

**Objective:** Evaluate whether harvest influences presence (or biomass) of a species

```
setwd("/Users/rebeccaselden/Documents/Collaborations/FishClim")
surv.dat <- readRDS("Output/dat.latbin2.rds")
land.dat <- readRDS("Output/total.land.latbin.rds")
setorder(land.dat, spp, year)
```

Assign catch=0 for areas surveyed and no catch observed =

Make sure each species has a landing record for each year and lat bin in the survey Get rid of southernmost survey bins because not observed in every year

```
combined.dat <- merge(surv.dat[!(is.na(lat.bin))& year<2015 &! (lat.bin %in% c("30.5", "31.5", "32.5", "33.5", "34.5", "35.5", "36.5", "37.5", "38.5", "39.5", "40.5", "41.5", "42.5", "43.5", "44.5", "45.5", "46.5", "47.5", "48.5", "49.5", "50.5", "51.5", "52.5", "53.5", "54.5", "55.5", "56.5", "57.5", "58.5", "59.5", "60.5", "61.5", "62.5", "63.5", "64.5", "65.5", "66.5", "67.5", "68.5", "69.5", "70.5", "71.5", "72.5", "73.5", "74.5", "75.5", "76.5", "77.5", "78.5", "79.5", "80.5", "81.5", "82.5", "83.5", "84.5", "85.5", "86.5", "87.5", "88.5", "89.5", "90.5", "91.5", "92.5", "93.5", "94.5", "95.5", "96.5", "97.5", "98.5", "99.5")), land.dat, by=c("lat.bin", "year", "spp"), all.x=T)
setorder(combined.dat, spp, year, lat.bin)

combined.dat[, "land_mt" := ifelse(is.na(tl_mt), 0, tl_mt)]
```

Standardize catch so mean of 0 and sd=1 within a species

If no catch in any year, give value of 0

```
catch.scale <- combined.dat[,list(mean_land=mean(land_mt, na.rm=T), sd_land=sd(land_mt, na.rm=T)), by=1]
combined.dat2 <- merge(combined.dat, catch.scale, by="spp")

combined.dat2[,"land_scaled":=ifelse(mean_land==0, 0, scale(land_mt, center=T, scale=T)), by=list(spp)]

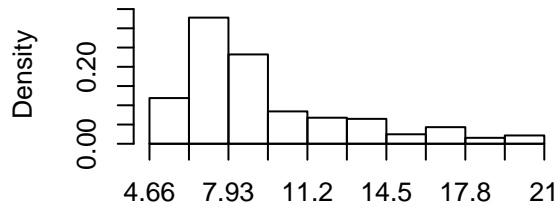
#land_yr <- combined.dat2[,list(mean_land=mean(land_mt), mean_land_scaled=mean(land_scaled)), by=list(spp)]
```

### Standardize Temperature within spp

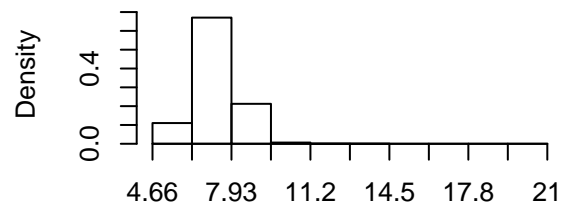
Normalize each temperature for each species range to a mean of 0 and SD of 1 across spp geo range Subtract its mean temperature of occurrence (or of biomass) from the observed temperature Then divide by standard deviation of temperature of occurrence (or of biomass)

```
par(mfrow=c(2,2))
weighted.hist(combined.dat2[spp=="Gadus morhua"]$SBT.actual,
              w=combined.dat2[spp=="Gadus morhua"]$wtcpuenal, freq=F, main="Cod SBT, w=Biomass")
weighted.hist(combined.dat2[spp=="Gadus morhua"]$SBT.actual,
              w=combined.dat2[spp=="Gadus morhua"]$pres2, freq=F, main="Cod SBT, w=Presence")
weighted.hist(combined.dat2[spp=="Centropristis striata"]$SBT.actual,
              w=combined.dat2[spp=="Centropristis striata"]$wtcpuenal, freq=F, main="Dogfish SBT, w=Biomass")
weighted.hist(combined.dat2[spp=="Centropristis striata"]$SBT.actual,
              w=combined.dat2[spp=="Centropristis striata"]$pres2, freq=F, main="Dogfish SBT, w=Presence")
```

