DLMs and GAMs for Lake Washington data

```
#install.packages('marssTMB', repos = c('https://atsa-es.r-universe.dev', #'https://cloud.r-p
library(MARSS)
library(mgcv)
library(dplyr)
library(forecast)
library(ggplot2)
library(lubridate)
library(marssTMB)
library(paletteer)
library(ggpubr)
library(grid)
col<-paletteer_d("nationalparkcolors::Everglades")
```

Data

```
dat_lakewa <- readRDS("data_for_lakeWA_example.rds")
dat_aksalmon <- readRDS("data_for_AK_salmon_example.rds")

# renaming
dat_lakewa <- dplyr::rename(dat_lakewa,
                            response = y_adj,
                            driver = lagged_zp)
dat_aksalmon <- dplyr::rename(dat_aksalmon,
                              response = salmon_DFA,
                              driver = z_sst,
                              date = year)
```

DLM with time-varying intercepts and time-varying slopes

Define the model – this block is basically copied from the MARSS book (salmon survival case study). First for the Lake WA example:

```r
dat <- dat_lakewa
m <- 2
TT <- nrow(dat)
B <- diag(m)  ## 2x2; Identity
U <- matrix(0, nrow = m, ncol = 1)  ## 2x1; both elements = 0
Q <- matrix(list(0), m, m)  ## 2x2; all 0 for now
diag(Q)[1] <- c("q.alpha")
diag(Q)[2] <- c("q.beta")
Z <- array(NA, c(1, m, TT))  ## NxMxT; empty for now
Z[1, 1, ] <- rep(1, TT)  ## Nx1; 1's for intercept
Z[1, 2, ] <- dat$driver  ## Nx1; predictor variable
A <- matrix(0)  ## 1x1; scalar = 0
R <- matrix("r")  ## 1x1; scalar = r
## only need starting values for regr parameters
inits_list <- list(x0 = matrix(c(0, 0), nrow = m))
## list of model matrices & vectors
mod_list <- list(B = B, U = U, Q = Q, Z = Z, A = A, R = R)
# convert response to matrix
dat_mat <- matrix(dat$response, nrow = 1)
# fit the model -- crank up the maxit to ensure convergence
dlm_lakewa_1 <- MARSS(dat_mat, inits = inits_list, model = mod_list,
               control = list(maxit=4000), method="TMB")
```

```
MARSS fit is
Estimation method: TMB
Estimation converged in 42 iterations.
Log-likelihood: -143.7477
AIC: 297.4954   AICc: 297.9466

          Estimate
R.r       1.49e-01
Q.q.alpha 2.15e-01
Q.q.beta  2.07e-11
x0.X1     1.64e+00
x0.X2     6.92e-02
Initial states (x0) defined at t=0
```

```
Standard errors have not been calculated.
Use MARSSparamCIs to compute CIs and bias estimates.
```

```
# Fit a second model with just time-varying slope
diag(Q)[1] <- 1e-10
mod_list <- list(B = B, U = U, Q = Q, Z = Z, A = A, R = R)
dlm_lakewa_2 <- MARSS(dat_mat, inits = inits_list, model = mod_list,
               control = list(maxit=4000), method="TMB")
```

```
MARSS fit is
Estimation method: TMB
Estimation converged in 18 iterations.
Log-likelihood: -155.082
AIC: 318.164   AICc: 318.4625

         Estimate
R.r         0.226
Q.q.beta    0.235
x0.X1       0.498
x0.X2       1.422
Initial states (x0) defined at t=0

Standard errors have not been calculated.
Use MARSSparamCIs to compute CIs and bias estimates.
```

Now for the AK salmon example

```
dat <- dat_aksalmon
m <- 2
TT <- nrow(dat)
B <- diag(m)  ## 2x2; Identity
U <- matrix(0, nrow = m, ncol = 1)  ## 2x1; both elements = 0
Q <- matrix(list(0), m, m)  ## 2x2; all 0 for now
diag(Q)[1] <- c("q.alpha")
diag(Q)[2] <- c("q.beta")
Z <- array(NA, c(1, m, TT))  ## NxMxT; empty for now
Z[1, 1, ] <- rep(1, TT)  ## Nx1; 1's for intercept
Z[1, 2, ] <- dat$driver  ## Nx1; predictor variable
A <- matrix(0)  ## 1x1; scalar = 0
```

```r
R <- matrix("r")  ## 1x1; scalar = r
## only need starting values for regr parameters
inits_list <- list(x0 = matrix(c(0, 0), nrow = m))
## list of model matrices & vectors
mod_list <- list(B = B, U = U, Q = Q, Z = Z, A = A, R = R)
# convert response to matrix
dat_mat <- matrix(dat$response, nrow = 1)
# fit the model -- crank up the maxit to ensure convergence
dlm_aksalmon_1 <- MARSS(dat_mat, inits = inits_list, model = mod_list,
               control = list(maxit=4000), method="TMB")
```

```
MARSS fit is
Estimation method: TMB
Estimation converged in 34 iterations.
