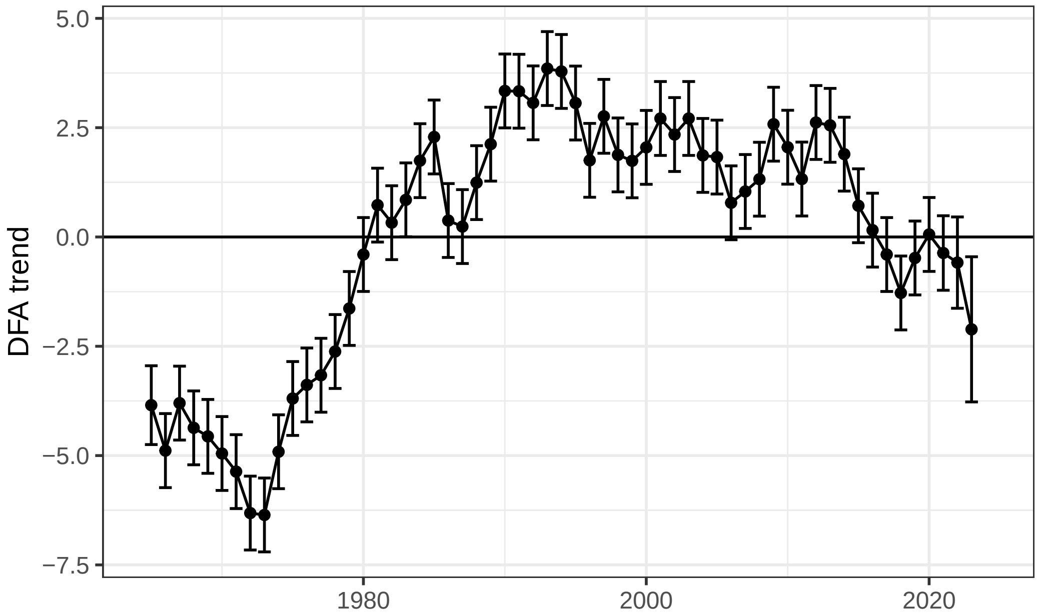
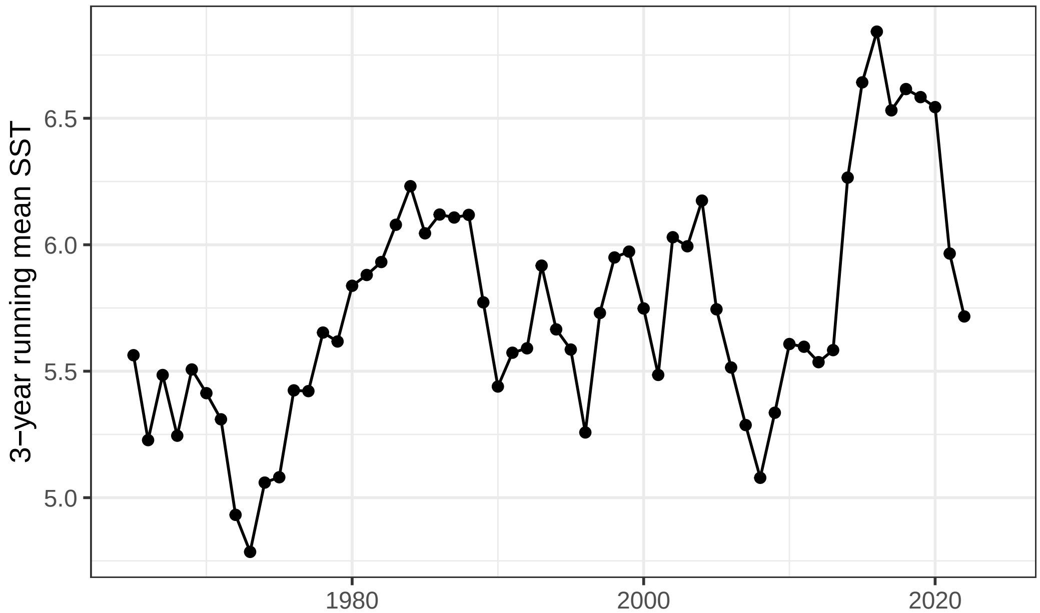
Temporally structured GAMs/LMs for GOA salmon