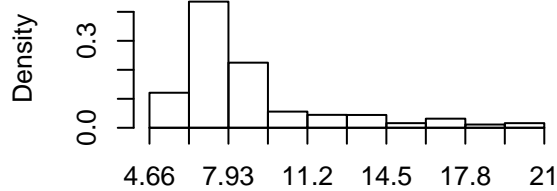
**Cod SBT, w=Biomass**



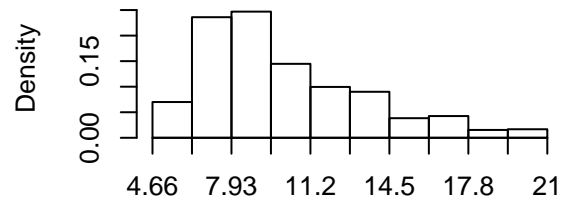
**Cod SBT, w=Presence**



**Dogfish SBT, w=Biomass**



**Dogfish SBT, w=Presence**



```
combined.dat2[,"mean.SBTpres"]:=weighted.mean(SBT.actual, w=pres2), by=list(spp)]
combined.dat2[,"mean.SSTpres"]:=weighted.mean(SST.actual, w=pres2), by=list(spp)]

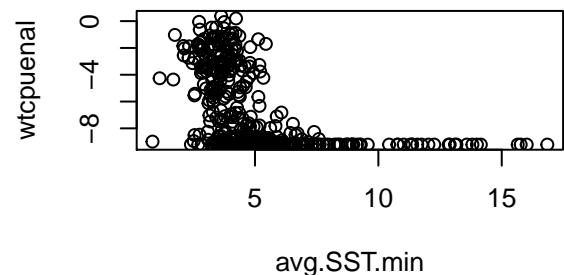
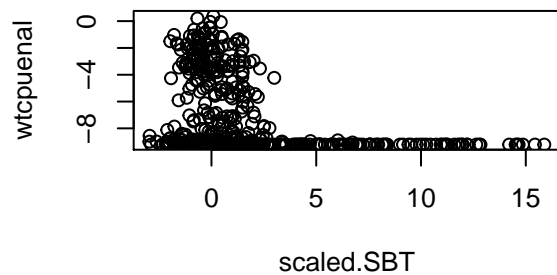
combined.dat2[,"sd.SBTpres"]:=sqrt(wtd.var(SBT.actual, weights=pres2)), by=list(spp)]
combined.dat2[,"sd.SSTpres"]:=sqrt(wtd.var(SST.actual, weights=pres2)), by=list(spp)]

combined.dat2[,"mean.SBTbio"]:=weighted.mean(SBT.actual, w=wtcpuenal), by=list(spp)]
combined.dat2[,"mean.SSTbio"]:=weighted.mean(SST.actual, w=wtcpuenal), by=list(spp)]

combined.dat2[,"sd.SBTbio"]:=sqrt(wtd.var(SBT.actual, weights=wtcpuenal)), by=list(spp)]
combined.dat2[,"sd.SSTbio"]:=sqrt(wtd.var(SST.actual, weights=wtcpuenal)), by=list(spp)]

combined.dat2[,"scaled.SBT"]:=(SBT.actual - mean.SBTpres)/sd.SBTpres]
combined.dat2[,"scaled.SST"]:=(SST.actual - mean.SSTpres)/sd.SSTpres]

plot(wtcpuenal ~ scaled.SBT, combined.dat2[spp=="Gadus morhua"])
plot(wtcpuenal ~ avg.SST.min, combined.dat2[spp=="Gadus morhua"])
```



###

Lag variables in time

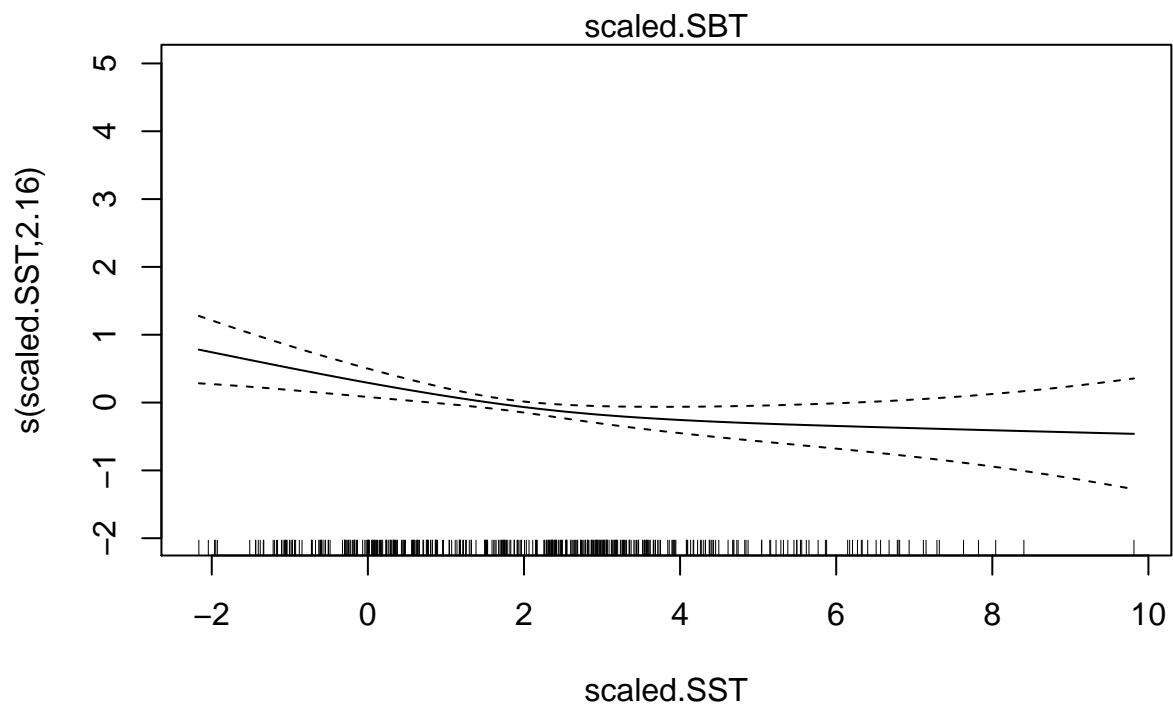
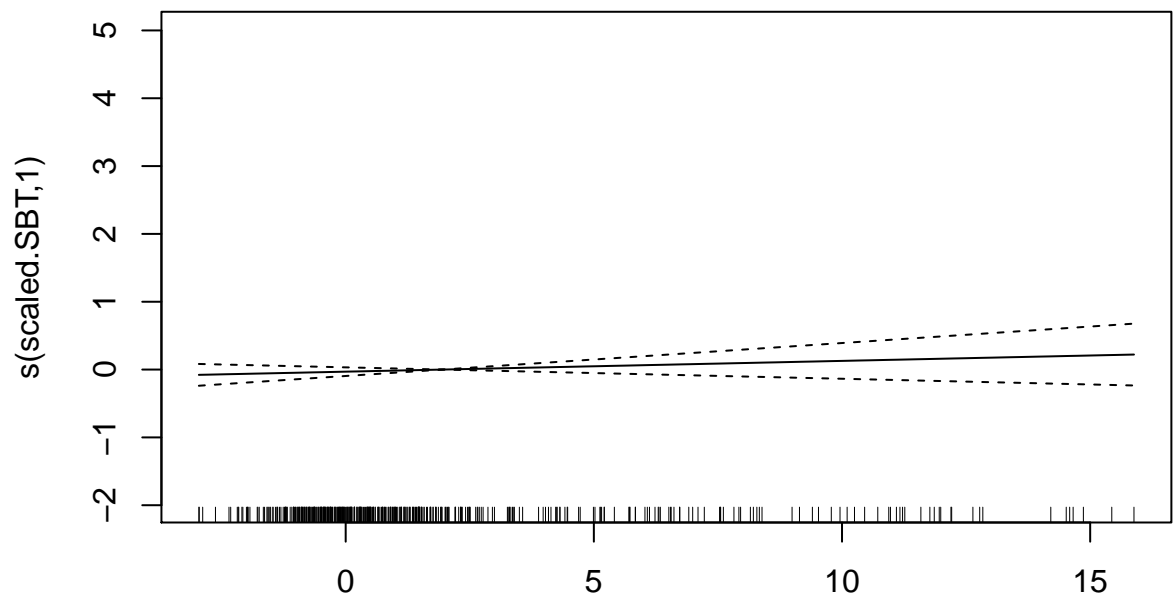
Lag catch, biomass, presence, and environment by one year