Log-likelihood: -67.30025
AIC: 144.6005   AICc: 145.7543

           Estimate
R.r        7.54e-09
Q.q.alpha  5.43e-01
Q.q.beta   1.27e-02
x0.X1     -3.52e+00
x0.X2      8.37e-01
Initial states (x0) defined at t=0

Standard errors have not been calculated.
Use MARSSparamCIs to compute CIs and bias estimates.
```

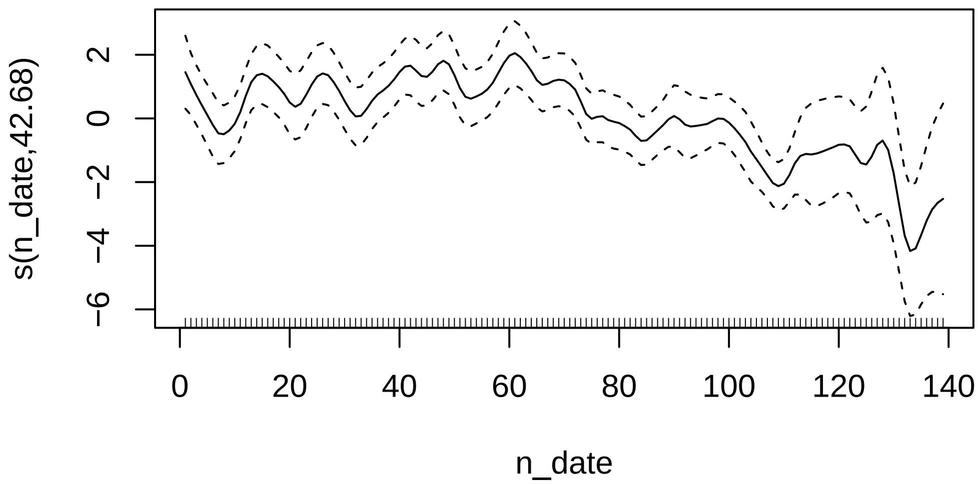
```r
# Fit a second model with just time-varying slope
diag(Q)[1] <- 1e-10
mod_list <- list(B = B, U = U, Q = Q, Z = Z, A = A, R = R)
dlm_aksalmon_2 <- MARSS(dat_mat, inits = inits_list, model = mod_list,
               control = list(maxit=4000), method="TMB")
```

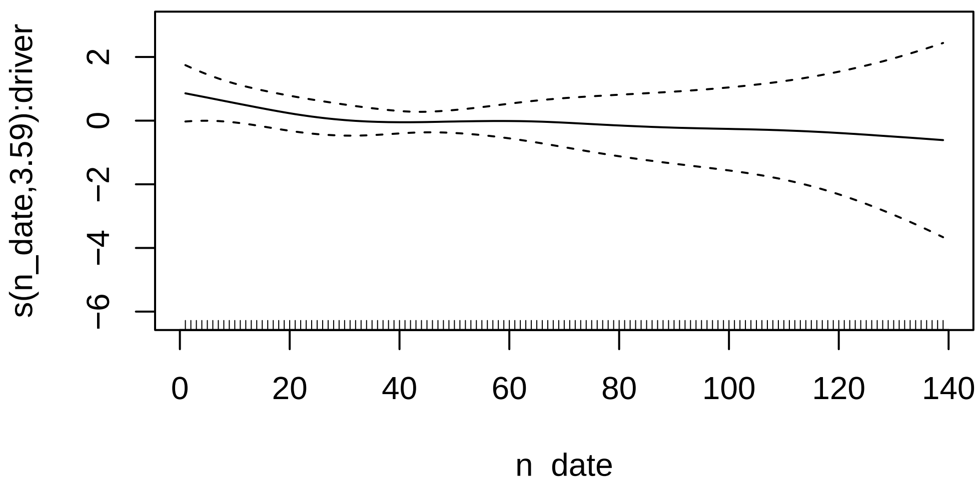
```
MARSS fit is
Estimation method: TMB
Estimation converged in 16 iterations.
Log-likelihood: -114.4286
AIC: 236.8572   AICc: 237.6119
```

```
         Estimate
R.r         2.143
Q.q.beta    0.365
x0.X1       1.269
x0.X2       6.343
Initial states (x0) defined at t=0

Standard errors have not been calculated.
Use MARSSparamCIs to compute CIs and bias estimates.