Data

```r
# Load salmon timeseries collapsed into DFA trend
salm.dat <- readRDS(here::here("GOA_salmon_case_study/Data/dfa_trend.rds"))

# Load 3-year running mean winter SST timeseries
sst.dat <- readRDS(here::here("GOA_salmon_case_study/Data/winterSST_3yr_running_mean.rds"))

# Bind
right_join(salm.dat, sst.dat) -> dat

# Plot salmon DFA
ggplot(dat, aes(year, salmon_DFA)) +
  theme_bw() +
  geom_line() +
  geom_hline(yintercept = 0) +
  geom_point() +
  geom_errorbar(aes(x=year, ymin=conf.low, ymax=conf.high)) +
  xlab("") +
  ylab("DFA trend")
```




```
# Plot sst
ggplot(dat, aes(year, sst_3yr_running_mean)) +
  theme_bw() +
  geom_line() +
  geom_point() +
  xlab("") +
  ylab("3-year running mean SST")
```




```r
# Scale predictor sst
dat %>%
  mutate(z_sst = scale(sst_3yr_running_mean)) %>%
  na.omit() -> dat2
```

Start with evaluate salmon catch using linear models with periods defined a-priori as before and after 1988/1989 (from Litzow et al. 2018)

```r
dat2 %>%
  mutate(period = case_when((year <= 1988) ~ "early",
                            (year > 1988) ~ "late")) -> dat2

mod_period<-lm(salmon_DFA~as.factor(period), data=dat2) # with period
mod_int<-lm(salmon_DFA~1, data=dat2) # intercept only

AIC(mod_period)
```

```
[1] 251.4338
```

```r
AIC(mod_int)
```

```
[1] 289.4622
```

A linear model that includes period is better than an intercept-only model. Now include sst as the slope parameter interacting with period

```r
mod_sst.period.int<-lm(salmon_DFA~z_sst*as.factor(period), data=dat2) # with interaction
mod_sst.period<-lm(salmon_DFA~z_sst+as.factor(period), data=dat2) # no interaction

mod.sum <- data.frame(Covariates = "SST", Period = round(AIC(mod_sst.period), 2),
                      Interaction = round(AIC(mod_sst.period.int), 2))

mod.sum %>%
  knitr::kable(.)
```

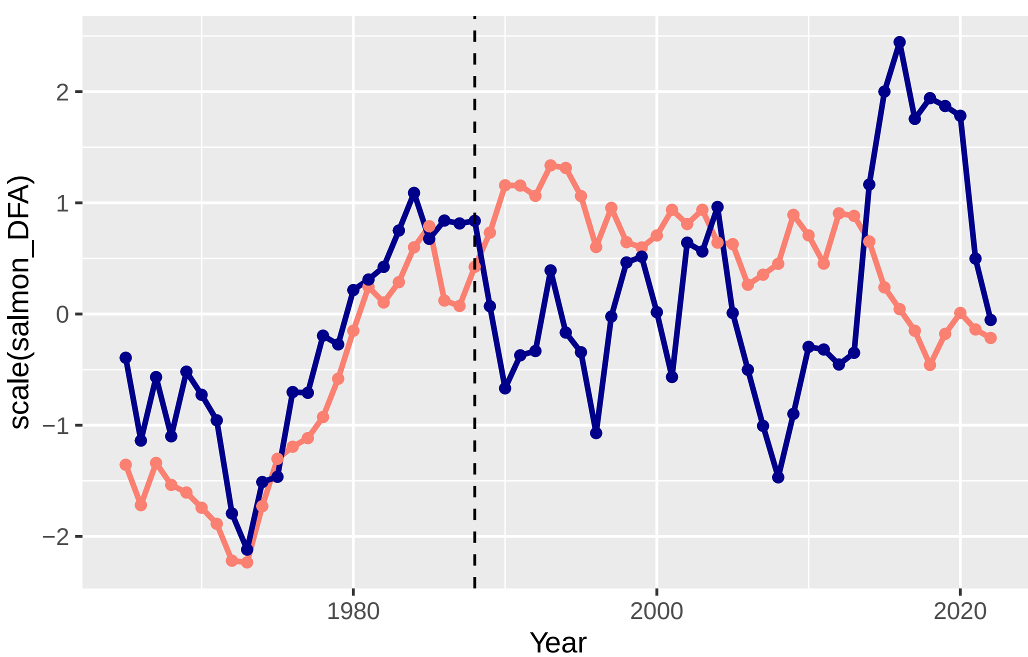
Covariates	Period	Interaction
SST	248.8	179.02