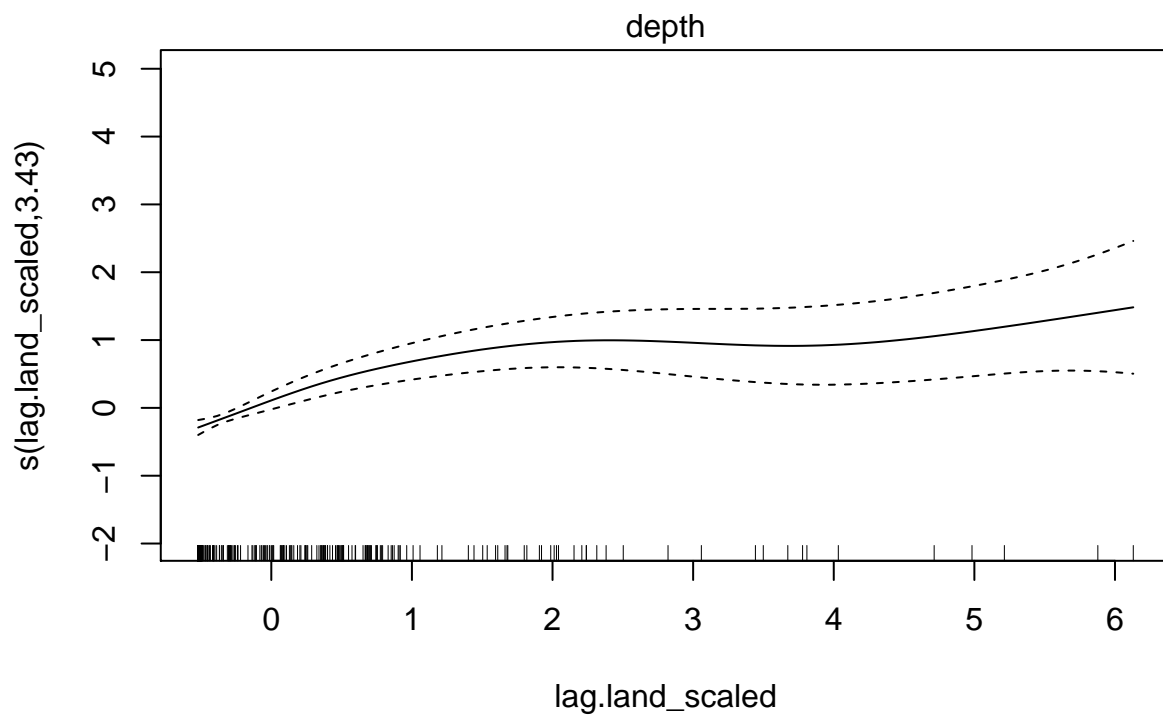
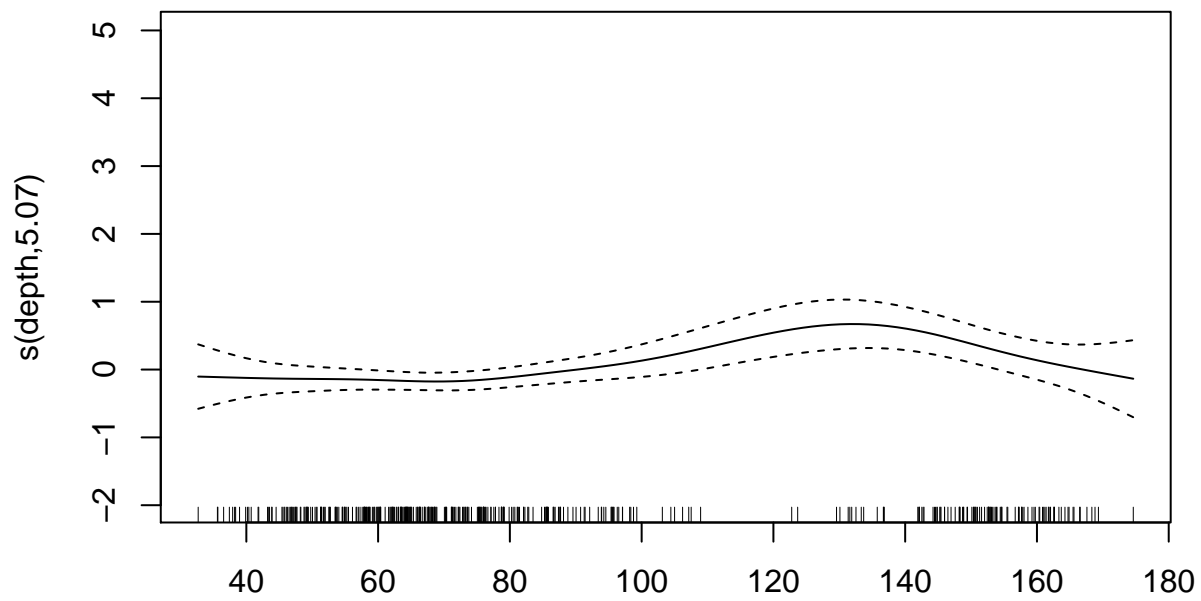
```
vars <- c("pres2", "wtcpuenal", "land_mt", "land_scaled",
          "scaled.SBT", "avg.SBT.min", "avg.SBT.max",
          "scaled.SST", "avg.SST.min", "avg.SST.max")
lagcols <- paste("lag", vars, sep=".")
lag.dt <- combined.dat2[, (lagcols) := shift(.SD), by=list(spp, lat.bin), .SDcols=vars]
```

