Streamlined Walleye Pollock Analyses & Figures

Laura Vary

1/24/2022

Contents

Purpose of This Document:

The purpose of this document is to streamline files and associated analyses for the creation of generalized additive models that investigate spawning behavior and larval biogeography among fishes in the Bering Sea. This is mainly an automation document, with the goal of minimizing the back-and-forth between code files should data need to be modified or analyses re-ran.

Loading Data:

Walleye pollock: both egg and larval data are included for this species. Pollock spawn from February to May, live roughly 12 years, and transform to juveniles at standard lengths between 25 and 40 mm.

These data have been trimmed. The egg data are constrained to depths between 54 and 221 meters; temporally, the egg data are constrained to above the 99th day of year and below the 160th day of the year (temporally centered on the spawning period of pollock). The egg data are constrained to latitudes between 54 and 62 degrees north and longitudes less than -175 degrees west. The egg data are also joined to regional temperature indices for each year (the reg.sst dataset). The larval data are constrained to latitudes between 53.5 and 61.5 degrees north and longitudes less than -175 degrees west. Larvae are linked to CTD-derived, *in situ* temperature and salinity measurements and constrained to days of year between doy 90 and 159.

The regional temperature index data are constrained to (-180, -151) degrees W and (50.5, 67.5) degrees N and reflect the average March temperature for each year across that region. March temperatures are chosen to estimate the conditions spawning pollock may have experienced, roughly two months before the peak amount of eggs in the water column occurs.

Descriptive Information:

Table 1: Descriptive Metrics for Pollock Egg Data

Lat Range	Lon Range	Day of Year Range	Bottom Depth Range
54-60.1	-173 to -158.2	100-159	0-3506

Table 2: Descriptive Metrics for Pollock Larval Data

Lat Range	Lon Range	Day of Year Range	Bottom Depth Range
53.6-60.3	-174 to -158.2	99-159	23-3365

The following two plots show *the day of year distribution for positive pollock egg catch* (left) and *the year distribution for positive pollock egg catch* (right). Analogous plots for larval data are following.