```

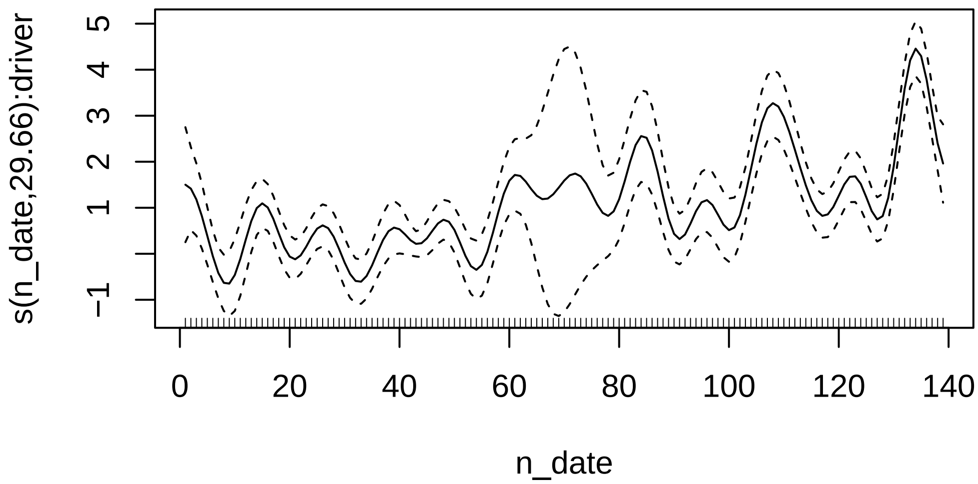
Time varying models with MGCV

```
K_WA <- round(0.99*nrow(dat_lakewa))
K_AK <- round(0.99*nrow(dat_aksalmon))
dat_lakewa$n_date <- seq(1,nrow(dat_lakewa))
gam_lakewa_1 <- gam(response ~ s(n_date,k=K_WA) + s(n_date, by = driver, bs = "gp",k=K_WA), d
gam_lakewa_2 <- gam(response ~ s(n_date, by = driver, bs = "gp",k=K_WA), data = dat_lakewa)

dat_aksalmon$n_date <- seq(1,nrow(dat_aksalmon))
gam_aksalmon_1 <- gam(response ~ s(n_date,k=K_AK) + s(n_date, by = driver, bs = "gp",k=K_AK)
gam_aksalmon_2 <- gam(response ~ s(n_date, by = driver, bs = "gp",k=K_AK), data = dat_aksalmo

# extract the smooth estimates with SEs to compare to DLM
plot_1 <- plot(gam_lakewa_1, seWithMean = TRUE, n = nrow(dat_lakewa))
```

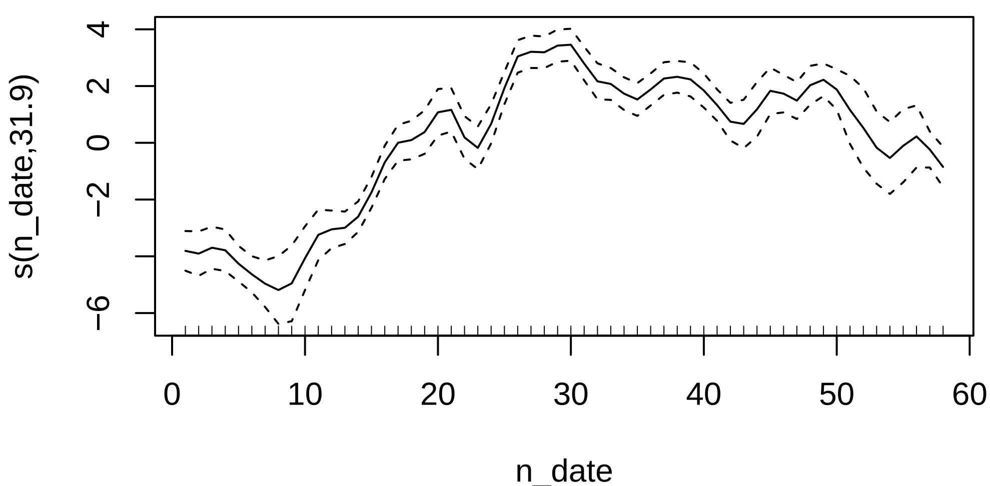


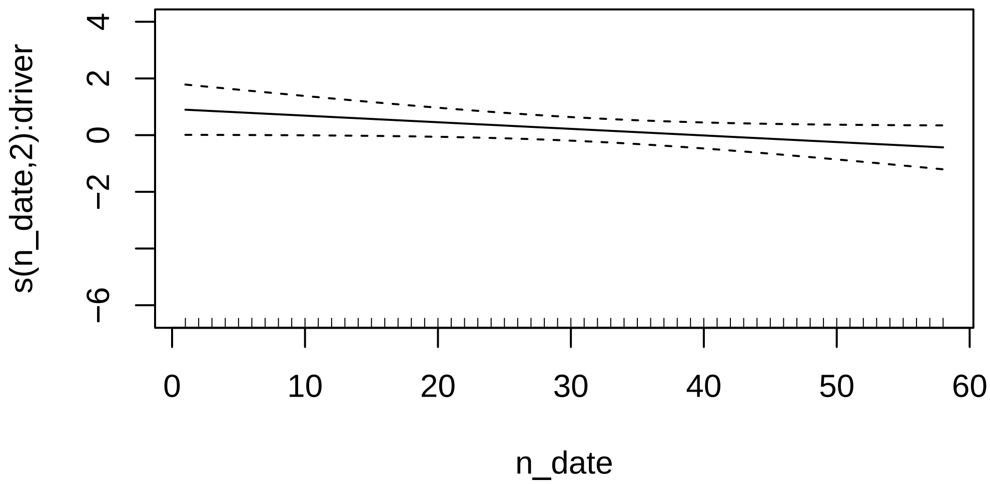



```
plot_2 <- plot(gam_lakewa_2, seWithMean = TRUE, n = nrow(dat_lakewa))
```

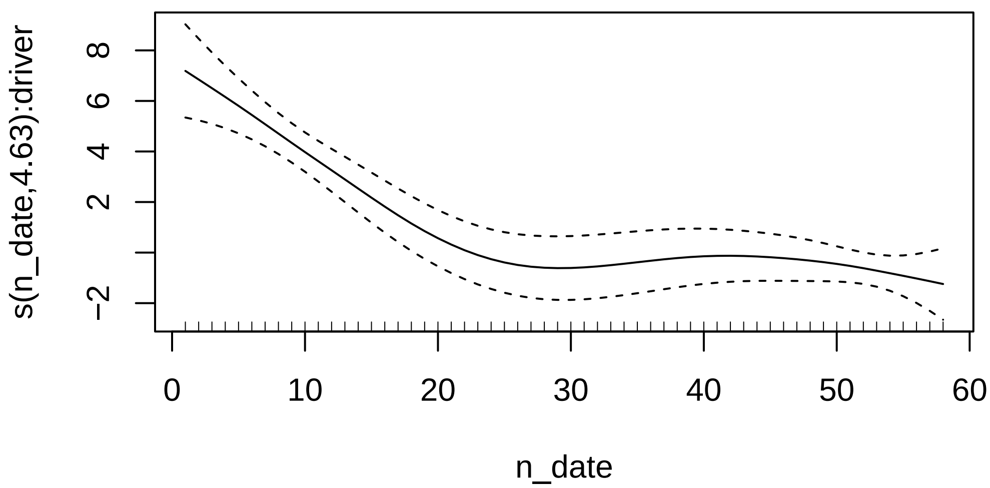



```r
plot_3 <- plot(gam_aksalmon_1, seWithMean = TRUE, n = nrow(dat_aksalmon))
```





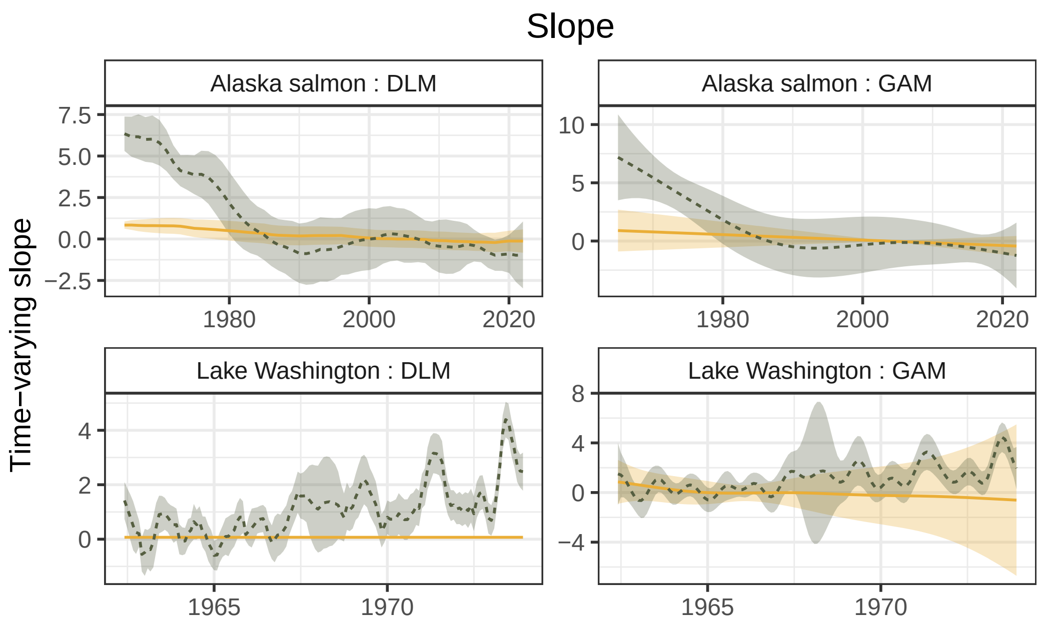

```
plot_4 <- plot(gam_aksalmon_2, seWithMean = TRUE, n = nrow(dat_aksalmon))
```



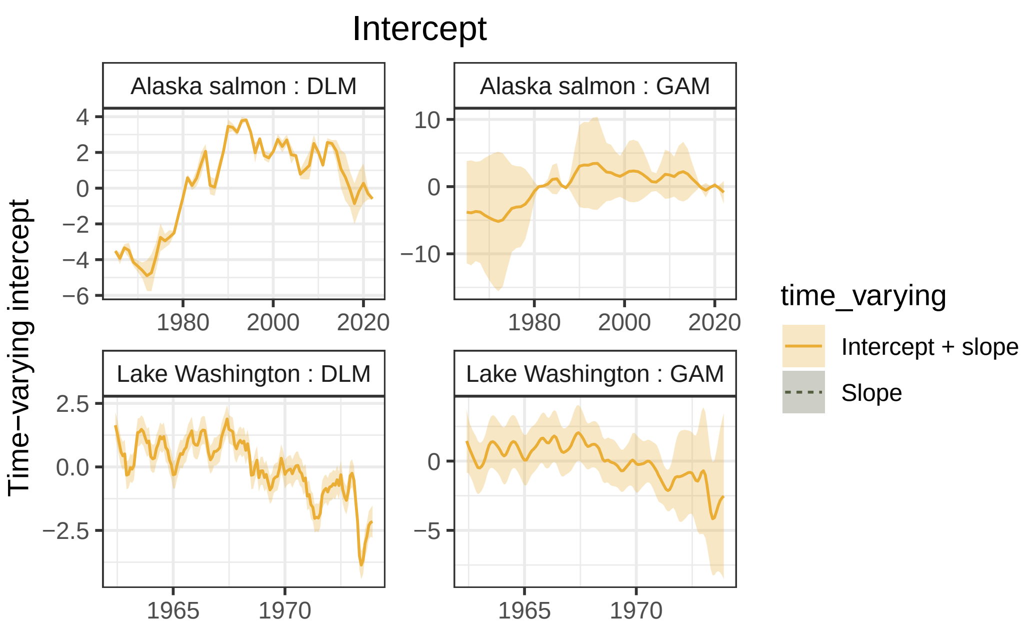
```
dat_aksalmon$date <- paste0(dat_aksalmon$date,"-01-01")

coef_dlm_lakeWA1 <- data.frame(date = dat_lakewa$date,
                        int_est = dlm_lakewa_1$states[1,],
                        int_se = dlm_lakewa_1$states.se[1,],
                        slope_est = dlm_lakewa_1$states[2,],
                        slope_se = dlm_lakewa_1$states.se[2,],
                        time_varying = "Intercept + slope",
                        dataset = "Lake Washington : DLM")
coef_dlm_lakeWA2 <- data.frame(date = dat_lakewa$date,
                        int_est = NA,