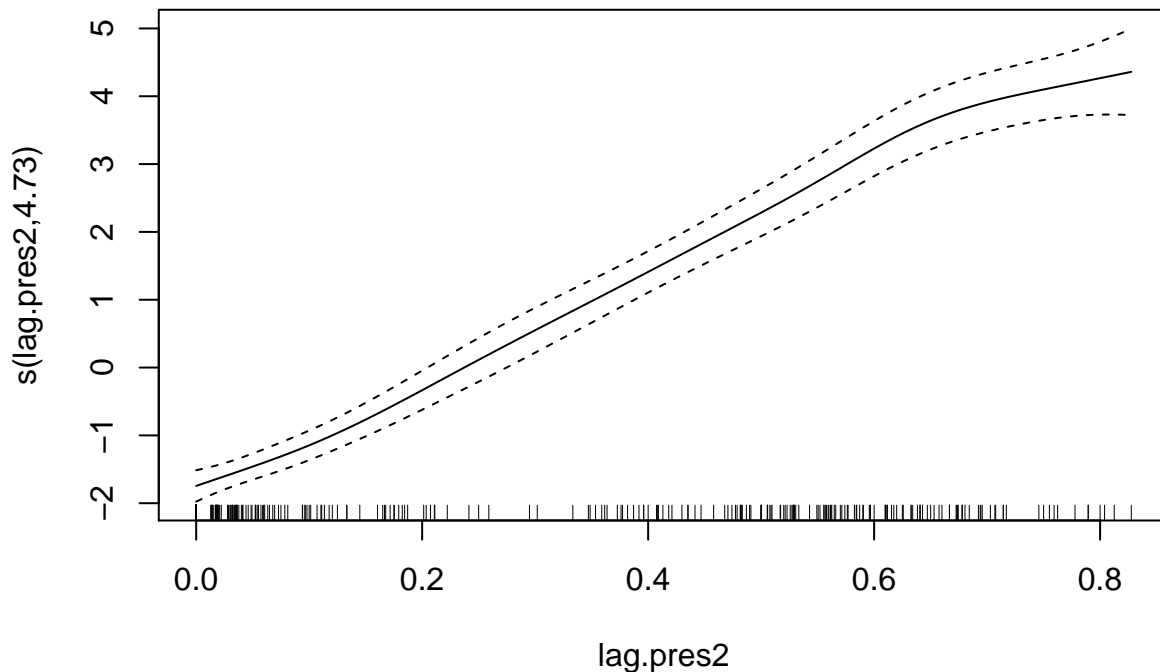
## GAM for Log Bio

Model (log bio) ~ scaled SBT + scaled SST + depth + landings in yr-1 + presence in yr-1 + random(spp)  
 Test with cod

```
cod.mod.bio <- gam(wtcpuenal ~ s(scaled.SBT) + s(scaled.SST) + s(depth) + s(lag.land_scaled) + s(lag.pres)
                  family=gaussian, data=lag.dt[spp%in%c("Gadus morhua")])
plot(cod.mod.bio)
```







```
summary(cod.mod.bio)
```

```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## wtcpuenal ~ s(scaled.SBT) + s(scaled.SST) + s(depth) + s(lag.land_scaled) +
##           s(lag.pres2)
##
## Parametric coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -6.85287    0.03415  -200.6   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##               edf Ref.df      F  p-value
## s(scaled.SBT)    1.000  1.000  0.943   0.3321
## s(scaled.SST)    2.160  2.818  3.677   0.0183 *
## s(depth)         5.065  6.198  2.729   0.0124 *
## s(lag.land_scaled) 3.435  4.245  8.292 1.11e-06 ***
## s(lag.pres2)     4.728  5.776 65.266 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.943   Deviance explained = 94.5%
## GCV = 0.51926   Scale est. = 0.49812    n = 427
# pdf("Figures/FishClimGAMOut.bio.pdf", height=7, width=5)
# mod.fish.clim.bio <- lag.dt[!(spp %in% catch.scale[mean_land<1]$spp),j={
#   print(spp)
#   t.dt <- .SD
```

```

# num.rec.fished <- length(t.dt$land_mt>0)
# if(num.rec.fished >5){
#   mod.b <- gam(wtcpuenal ~ s(scaled.SBT)+ s(scaled.SST)+ s(depth) + s(lag.land_scaled) + s(lag.pres
#     family=gaussian, data=t.dt)
#   chi <- as.vector(summary(mod.b)$chi.sq)
#   fval <- as.vector(summary(mod.b)$s.table[,3])
#   pval <- as.vector(summary(mod.b)$s.pv)
#   vars <- names(summary(mod.b)$chi.sq)
#
#   par(mfrow=c(3,2))
#   plot(mod.b)
#   plot(fitted(mod.b), napredict(mod.b$na.action, mod.b$y), main=spp,
#     ylab="Observed", xlab="Fitted")
#
#   list(vars=vars, chi=chi, fval=fval, pval=pval)
# }
# else{list(vars=NA, chi=NA, fval=NA, pval=NA)}
# }, by=list(spp)]
# dev.off()