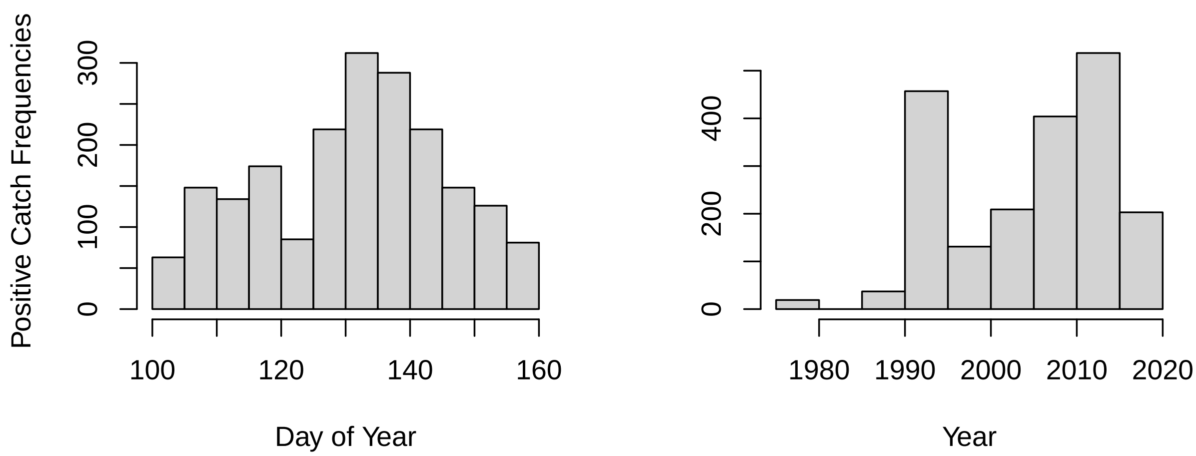
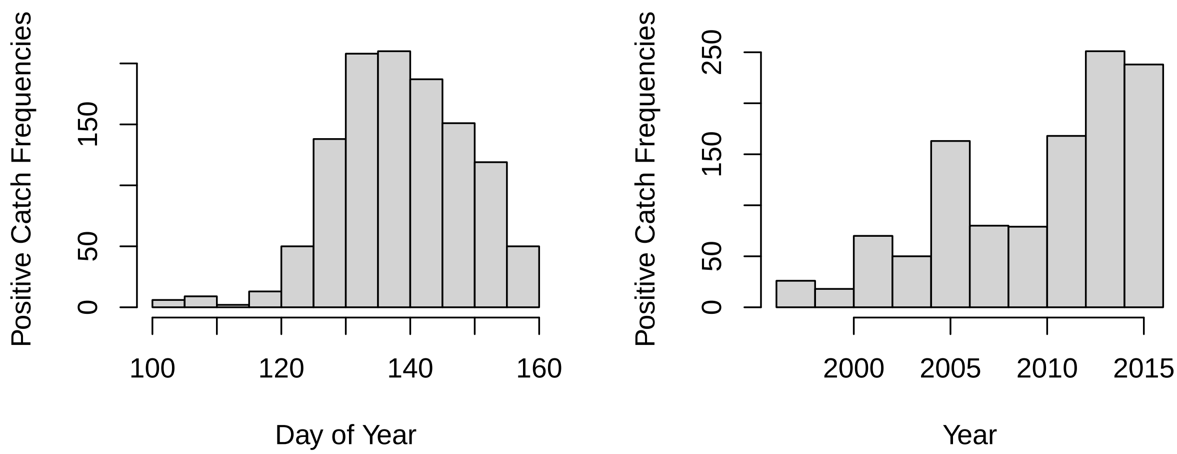
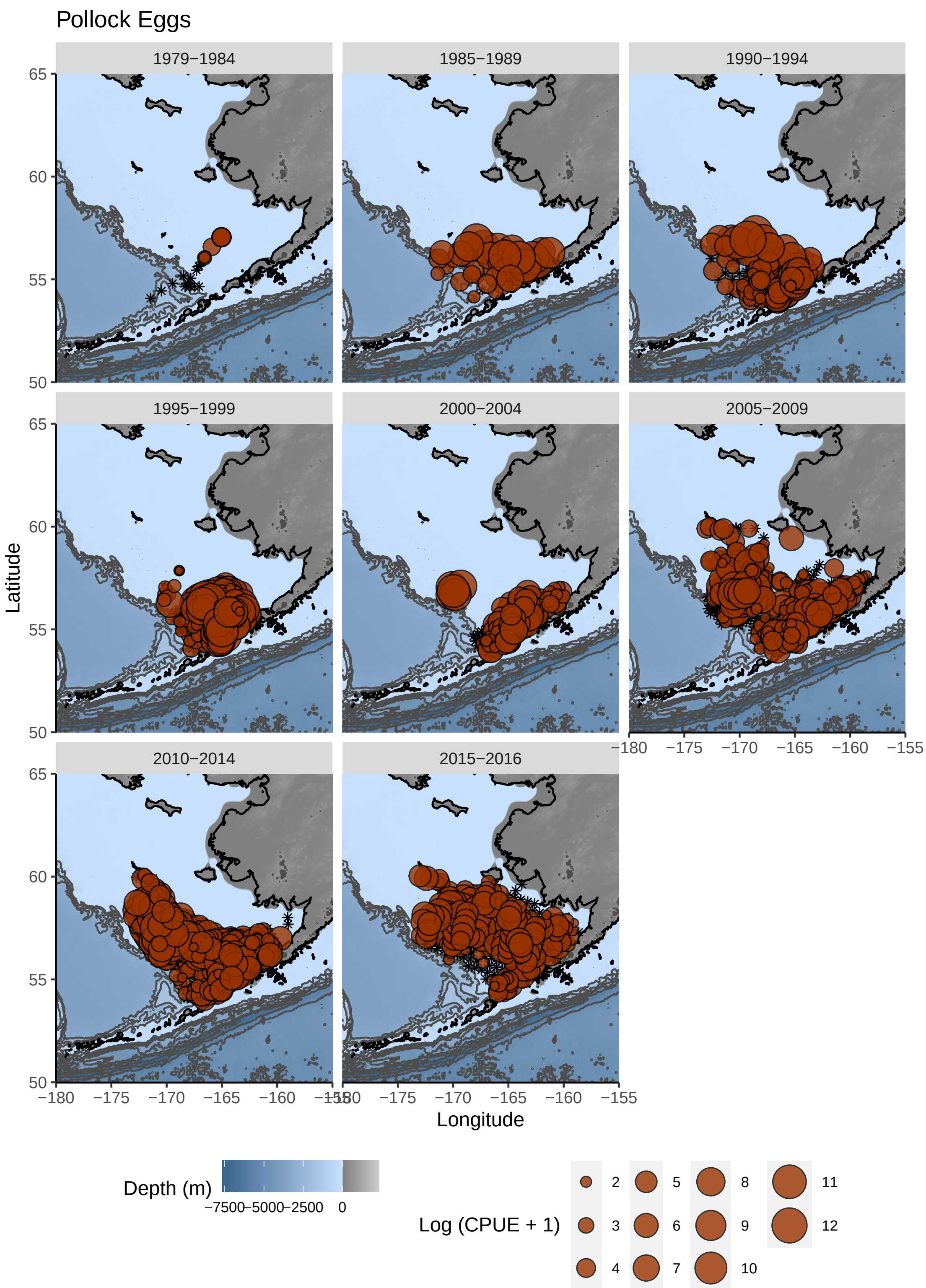