                        int_se = NA,
                        slope_est = dlm_lakewa_2$states[2,],
                        slope_se = dlm_lakewa_2$states.se[2,],
                        time_varying = "Slope",
                        dataset = "Lake Washington : DLM")
coef_dlm_AK1 <- data.frame(date = dat_aksalmon$date,
                        int_est = dlm_aksalmon_1$states[1,],
                        int_se = dlm_aksalmon_1$states.se[1,],
                        slope_est = dlm_aksalmon_1$states[2,],
                        slope_se = dlm_aksalmon_1$states.se[2,],
                        time_varying = "Intercept + slope",
                        dataset = "Alaska salmon : DLM")
coef_dlm_AK2 <- data.frame(date = dat_aksalmon$date,
                        int_est = NA,
                        int_se = NA,
                        slope_est = dlm_aksalmon_2$states[2,],
                        slope_se = dlm_aksalmon_2$states.se[2,],
                        time_varying = "Slope",
                        dataset = "Alaska salmon : DLM")


coef_gam_lakeWA1 <- data.frame(date = dat_lakewa$date,
                        int_est = plot_1[[1]]$fit,
                        int_se = plot_1[[1]]$se,
                        slope_est = plot_1[[2]]$fit,
                        slope_se = plot_1[[2]]$se,
                        time_varying = "Intercept + slope",
                        dataset = "Lake Washington : GAM")
coef_gam_lakeWA2 <- data.frame(date = dat_lakewa$date,
                        int_est = NA,
                        int_se = NA,
```

```
                        slope_est = plot_2[[1]]$fit,
                        slope_se = plot_2[[1]]$se,
                        time_varying = "Slope",
                        dataset = "Lake Washington : GAM")