Including sst * period is better than sst + period. View full model results.

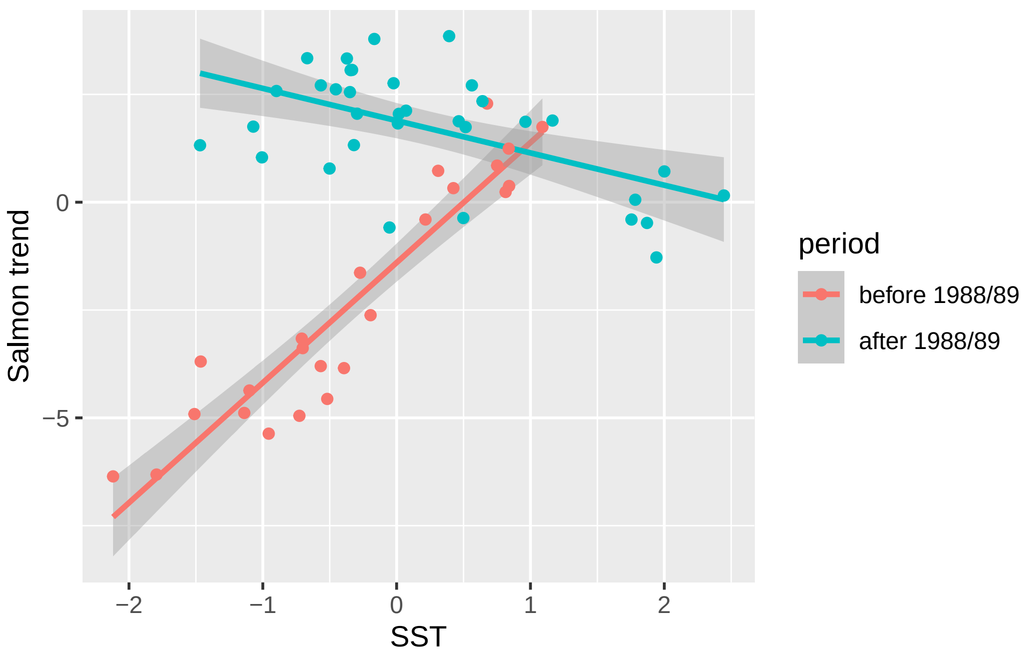
```r
mod.results <-data.table(coef(summary(mod_sst.period.int)), keep.rownames = 'term')
mod.results %>%
  knitr::kable(.)
```

term	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-1.400754	0.2348701	-5.963951	2e-07
z_sst	2.784330	0.2423900	11.486981	0e+00
as.factor(period)late	3.292394	0.3022143	10.894236	0e+00
z_sst:as.factor(period)late	-3.532771	0.3073001	-11.496158	0e+00

```
Fit_slm<-ggplot(data=dat2, aes(group=period,col=period, y=salmon_DFA,x=z_sst)) +
  stat_smooth(method = "lm")+
  scale_colour_discrete(labels=c("before 1988/89", "after 1988/89"))+
  geom_point()+
  ylab("Salmon trend")+
  xlab("SST")

TS_slm<- ggplot()+
  geom_line(dat2, mapping = aes(x=year, y=scale(salmon_DFA)), color = "salmon", size = 1) +
  geom_line(dat2, mapping = aes(x = year, y = z_sst), color = "darkblue", size = 1)+
  geom_point(dat2, mapping = aes(x=year, y=scale(salmon_DFA)), color = "salmon") +
  geom_point(dat2, mapping = aes(x = year, y = z_sst), color = "darkblue")+
  scale_colour_discrete(labels=c("salmon trend", "SST"))+
  geom_vline(xintercept=1988, lty=2)+
  xlab("Year")

TS_slm
```