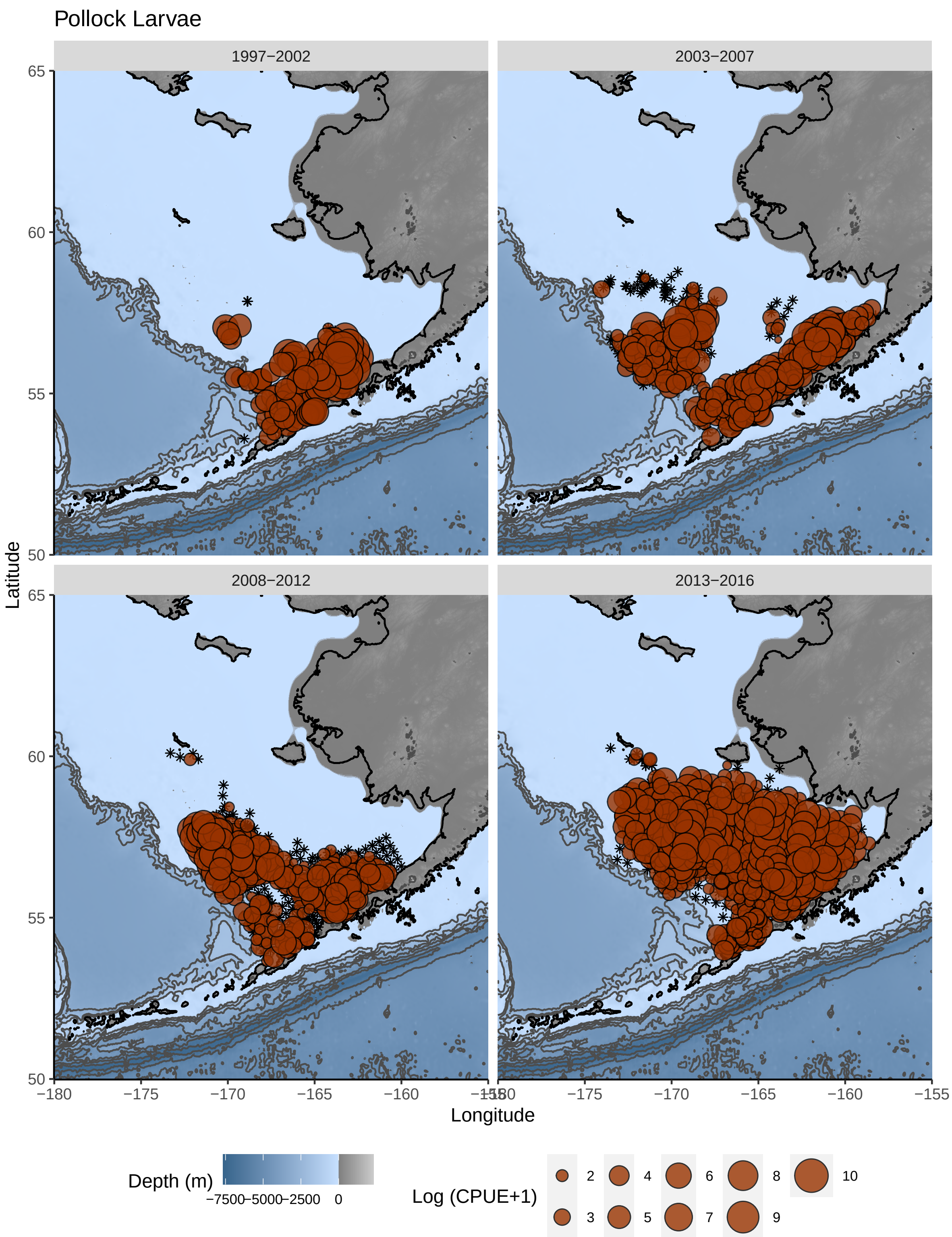
Figure 1: Pollock Eggs



The following plots show pollock egg and larval catch distributions (Catch per unit effort, or per 10m^2) across five year increments from 1979 to 2016.



Pollock Eggs
1979−1984
1985−1989
1990−1994
1995−1999
2000−2004
2005−2009
2010−2014
2015−2016
Latitude
Longitude
65
60
55
50
−180
−175
−170
−165
−160
−155
Depth (m)
−7500
−5000
−2500
0
Log (CPUE + 1)
2
3
4
5
6
7
8
9
10
11
12



Now we'll move into the GAMs. The following code *is only necessary if the data were re-trimmed and new GAMs need to be run.* In this case, modify markdown document such that "{eval = TRUE}". The other model figures are marked as "eval = FALSE" if they, as of the last model run, do not produce the best model results. **Make sure to save the new models as RDS objects.**

Pollock eggs were best explained by the threshold geography model, in which the spatial distribution of eggs varied differently below and above 1.08 degrees Celsius.