coef_gam_AK1 <- data.frame(date = dat_aksalmon$date,
                        int_est = plot_3[[1]]$fit,
                        int_se = plot_3[[1]]$fit,
                        slope_est = plot_3[[2]]$fit,
                        slope_se = plot_3[[2]]$fit,
                        time_varying = "Intercept + slope",
                        dataset = "Alaska salmon : GAM")
coef_gam_AK2 <- data.frame(date = dat_aksalmon$date,
                        int_est = NA,
                        int_se = NA,
                        slope_est = plot_4[[1]]$fit,
                        slope_se = plot_4[[1]]$se,
                        time_varying = "Slope",
                        dataset = "Alaska salmon : GAM")

coefs <- rbind(coef_dlm_lakeWA1, coef_dlm_lakeWA2, coef_dlm_AK1, coef_dlm_AK2,
               coef_gam_lakeWA1, coef_gam_lakeWA2, coef_gam_AK1, coef_gam_AK2)

annotate_Slope<-data.frame(label=c("A.", "B.","E.", "F"),
                       date=c(1975,1975,dat_lakewa$date[25],dat_lakewa$date[25]),
                       y=c(7.5,10,4.5,7.75), time_varying=c("slope","slope","slope","slope")

slope_panel <-ggplot(coefs, aes(date, slope_est, group = time_varying, fill=time_varying, co
  geom_ribbon(aes(ymin=slope_est-2*slope_se, ymax = slope_est+2*slope_se), alpha=0.3, col = 
  geom_line(aes(lty=time_varying)) +
  ylab("Time-varying slope") +
  xlab("") + theme_bw() +
  ggtitle("Slope")+
  #ylim(c(-5,11))+
  facet_wrap(~ dataset, scale="free") +
  #scale_color_viridis_d(option="magma",begin=0.2, end=0.8, name = "Time-varying") +
  #scale_fill_viridis_d(option="magma",begin=0.2, end=0.8, name = "Time-varying") +
  scale_fill_manual(values=c(col[4],col[5]))+
  scale_colour_manual(values=c(col[4],col[5]))+
  guides(fill="none",colour="none",lty="none")+
  theme(