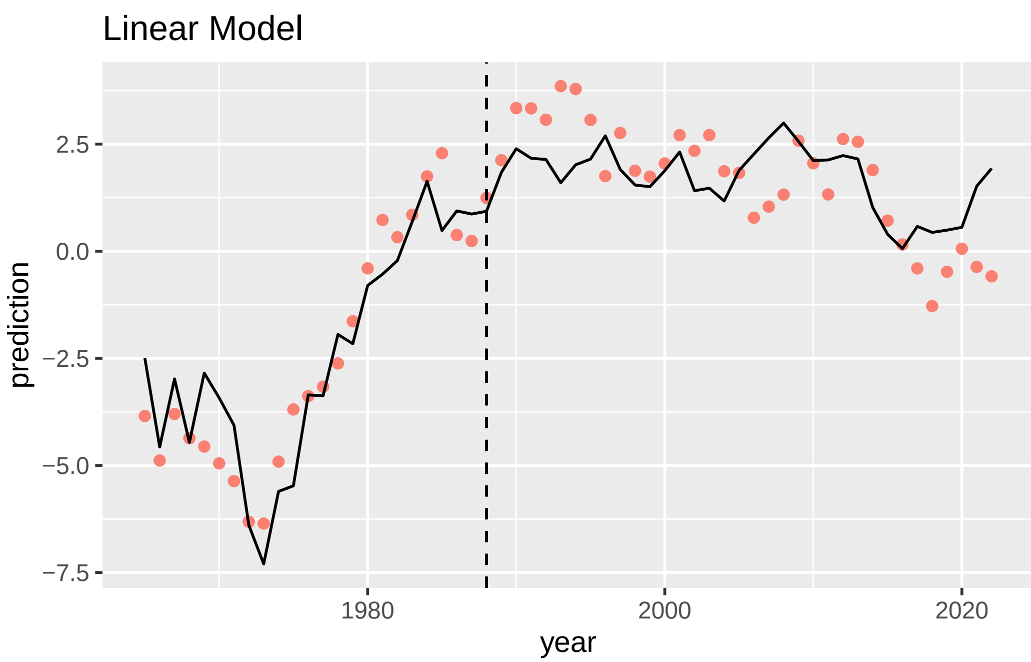



```
Fit_slm
```



Time to fit a temporally-structured GAM

```
prediction=data.frame(prediction=predict(mod_sst.period.int),year=unique(dat2$year))
ggplot(data=prediction,aes(y=prediction, x=year)) +
  #facet_wrap(~area1,scales="free_y") +
 # geom_point()+
  geom_point(data=dat2, aes(y=salmon_DFA, x=year), color = "salmon")+
  geom_line()+
  ggtitle('Linear Model')+
  geom_vline(xintercept=1988, lty=2)
```




```r
y.ts <- ts(data=prediction, frequency=1)
# fit breakpoint model
bp.y <- breakpoints(y.ts ~ 1)
summary(bp.y)
```