```

## Mixed Effects GAM for Log Bio

Species as random effect See random effects in mgcv <https://stat.ethz.ch/R-manual/R-devel/library/mgcv/html/random.effects.html> Model (log bio) ~ scaled SBT + scaled SST + depth + landings in yr-1 + presence in yr-1 + random(spp)

```

### All species single model
spplim.mod.bio <- gam(wtcpuenal ~ s(scaled.SBT)+ s(scaled.SST)+ s(depth) + s(lag.land_scaled) + s(lag.p
  family=gaussian, data=lag.dt[spp%in%c("Gadus morhua", "Squalus acanthias", "Limanda ferru
summary(spplim.mod.bio)

```

```

##
## Family: gaussian
## Link function: identity
##
## Formula:
## wtcpuenal ~ s(scaled.SBT) + s(scaled.SST) + s(depth) + s(lag.land_scaled) +
##   s(lag.pres2) + s(spp, bs = "re")
##
## Parametric coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -4.6517      0.0218  -213.3   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##               edf Ref.df      F  p-value
## s(scaled.SBT)    1.000   1.000  11.742  0.00063 ***
## s(scaled.SST)    5.018   6.229   9.623  1.30e-10 ***
## s(depth)         2.887   3.609   1.012  0.37762
## s(lag.land_scaled) 2.814   3.497  10.150  3.94e-07 ***
## s(lag.pres2)     1.000   1.000 1002.383 < 2e-16 ***
## s(spp)           2.250   3.000  88.173 < 2e-16 ***

```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Rank: 48/49
## R-sq.(adj) =  0.852   Deviance explained = 85.3%
## GCV = 1.0961   Scale est. = 1.0826     n = 1281
```

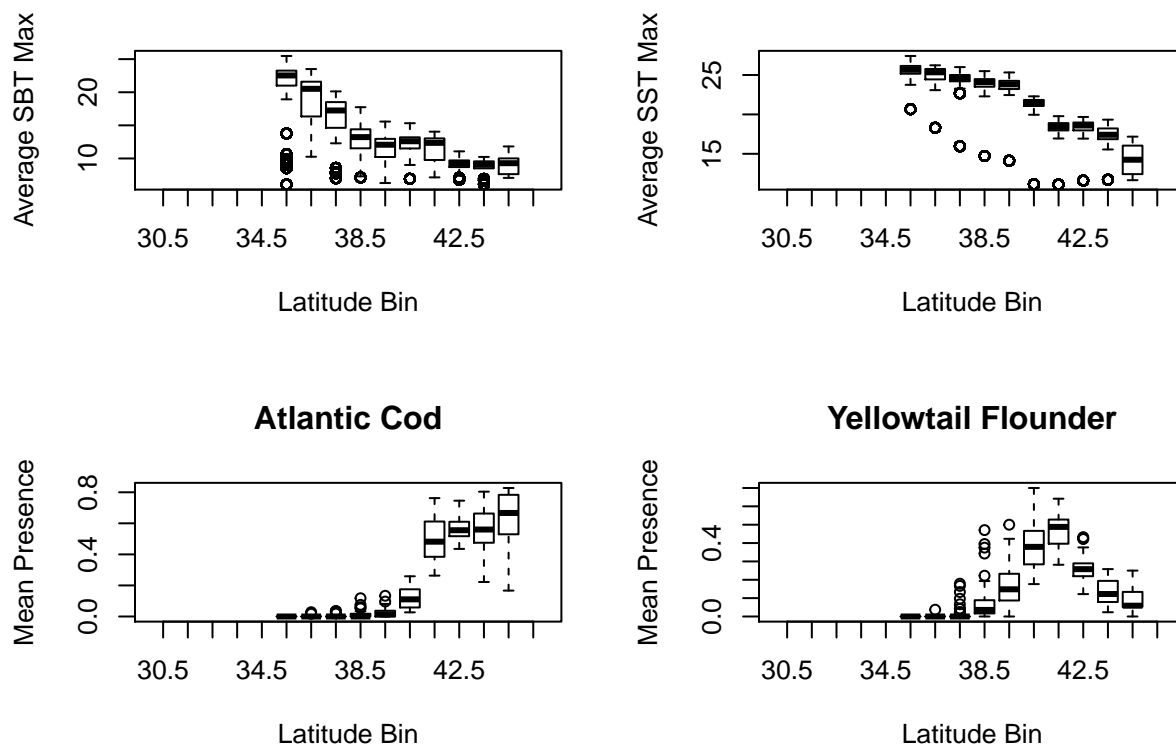
```
gam.vcomp(spllim.mod.bio)
```