Generalized Additive Models: Pollock Eggs

The base model formulation:

```
eg.base<-readRDS("./GAM Models/pk_egg_base.rds")
summary(eg.base)
```

```
##
## Family: Tweedie(p=1.99)
## Link function: log
##
## Formula:
## (Cper10m2 + 1) ~ factor(year) + s(lon, lat) + s(doy) + s(bottom_depth,
##     k = 5)
##
## Parametric coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       3.893065   0.327355  11.892  < 2e-16 ***
## factor(year)1988  5.262118   0.448779  11.725  < 2e-16 ***
## factor(year)1991  3.734567   0.436803   8.550  < 2e-16 ***
## factor(year)1992  3.747396   0.495497   7.563 5.52e-14 ***
## factor(year)1993  4.172129   0.395443  10.551  < 2e-16 ***
## factor(year)1994  2.755748   0.391602   7.037 2.53e-12 ***
## factor(year)1995  2.581079   0.352858   7.315 3.47e-13 ***
## factor(year)1996 -2.468430   0.829333  -2.976 0.002945 **
## factor(year)1997  0.934351   0.405810   2.302 0.021393 *
## factor(year)1998  0.003856   0.665242   0.006 0.995376
## factor(year)1999  3.050271   0.423969   7.195 8.27e-13 ***
## factor(year)2000  2.490703   0.580076   4.294 1.82e-05 ***
## factor(year)2002  1.320512   0.385857   3.422 0.000631 ***
## factor(year)2003  0.656702   0.401769   1.635 0.102276
## factor(year)2005  0.106564   0.374938   0.284 0.776266
## factor(year)2006  0.254813   0.374133   0.681 0.495886
## factor(year)2007  2.152415   0.388722   5.537 3.40e-08 ***
## factor(year)2008  2.460322   0.378408   6.502 9.57e-11 ***
## factor(year)2009 -0.583625   0.380339  -1.534 0.125038
## factor(year)2010  0.418674   0.377640   1.109 0.267684
## factor(year)2011  0.366685   0.429801   0.853 0.393659
## factor(year)2012  2.666663   0.342525   7.785 1.01e-14 ***
## factor(year)2013 -0.099193   0.464814  -0.213 0.831030
## factor(year)2014  0.820757   0.326444   2.514 0.011992 *
## factor(year)2015  0.126236   0.455995   0.277 0.781929
## factor(year)2016  1.156368   0.329401   3.511 0.000455 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                     edf Ref.df     F p-value
## s(lon,lat)       27.297 28.820 40.59  <2e-16 ***
## s(doy)            6.558  7.694 27.54  <2e-16 ***
```

```
## s(bottom_depth)  1.003  1.007 47.30  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.126   Deviance explained = 50.4%
## -REML =  15361  Scale est. = 2.8401    n = 2535
```

```
AIC(eg.base)
```

```
## [1] 30823.28
```