    strip.background = element_rect(fill = "white"),
    plot.title = element_text(hjust = 0.5)
```

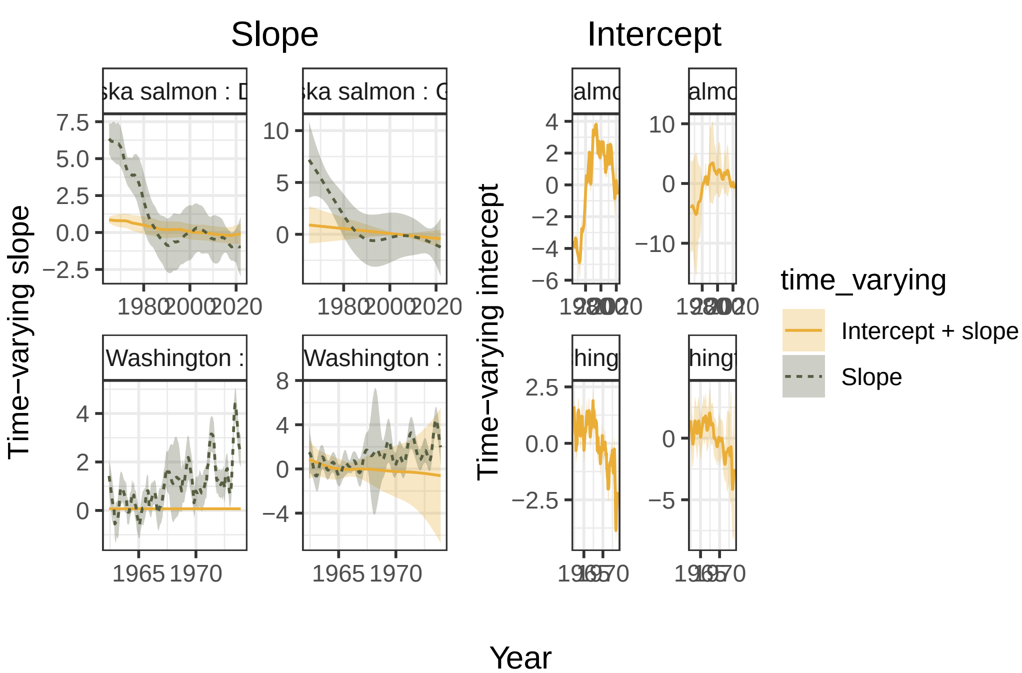
```
  )
slope_panel
```




```
int_panel <- ggplot(coefs, aes(date, int_est, group = time_varying, fill=time_varying, col=ti
  geom_ribbon(aes(ymin=int_est-2*int_se, ymax = int_est+2*int_se), alpha=0.3, col = NA) +
  geom_line(aes(lty=time_varying)) +
  ylab("Time-varying intercept") +
  xlab("") + theme_bw() +
  facet_wrap(~ dataset, scale="free") +
  ggtitle("Intercept")+
  #ylim(-20,10.5)+
  #scale_color_viridis_d(option="magma",begin=0.2, end=0.8, name = "Time-varying") +
  #scale_fill_viridis_d(option="magma",begin=0.2, end=0.8, name = "Time-varying") +
  scale_fill_manual(values=c(col[4],col[5]))+
  scale_colour_manual(values=c(col[4],col[5]))+

  theme(
    strip.background = element_rect(fill = "white"),
    plot.title = element_text(hjust = 0.5)
  )
int_panel
```




```
Fig4<-ggarrange(slope_panel,int_panel, nrow=1,ncol=2, widths=c(1,1.25))
annotate_figure(Fig4, bottom = textGrob("Year", gp = gpar(cex = 1)))
```



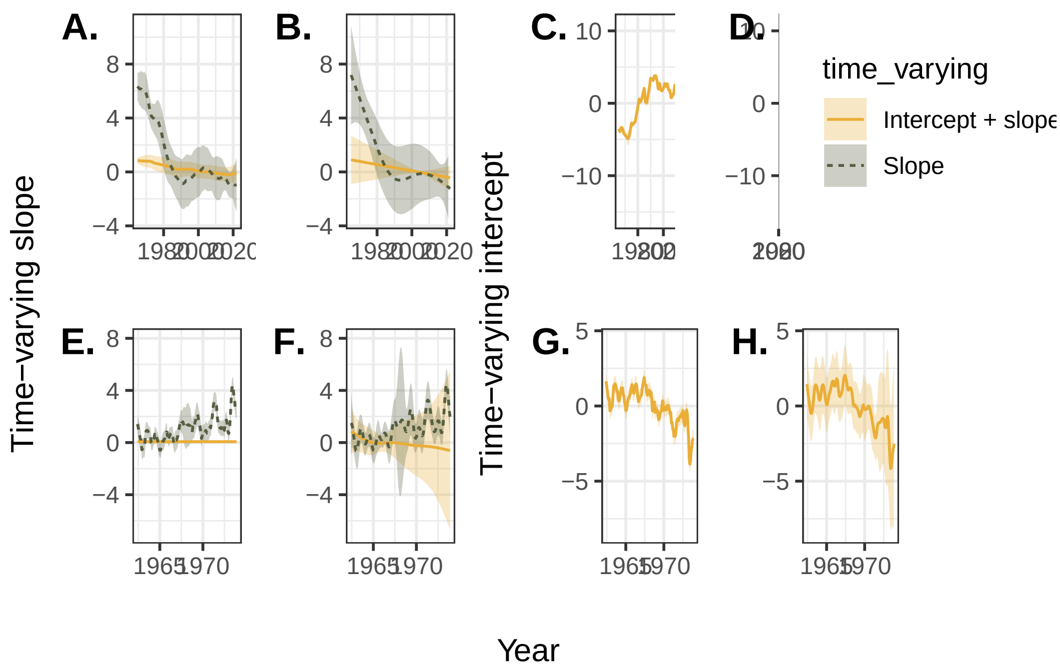
```
ggsave("Figure_4_dlmgam_comparison.png", height = 7, width = 11)

#ggsave("Figure_2_dlmgam_comparison.png", height = 7, width = 7)
```

```
panel_a<-ggplot(coefs%>%filter(dataset=="Alaska salmon : DLM"), aes(date, slope_est, group =
  geom_ribbon(aes(ymin=slope_est-2*slope_se, ymax = slope_est+2*slope_se), alpha=0.3, col = N
  geom_line(aes(lty=time_varying)) +
  ylab("") +
  xlab("") + theme_bw() +
  ylim(c(-3.5,11))+
  scale_fill_manual(values=c(col[4],col[5]))+
  scale_colour_manual(values=c(col[4],col[5]))+
  guides(fill="none",colour="none",lty="none")+
  theme(
    strip.background = element_rect(fill = "white"),
    plot.title = element_text(hjust = 0.5)