```
##      s(scaled.SBT)      s(scaled.SST)      s(depth)
##      9.634996e-07      2.292728e-01      1.603248e-03
## s(lag.land_scaled)  s(lag.pres2)      s(spp)
##      2.170153e-01      1.082680e-05      3.104446e+03
```

### Beta Regression for Presence

```
### To use beta regression, need pres2 to be (0,1)
lag.dt[, "pres2_beta" := ifelse(pres2 == 0, 0.00001, ifelse(pres2 == 1, 0.99999, pres2))]

par(mfrow=c(2,2))
plot(avg.SBT.max ~ lat.bin, lag.dt, ylab="Average SBT Max", xlab="Latitude Bin")
plot(avg.SST.max ~ lat.bin, lag.dt, ylab="Average SST Max", xlab="Latitude Bin")
boxplot(pres2 ~ lat.bin, lag.dt[spp=="Gadus morhua"], main="Atlantic Cod",
        ylab="Mean Presence", xlab="Latitude Bin")
boxplot(pres2 ~ lat.bin, lag.dt[spp=="Limanda ferruginea"], main="Yellowtail Flounder",
        ylab="Mean Presence", xlab="Latitude Bin")
```



```
# par(mfrow=c(3,2))
# plot(pres2 ~ avg.SST.max, lag.dtN[spp=="Gadus morhua"], main="Atlantic Cod",
#      ylab="Mean Presence", xlab="Average SST Max")
```



```

# plot(pres2 ~ avg.SST.max, lag.dtN[spp=="Limanda ferruginea"], main="Yellowtail Flounder",
#       ylab="Mean Presence", xlab="Average SST Max")
# plot(pres2 ~ lag.land_scaled, lag.dtN[spp=="Gadus morhua"], main="Atlantic Cod",
#       ylab="Mean Presence", xlab="Scaled Landings t-1")
# plot(pres2 ~ lag.land_scaled, lag.dtN[spp=="Limanda ferruginea"], main="Yellowtail Flounder",
#       ylab="Mean Presence", xlab="Scaled Landings t-1")
# plot(lag.land_scaled ~ lag.pres2, lag.dtN[spp=="Gadus morhua"], main="Atlantic Cod",
#       ylab="Scaled Landings t-1", xlab="Mean Presence t-1")
# plot(lag.land_scaled ~ lag.pres2, lag.dtN[spp=="Limanda ferruginea"], main="Yellowtail Flounder",
#       ylab="Scaled Landings t-1", xlab="Mean Presence t-1")

### Model (pres) ~ max SBT + max SST + depth + landings in yr-1 + presence in yr-1
#### Test with cod
cod.mod <- gam(pres2_beta ~ s(avg.SBT.max) + s(avg.SST.max) + s(depth) + s(lag.land_scaled) + s(lag.pres2),
              family=betar(link="logit"), data=lag.dt[spp=="Gadus morhua"])

summary(cod.mod)

##
## Family: Beta regression(25.977)
## Link function: logit
##
## Formula:
## pres2_beta ~ s(avg.SBT.max) + s(avg.SST.max) + s(depth) + s(lag.land_scaled) +
##           s(lag.pres2)
##
## Parametric coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.79948    0.03755  -74.56  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##               edf Ref.df  Chi.sq p-value
## s(avg.SBT.max)   2.014  2.535   1.063  0.5719
## s(avg.SST.max)   4.286  5.319  42.458 8.3e-08 ***
## s(depth)         2.781  3.494   8.217  0.0646 .
## s(lag.land_scaled) 2.369  2.948   7.751  0.0518 .
## s(lag.pres2)     6.435  7.567 224.985 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) = 0.955  Deviance explained = 98.3%
## -REML = -1723.8  Scale est. = 1          n = 427

# #### Remove unfished and very lightly fished species
# #### use scaled SST & SBT
# pdf("Figures/FishClimGAMOut.pdf", height=7, width=5)
# mod.fish.clim <- lag.dt[!(spp %in% catch.scale[mean_land<1]$spp),j={
#   print(spp)
#   t.dt <- .SD
#   num.rec.fished <- length(t.dt$land_mt>0)

```

```

#   if(num.rec.fished >5){
#     mod.p <- gam(pres2_beta ~ s(scaled.SBT)+ s(scaled.SST)+ s(depth) + s(lag.land_scaled) + s(lag.pres2),
#                 family=betar(link="logit"), data=t.dt)
#     chi <- as.vector(summary(mod.p)$chi.sq)
#     fval <- as.vector(summary(mod.p)$s.table[,3])
#     pval <- as.vector(summary(mod.p)$s.pv)
#     vars <- names(summary(mod.p)$chi.sq)
#
#     par(mfrow=c(3,2))
#     plot(mod.p)
#     plot(fitted(mod.p), napredict(mod.p$na.action, mod.p$y), main=spp,
#          ylab="Observed", xlab="Fitted")
#
#     list(vars=vars, chi=chi, fval=fval, pval=pval)
#   }
#   else{list(vars=NA, chi=NA, fval=NA, pval=NA)}
# }, by=list(spp)]
# dev.off()
#
# mod.fish.clim.nopres <- lag.dt[!(spp %in% catch.scale[mean_land<1]$spp),j]={
#   print(spp)
#   t.dt <- .SD
#   num.rec.fished <- length(t.dt$land_mt>0)
#   if(num.rec.fished >5){
#     mod.p <- gam(pres2_beta ~ s(avg.SBT.max)+ s(avg.SST.max)+ s(depth) + s(lag.land_scaled),
#                 family=betar(link="logit"), data=t.dt)
#     chi <- as.vector(summary(mod.p)$chi.sq)
#     pval <- as.vector(summary(mod.p)$s.pv)
#     vars <- names(summary(mod.p)$chi.sq)
#
#     par(mfrow=c(3,2))
#     plot(mod.p)
#     plot(fitted(mod.p), napredict(mod.p$na.action, mod.p$y), main=spp,
#          ylab="Observed", xlab="Fitted")
#     qq.gam(mod.p)
#
#     list(vars=vars, chi=chi, pval=pval)
#   }
#   else{list(vars=NA, chi=NA, pval=NA)}
# }, by=list(spp)]

```

## Regression for Delta Presence