The variable-coefficient geography formulation (in which geographic egg distributions vary differently in relation to regional SST indices). This was the second best-performing model for pollock egg variance.

```
##
## Family: Tweedie(p=1.99)
## Link function: log
##
## Formula:
## (Cper10m2 + 1) ~ factor(year) + s(lon, lat) + s(doy) + s(bottom_depth,
##     k = 5) + s(lon, lat, by = reg.SST)
##
## Parametric coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.0000     0.0000     NaN      NaN
## factor(year)1988   5.9004     0.3874  15.232  < 2e-16 ***
## factor(year)1991   4.6634     0.3788  12.310  < 2e-16 ***
## factor(year)1992   5.0919     0.4418  11.525  < 2e-16 ***
## factor(year)1993   4.5153     0.3602  12.537  < 2e-16 ***
## factor(year)1994   3.3823     0.3435   9.847  < 2e-16 ***
## factor(year)1995   3.2106     0.2956  10.860  < 2e-16 ***
## factor(year)1996  -3.5963     0.8217  -4.377 1.26e-05 ***
## factor(year)1997   1.5719     0.3763   4.178 3.05e-05 ***
## factor(year)1998   0.3880     0.6499   0.597  0.55055
## factor(year)1999   3.7645     0.3568  10.552  < 2e-16 ***
## factor(year)2000   3.0767     0.5213   5.902 4.09e-09 ***
## factor(year)2002   2.0380     0.3520   5.789 7.99e-09 ***
## factor(year)2003   0.9099     0.4221   2.155  0.03122 *
## factor(year)2005   0.1721     0.3906   0.441  0.65958
## factor(year)2006   0.8630     0.3529   2.446  0.01452 *
## factor(year)2007   3.2945     0.3243  10.159  < 2e-16 ***
## factor(year)2008   3.4996     0.2771  12.629  < 2e-16 ***
## factor(year)2009   0.2262     0.3126   0.724  0.46942
## factor(year)2010   0.8378     0.2794   2.999  0.00274 **
## factor(year)2011   0.8121     0.4112   1.975  0.04841 *
## factor(year)2012   3.2387     0.2113  15.331  < 2e-16 ***
## factor(year)2013   0.8083     0.3949   2.047  0.04076 *
## factor(year)2014   1.6602     0.3054   5.437 5.97e-08 ***
## factor(year)2015  -1.2502     0.4812  -2.598  0.00943 **
## factor(year)2016   1.0347     0.3439   3.009  0.00265 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

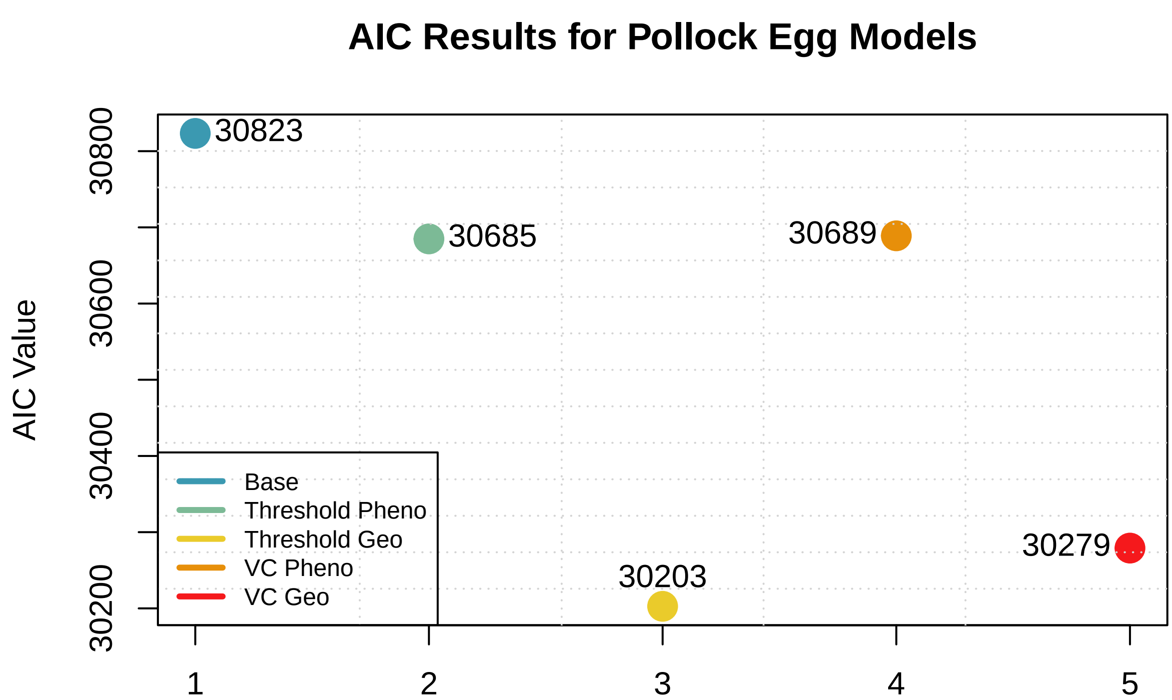
```
