_rlm)

# boxplot
inner_join(dono17, gest) %>%
  inner_join(., lest) %>%
  inner_join(., dono17 %>% rename(dNDVI_dono17 = dL)) %>%
  inner_join(., kop) %>%
  pivot_longer(cols=starts_with("dNDVI")) %>%
  ggplot(data=.,aes(value,zone,fill=name))+
  geom_boxplot(outlier.colour = NA)+
  coord_cartesian(xlim=c(0,0.25),clip = 'on')+
  scale_fill_viridis_d()
```




```
gest %>% rename(GAM = dNDVI_co2) %>%
  inner_join(., lest %>% rename(RLM=dNDVI_co2_rlm)) %>%
  inner_join(., dono17 %>% rename(PETA = dL)) %>%
  inner_join(., kop) %>%
  pivot_longer(cols=c(starts_with("dNDVI"),'PETA','RLM','GAM')) %>%
  ggplot(data=.,aes(lai,value,color=name))+
  geom_point(data=. %>% filter(name %in% c("RLM","GAM")),
             aes(lai,value,color=name),
    alpha=0.03,size=0.5)+
  geom_smooth(method='gam',
              formula=y~s(x,bs='cs'),
              method.args=list(method='ML',
                               select=TRUE))+
  scale_color_manual(#option='D',end=0.6,direction = 1,
    values=c("darkgreen",
             "black",
             "blue"),
    breaks=c("PETA","GAM","RLM"),
    labels=c("PETA Prediction","GAM Estimate","RLM Estimate"),
    # drop=TRUE,
    na.translate=FALSE)+
  labs(x=expression(paste('Leaf Area Index ',(m**2/m**2))),
       color='',
       y=expression(paste(CO[2]~'effect on NDVI')))+
  coord_cartesian(ylim=c(0,0.25),
                  xlim=c(0,4.5),
```

```
                  expand=FALSE,
                  default = FALSE,
                  clip = 'on')+
  theme_linedraw()+
  theme(legend.position = c(1,1),
        legend.justification = c(0.99,0.99),
        legend.background = element_rect(fill=NA),
        panel.grid = element_blank())
```