```

lag.dt[, "pres2_delta" := pres2 - lag.pres2]
lag.dt[, "SBT.delta" := avg.SBT.max - lag.avg.SBT.max]
lag.dt[, "SST.delta" := avg.SST.max - lag.avg.SST.max]
lag.dt[, "lag.land.delta" := land_scaled - lag.land_scaled]

### Low deviance explained for delta temperatures
delta.cod.mod <- gam(pres2_delta ~ s(SBT.delta) + s(SST.delta) + s(depth) + s(lag.land_scaled),
                    data=lag.dt[spp=="Gadus morhua"])

```

```

### Really low deviance explained for absolute temperatures
delta.cod.mod2 <- gam(pres2_delta ~ s(avg.SBT.max) + s(avg.SST.max) + s(depth) + s(lag.land_scaled),
                      data=lag.dt[spp=="Gadus morhua"])
summary(delta.cod.mod2)

##
## Family: gaussian
## Link function: identity
##
## Formula:
## pres2_delta ~ s(avg.SBT.max) + s(avg.SST.max) + s(depth) + s(lag.land_scaled)
##
## Parametric coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.002929   0.003220  -0.91    0.364
##
## Approximate significance of smooth terms:
##               edf Ref.df      F p-value
## s(avg.SBT.max)    1.000   1.000 0.066   0.798
## s(avg.SST.max)    1.725   2.161 1.214   0.263
## s(depth)          1.000   1.000 1.217   0.271
## s(lag.land_scaled) 1.000   1.000 1.136   0.287
##
## R-sq.(adj) =  0.0017   Deviance explained = 1.28%
## GCV = 0.0044883   Scale est. = 0.0044281   n = 427

```

## First Differences Econometrics Model

```

lag.dt[, "bio_delta" := wtcpuenal - lag.wtcpuenal]
fd.bio.cod <- gam(bio_delta ~ s(SBT.delta) + s(SST.delta) + s(lag.land.delta),
                  data=lag.dt[spp=="Gadus morhua"])
summary(fd.bio.cod)

##
## Family: gaussian
## Link function: identity
##
## Formula:
## bio_delta ~ s(SBT.delta) + s(SST.delta) + s(lag.land.delta)
##
## Parametric coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.04217    0.03448  -1.223   0.222
##
## Approximate significance of smooth terms:
##               edf Ref.df      F p-value
## s(SBT.delta)     1      1 0.007   0.936
## s(SST.delta)     1      1 0.002   0.962
## s(lag.land.delta) 1      1 0.109   0.741
##
## R-sq.(adj) = -0.00681   Deviance explained = 0.0278%
## GCV = 0.51243   Scale est. = 0.50763   n = 427

```

## Fixed Effects Econometrics Model

Substract off mean for each latitude bin and species

```
fe.dt <- lag.dt[,list(m.pres2=mean(pres2, na.rm=T),
                     m.wtcpuenal=mean(wtcpuenal, na.rm=T),
                     m.land =mean(land_scaled, na.rm=T),
                     m.SBT=mean(SBT.actual, na.rm=T),
                     m.SST=mean(SST.actual, na.rm=T)),
                 by=list(spp, lat.bin)]

fe.dt2 <- merge(lag.dt[!(lat.bin %in% c("30.5", "31.5", "32.5", "33.5", "34.5"))],
               fe.dt, by=c("spp", "lat.bin"))