  )

panel_b<-ggplot(coefs%>%filter(dataset=="Alaska salmon : GAM"), aes(date, slope_est, group =
  geom_ribbon(aes(ymin=slope_est-2*slope_se, ymax = slope_est+2*slope_se), alpha=0.3, col = N
  geom_line(aes(lty=time_varying)) +
  ylab("") +
  xlab("") + theme_bw() +
  ylim(c(-3.5,11))+
  scale_fill_manual(values=c(col[4],col[5]))+
  scale_colour_manual(values=c(col[4],col[5]))+
  guides(fill="none",colour="none",lty="none")+
  theme(
    strip.background = element_rect(fill = "white"),
    plot.title = element_text(hjust = 0.5)
  )

panel_e <- ggplot(coefs%>%filter(dataset=="Lake Washington : DLM"), aes(date, slope_est, grou
  geom_ribbon(aes(ymin=slope_est-2*slope_se, ymax = slope_est+2*slope_se), alpha=0.3, col = N
  geom_line(aes(lty=time_varying)) +
  ylab("") +
  xlab("") + theme_bw() +
  #ylim(c(-1.5,5.1))+
   ylim(c(-7,8))+
  scale_fill_manual(values=c(col[4],col[5]))+
  scale_colour_manual(values=c(col[4],col[5]))+
  guides(fill="none",colour="none",lty="none")+
```

```
  theme(
    strip.background = element_rect(fill = "white"),
    plot.title = element_text(hjust = 0.5)
  )

panel_f <-ggplot(coefs%>%filter(dataset=="Lake Washington : GAM"), aes(date, slope_est, group
  geom_ribbon(aes(ymin=slope_est-2*slope_se, ymax = slope_est+2*slope_se), alpha=0.3, col = N
  geom_line(aes(lty=time_varying)) +
  ylab("") +
  xlab("") + theme_bw() +
  #ylim(c(-1.5,5.1))+
   ylim(c(-7,8))+
  scale_fill_manual(values=c(col[4],col[5]))+
  scale_colour_manual(values=c(col[4],col[5]))+
  guides(fill="none",colour="none",lty="none")+
  theme(
    strip.background = element_rect(fill = "white"),
    plot.title = element_text(hjust = 0.5)
  )


fig_slope=ggarrange(panel_a, panel_b, panel_e,panel_f, nrow=2,ncol=2,
           labels = c("A.", "B.", "E.", "F."))