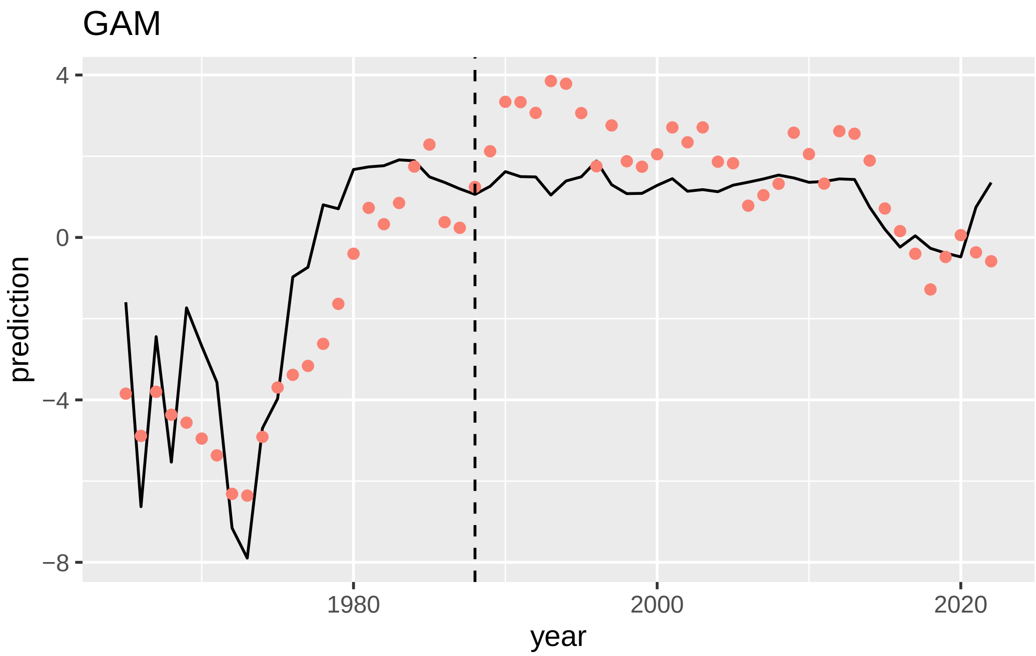
```
	 Optimal (m+1)-segment partition:

Call:
breakpoints.formula(formula = y.ts ~ 1)

Breakpoints at observation number:

m = 1      15
m = 2      13 21
m = 3      15 24          49
m = 4      15 24       41 49
m = 5      15 24 33 41 49
m = 6    8 16 24 33 41 49

Corresponding to breakdates:
```

```
m = 1      15
m = 2      13 21
m = 3      15 24        49
m = 4      15 24     41 49
m = 5      15 24 33 41 49
m = 6    8 16 24 33 41 49

Fit:

m   0      1      2      3      4      5      6
RSS 403.47  70.05  58.00  47.93  46.57  45.70  53.65
BIC 285.22 191.79 188.96 186.01 192.47 199.50 216.92
```

```
prediction[26,]
```

```
   prediction year
26    2.39179 1990
```

```
gam1<-gam(salmon_DFA~s(year, by=z_sst,k=10), data=dat2)
#plot(gam1)
gam2<-gam(salmon_DFA~s(year,k=10), data=dat2)
gam3<-gam(salmon_DFA~s(z_sst,k=10), data=dat2)


prediction2=data.frame(prediction=predict(gam1),year=unique(dat2$year))
ggplot(data=prediction2,aes(y=prediction, x=year)) +
  #facet_wrap(~area1,scales="free_y") +
  #geom_point()+
  geom_line()+
    geom_point(data=dat2, aes(y=salmon_DFA, x=year), color = "salmon")+
  ggtitle('GAM')+
  geom_vline(xintercept=1988, lty=2)
```




```
y.ts <- ts(data=prediction2, frequency=1)
# fit breakpoint model
bp.y <- breakpoints(y.ts ~ 1)
summary(bp.y)
```

```

	 Optimal (m+1)-segment partition:

Call:
breakpoints.formula(formula = y.ts ~ 1)

Breakpoints at observation number:

m = 1   11
m = 2   11                50
m = 3   11 19             49
m = 4   11 19     32      50
m = 5   11 19     32 41 49
m = 6   10 18 26 34 42 50

Corresponding to breakdates:
```

```
m = 1   11
m = 2   11              50
m = 3   11 19           49
m = 4   11 19    32     50
m = 5   11 19    32 41  49
m = 6   10 18 26 34 42  50

Fit:

m   0      1      2      3      4      5      6
RSS 333.32  72.50  64.08  62.55  62.40  62.25  80.41
BIC 274.14 193.78 194.74 201.46 209.44 217.42 240.39
```

```r
prediction2[26,]
```

```
   prediction year
26   1.620673 1990
```