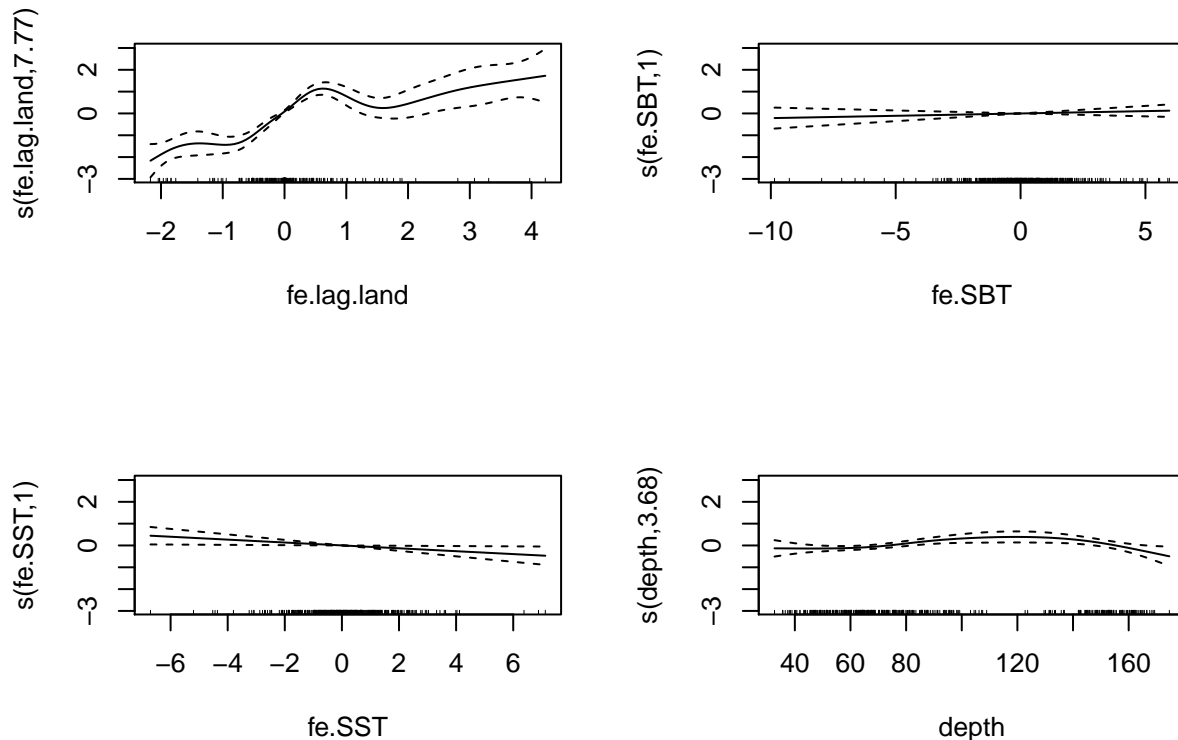
fe.dt2[, "fe.pres2" := pres2 - m.pres2]
fe.dt2[, "fe.wtcpuenal" := wtcpuenal - m.wtcpuenal]
fe.dt2[, "fe.lag.land" := lag.land_scaled - m.land]
fe.dt2[, "fe.SBT" := SBT.actual - m.SBT]
fe.dt2[, "fe.SST" := SST.actual - m.SST]

fe.mod.cod <- gam(fe.wtcpuenal ~ s(fe.lag.land) + s(fe.SBT) + s(fe.SST) + s(depth), data=fe.dt2[spp=="Gadus morhua"])
fe.mod.cod.lm <- lm(fe.wtcpuenal ~ fe.lag.land + fe.SBT + fe.SST + depth, data=fe.dt2[spp=="Gadus morhua"])

summary(fe.mod.cod)

##
## Family: gaussian
## Link function: identity
##
## Formula:
## fe.wtcpuenal ~ s(fe.lag.land) + s(fe.SBT) + s(fe.SST) + s(depth)
##
## Parametric coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.01424    0.03751  -0.38    0.704
##
## Approximate significance of smooth terms:
##              edf Ref.df      F p-value
## s(fe.lag.land) 7.77  8.583 29.346 < 2e-16 ***
## s(fe.SBT)      1.00  1.000  0.782 0.37714
## s(fe.SST)      1.00  1.000  4.980 0.02616 *
## s(depth)       3.68  4.568  4.445 0.00097 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.403   Deviance explained = 42.2%
## GCV = 0.62189   Scale est. = 0.60084    n = 427

par(mfrow=c(2,2))
plot(fe.mod.cod)
```



*# higher landings than average in the previous year associated with higher than average biomass this year*

### Panel Data Approaches Based on Ethan Addicott's code

```
library(plm)
df <- combined.dat2[spp=="Gadus morhua"]
setorder(df, lat.bin, year)

### Create the panel
p.df <- pdata.frame(df, index = c("lat.bin", "year"))

### Lag landings by one year (in same lat.bin)
p.df$land_mt.lag.1 <- lag(p.df$land_mt, k = 1)
```

Just do the Fixed Effects model (as I did manually in the previous section of code)

Documentation from plm vignette <https://cran.r-project.org/web/packages/plm/vignettes/plm.pdf> the standard way of estimating fixed effects models with, say, group (time) effects entails transforming the data by subtracting the average over time (group) to every variable, which is usually termed time-demeaning. Analogous to Ethan's m.4: `m.4 <- plm(wtcpue~pred.total+ER.lag.1, data = p.df, model = "within")`

```
m.fe <- plm(wtcpuenal ~ SBT.actual + SST.actual + depth + land_mt.lag.1, data=p.df, model="within")
summary(m.fe)
```

## Oneway (individual) effect Within Model

##

## Call:

## `plm(formula = wtcpuenal ~ SBT.actual + SST.actual + depth + land_mt.lag.1,`

```

##      data = p.df, model = "within")
##
## Unbalanced Panel: n = 10, T = 11-46, N = 419
##
## Residuals:
##      Min. 1st Qu.  Median 3rd Qu.    Max.
## -3.9000 -0.2300 -0.0331  0.1750  4.2000
##
## Coefficients:
##              Estimate Std. Error t-value Pr(>|t|)
## SBT.actual      1.1784e-02  2.6898e-02  0.4381  0.66155
## SST.actual     -7.0972e-02  3.2396e-02 -2.1908  0.02904 *
## depth           3.4058e-03  3.2269e-03  1.0554  0.29186
## land_mt.lag.1   1.6340e-04  1.3793e-05 11.8465 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:      387.14
## Residual Sum of Squares: 282.93
## R-Squared:      0.26917
## Adj. R-Squared: 0.24571
## F-statistic: 37.291 on 4 and 405 DF, p-value: < 2.22e-16

# Include instrumental variables in the fixed effects model instruments after the | if the model is y ~x1 + x2
+ x3 with x1 and x2 endogenous and z1 and z2 external instruments formula=y~x1+x2+x3 | x3 + z1 + z2
Or equivalently formula=y~x1+x2+x3 | . -x1-x2+z1+z2 (eg subtract out x1 and x1 and add in z1 and z2)

```