fig_slope_y<-annotate_figure(fig_slope, left = textGrob("Time-varying slope",rot=90,gp = gpa



panel_c<-ggplot(coefs%>%filter(dataset=="Alaska salmon : DLM"), aes(date, int_est, group = ti
  geom_ribbon(aes(ymin=int_est-2*int_se, ymax = int_est+2*int_se), alpha=0.3, col = NA) +
  geom_line(aes(lty=time_varying)) +
  ylab("") +
  xlab("") + theme_bw() +
  ylim(-16,11)+
    guides(fill="none",colour="none",lty="none")+
  scale_fill_manual(values=c(col[4],col[5]))+
  scale_colour_manual(values=c(col[4],col[5]))+
  theme(
    strip.background = element_rect(fill = "white"),
    plot.title = element_text(hjust = 0.5)
  )

panel_d<-ggplot(coefs%>%filter(dataset=="Alaska salmon : GAM"), aes(date, int_est, group = ti
```

```
  geom_ribbon(aes(ymin=int_est-2*int_se, ymax = int_est+2*int_se), alpha=0.3, col = NA) +
  geom_line(aes(lty=time_varying)) +
  ylab("") +
  xlab("") + theme_bw() +
  ylim(-16,11)+
  scale_fill_manual(values=c(col[4],col[5]))+
  scale_colour_manual(values=c(col[4],col[5]))+
  theme(
    strip.background = element_rect(fill = "white"),
    plot.title = element_text(hjust = 0.5)
  )

panel_g<-ggplot(coefs%>%filter(dataset=="Lake Washington : DLM"), aes(date, int_est, group =
  geom_ribbon(aes(ymin=int_est-2*int_se, ymax = int_est+2*int_se), alpha=0.3, col = NA) +
  geom_line(aes(lty=time_varying)) +
  ylab("") +
  xlab("") + theme_bw() +
    ylim(-8.5,4.5)+
    guides(fill="none",colour="none",lty="none")+
  scale_fill_manual(values=c(col[4],col[5]))+
  scale_colour_manual(values=c(col[4],col[5]))+
  theme(
    strip.background = element_rect(fill = "white"),
    plot.title = element_text(hjust = 0.5)
  )
panel_h<-ggplot(coefs%>%filter(dataset=="Lake Washington : GAM"), aes(date, int_est, group =
  geom_ribbon(aes(ymin=int_est-2*int_se, ymax = int_est+2*int_se), alpha=0.3, col = NA) +
  geom_line(aes(lty=time_varying)) +
  ylab("") +
  xlab("") + theme_bw() +
  ylim(-8.5,4.5)+
    guides(fill="none",colour="none",lty="none")+
  scale_fill_manual(values=c(col[4],col[5]))+
  scale_colour_manual(values=c(col[4],col[5]))+
  theme(
    strip.background = element_rect(fill = "white"),
    plot.title = element_text(hjust = 0.5)
  )


top<-ggarrange(panel_c, panel_d, nrow=1,ncol=2,
           labels = c("C.", "D."), widths=c(1,1.75))
```

```
bottom<-ggarrange(panel_g, panel_h,nrow=1,ncol=3,
           labels = c("G.", "H.",""), widths=c(1,1,0.71))
arranged_int<-ggarrange(top,bottom, nrow=2)

fig_int_y<-annotate_figure(arranged_int, left = textGrob("Time-varying intercept",rot=90,gp


Fig_4 <-ggarrange(fig_slope_y, fig_int_y,ncol=2,widths=c(1,1.25))
annotate_figure(Fig_4, bottom = textGrob("Year", gp = gpar(cex = 1)))
```




```
ggsave("Figure_4_dlmgam_comparison.png", height = 7, width = 11, bg = "white")
```