## Approximate significance of smooth terms:
##                         edf Ref.df      F  p-value
## s(lon,lat)           27.379 28.656 33.155  < 2e-16 ***
## s(doy)                7.186  8.214 35.618  < 2e-16 ***
## s(bottom_depth)       3.226  3.710  9.383 3.15e-06 ***
## s(lon,lat):reg.SST 28.228 29.626 33.828  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Rank: 97/98
## R-sq.(adj) =  0.137   Deviance explained = 58.3%
## -REML =  15187  Scale est. = 2.4682    n = 2535
```

The threshold geography model formulation (in which the geographic distribution of eggs vary differently above and below a threshold temperature of 1.08 degrees Celsius): *This is the best model to explain pollock egg distribution as of 1/10/2022*

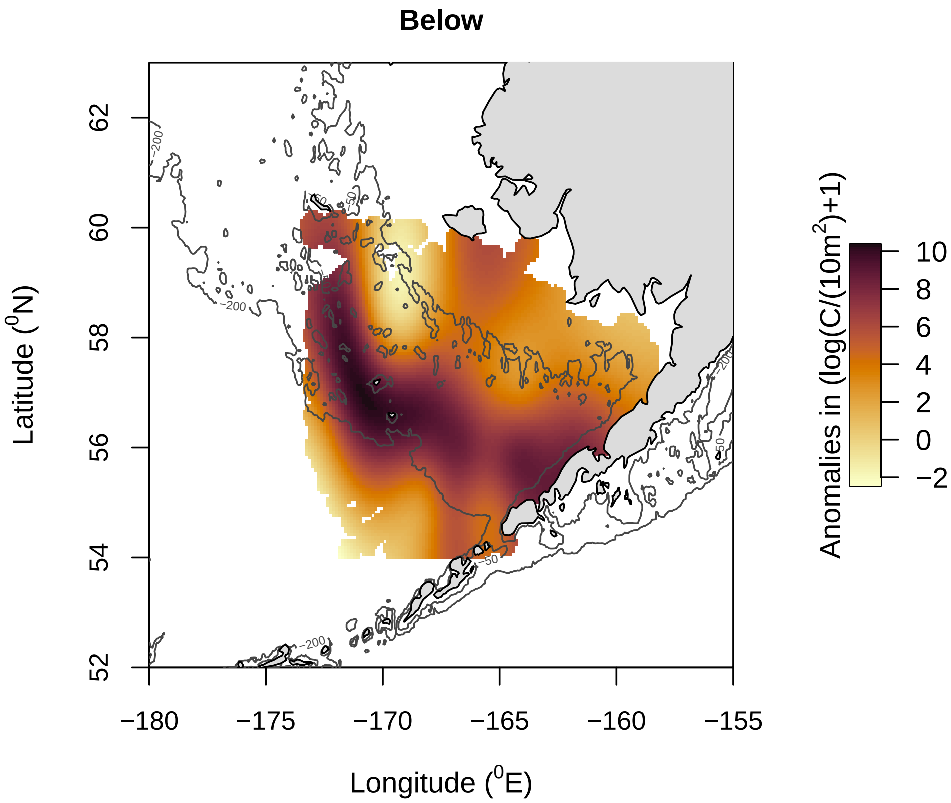
```
##
## Family: Tweedie(p=1.99)
## Link function: log
##
## Formula:
## (Cper10m2 + 1) ~ factor(year) + s(doy) + s(bottom_depth, k = 5) +
##     s(lon, lat, by = th)
##
## Parametric coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       3.022842   0.313413   9.645  < 2e-16 ***
## factor(year)1988  5.943797   0.424435  14.004  < 2e-16 ***
## factor(year)1991  4.866654   0.419649  11.597  < 2e-16 ***
## factor(year)1992  5.384985   0.587448   9.167  < 2e-16 ***
## factor(year)1993  4.714609   0.376491  12.522  < 2e-16 ***
## factor(year)1994  3.288787   0.372141   8.837  < 2e-16 ***
## factor(year)1995  3.339621   0.334889   9.972  < 2e-16 ***
## factor(year)1996 -1.805580   0.774783  -2.330   0.0199 *
## factor(year)1997  1.961116   0.384535   5.100 3.66e-07 ***
## factor(year)1998  0.013463   0.646647   0.021   0.9834
## factor(year)1999  3.305275   0.405758   8.146 5.94e-16 ***
## factor(year)2000  3.810297   0.552412   6.898 6.71e-12 ***
## factor(year)2002  2.209490   0.366156   6.034 1.84e-09 ***
## factor(year)2003  1.554264   0.382927   4.059 5.09e-05 ***
## factor(year)2005  0.896133   0.356708   2.512   0.0121 *
## factor(year)2006  0.873494   0.357539   2.443   0.0146 *
## factor(year)2007  1.896158   0.373175   5.081 4.04e-07 ***
## factor(year)2008  2.421118   0.363725   6.656 3.45e-11 ***
## factor(year)2009  0.003842   0.363250   0.011   0.9916
## factor(year)2010  0.035806   0.360923   0.099   0.9210
## factor(year)2011  0.540327   0.412821   1.309   0.1907
## factor(year)2012  2.458313   0.327965   7.496 9.17e-14 ***
## factor(year)2013  1.067817   0.451606   2.364   0.0181 *
## factor(year)2014  1.614340   0.316724   5.097 3.71e-07 ***
## factor(year)2015  0.390104   0.431208   0.905   0.3657
## factor(year)2016  1.576646   0.316774   4.977 6.90e-07 ***
## ---
```

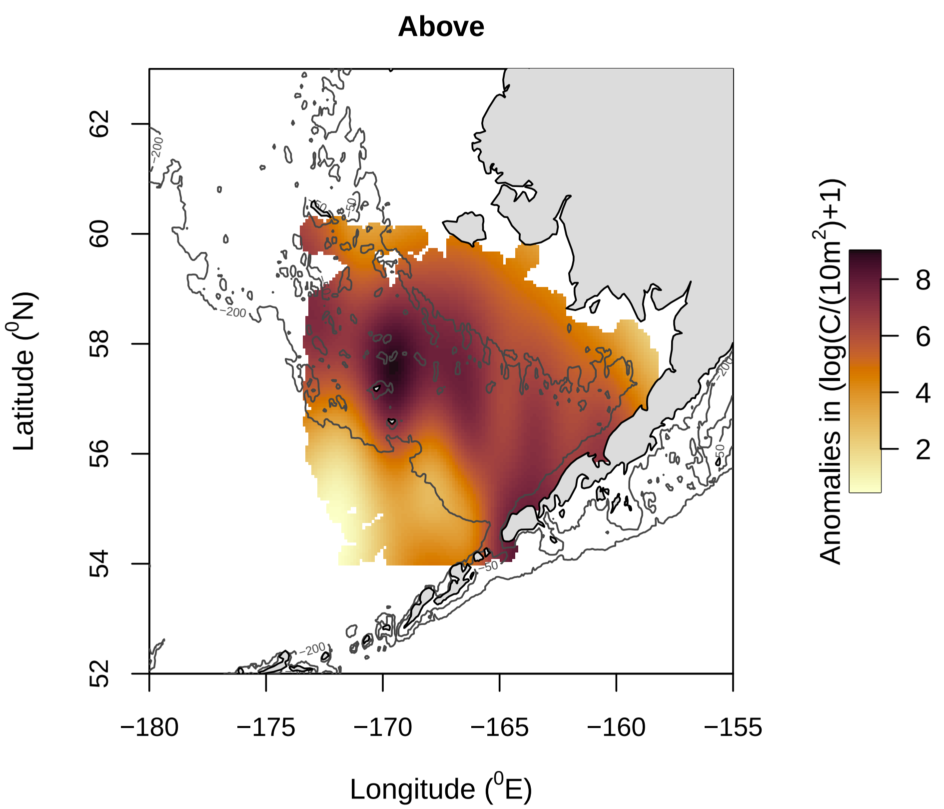
```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                         edf Ref.df      F  p-value
## s(doy)                7.356  8.334 24.950  < 2e-16 ***
## s(bottom_depth)       1.857  2.316  9.761 3.01e-05 ***
## s(lon,lat):thFALSE 27.351 28.821 33.554  < 2e-16 ***
## s(lon,lat):thTRUE  27.930 28.918 49.287  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.166   Deviance explained = 59.2%
## -REML =  15110  Scale est. = 2.423     n = 2535
```

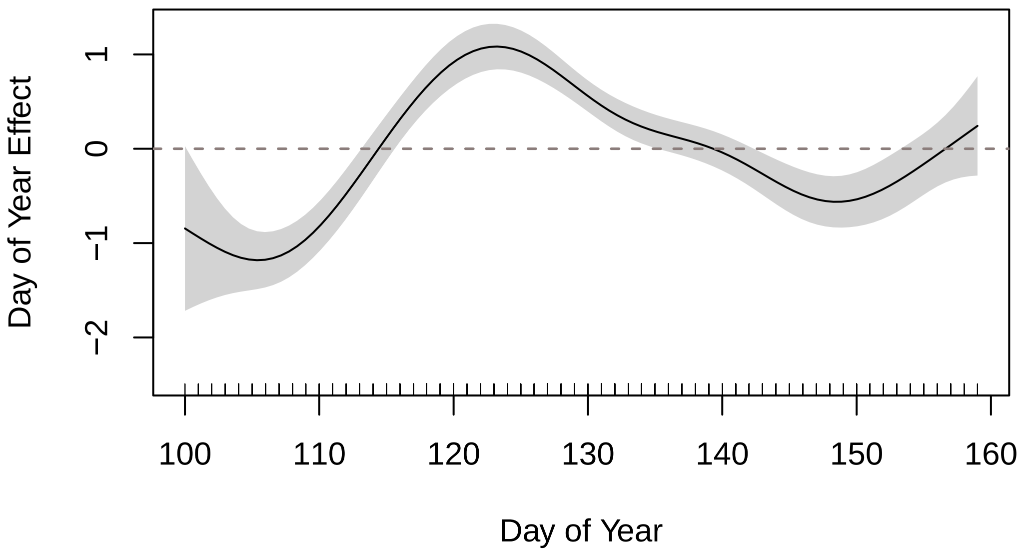
To confirm that the threshold geography model is indeed the best model, we can compare AIC values across all five tested models.

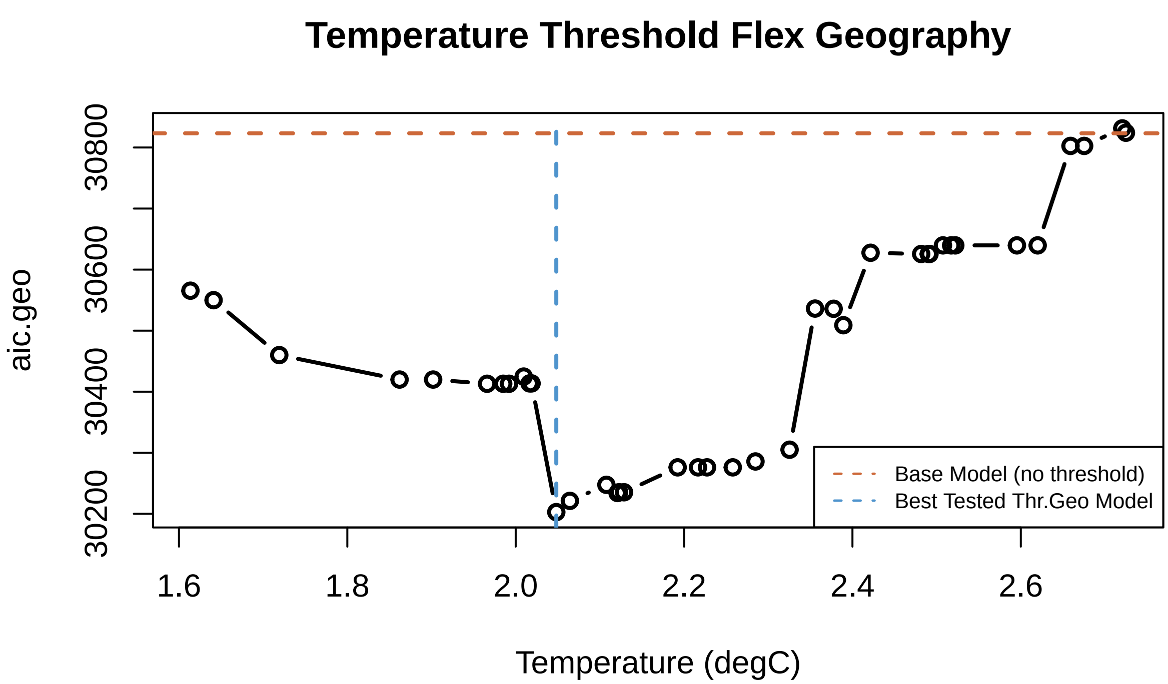


This is the below threshold (1.08 deg C) and above threshold geographic distribution of pollock eggs (based on the threshold geography model).

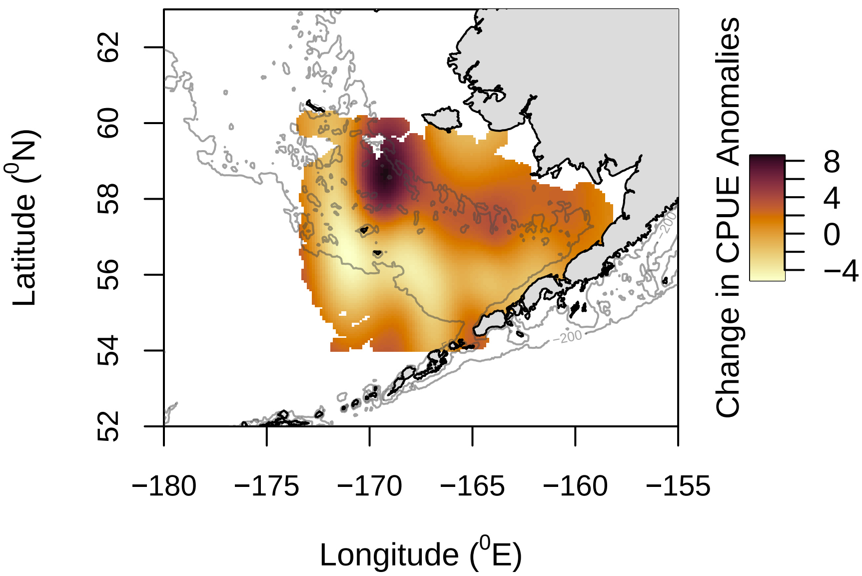

Below
Latitude (^{0}N)
Longitude (^{0}E)
Anomalies in (log(C/(10m^2)+1)
62
60
58
56
54
52
−180
−175
−170
−165
−160
−155
10
8
6
4
2
0
−2
−200
−50


Above
Latitude (^{0}N)
Longitude (^{0}E)
Anomalies in (log(C/(10m^2)+1)
62
60
58
56
54
52
−180
−175
−170
−165
−160
−155
8
6
4
2
−200
−50





With the threshold geography model, we can also see the *significant* differences across the temperature threshold in predicted catch anomalies. This figure shows the prediction of egg catch anomalies below the threshold subtracted from those above the threshold. Thus, more positive values indicate a higher predicted egg catch anomaly above the threshold compared with below the threshold.



Reduction in MSE (%):

```
## [1] 14.68774
```

Larval Generalized Additive Models:

The following code *is only necessary if the data were re-trimmed and new GAMs need to be run.* In this case, modify markdown document such that "{eval = TRUE}". The other model figures are marked as "eval = FALSE" if they, as of the last model run, do not produce the best model results. These models are produced using conductivity-temperature-depth derived temperature and salinity measurements.

We begin with the base larval model:

```
lv.base<-readRDS("./GAM Models/pk_larvae_base.rds")
summary(lv.base)
```

```
##
## Family: Tweedie(p=1.99)
## Link function: log
##
## Formula:
## (Cper10m2 + 1) ~ factor(year) + s(doy, k = 7) + s(lon, lat) +
##     s(bottom_depth, k = 5)
##
## Parametric coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          7.0156     0.2988  23.480  < 2e-16 ***
## factor(year)1998   -4.9083     1.3158  -3.730 0.000198 ***
## factor(year)1999   -1.6987     0.6569  -2.586 0.009804 **
```

```
## factor(year)2000   0.2270     0.5161   0.440 0.660077
## factor(year)2002   0.3765     0.3567   1.055 0.291467
## factor(year)2003  -1.5490     0.3831  -4.043 5.53e-05 ***
## factor(year)2005  -1.5683     0.3456  -4.537 6.14e-06 ***
## factor(year)2006  -1.9752     0.3317  -5.955 3.21e-09 ***
## factor(year)2007  -2.0773     0.3521  -5.900 4.46e-09 ***
## factor(year)2008  -4.5536     0.5289  -8.610  < 2e-16 ***
## factor(year)2009  -4.1886     0.3450 -12.141  < 2e-16 ***
## factor(year)2010  -3.2945     0.3299  -9.986  < 2e-16 ***
## factor(year)2011  -5.7507     0.5012 -11.473  < 2e-16 ***
## factor(year)2012  -2.1844     0.3239  -6.744 2.17e-11 ***
## factor(year)2013  -5.6945     0.4367 -13.040  < 2e-16 ***
## factor(year)2014  -2.0107     0.3278  -6.134 1.09e-09 ***
## factor(year)2015  -0.2141     0.4392  -0.488 0.625950
## factor(year)2016  -1.3706     0.3320  -4.128 3.85e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                    edf Ref.df      F p-value
## s(doy)           5.823  5.980 67.939  <2e-16 ***
## s(lon,lat)      26.598 28.637 25.929  <2e-16 ***
## s(bottom_depth)  1.821  2.238  1.552   0.224
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.212   Deviance explained = 50.7%
## -REML = 9129.1  Scale est. = 2.4193    n = 1596
```

```
AIC(lv.base)
```

```
## [1] 18280.29
```

```
##
## Family: Tweedie(p=1.99)
## Link function: log
##
## Formula:
## (Cper10m2 + 1) ~ factor(year) + s(doy, k = 7) + s(lon, lat) +
##     s(bottom_depth, k = 5) + s(temperature)
##
## Parametric coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        7.1514     0.2821  25.352  < 2e-16 ***
## factor(year)1998  -2.5699     1.0928  -2.352 0.018815 *
## factor(year)1999  -1.2457     0.6086  -2.047 0.040835 *
## factor(year)2000   0.8033     0.4871   1.649 0.099330 .
## factor(year)2002  -1.2176     0.3497  -3.482 0.000512 ***
## factor(year)2003  -2.8689     0.3749  -7.652 3.46e-14 ***
## factor(year)2005  -3.0337     0.3508  -8.648  < 2e-16 ***
## factor(year)2006  -1.8960     0.3131  -6.056 1.75e-09 ***
## factor(year)2007  -1.6325     0.3382  -4.827 1.53e-06 ***
## factor(year)2008  -3.3440     0.5181  -6.454 1.46e-10 ***
```

```
## factor(year)2009  -4.3928     0.3245 -13.537  < 2e-16 ***
## factor(year)2010  -2.2269     0.3155  -7.059 2.54e-12 ***
## factor(year)2011  -4.3868     0.4763  -9.210  < 2e-16 ***
## factor(year)2012  -1.0461     0.3351  -3.122 0.001830 **
## factor(year)2013  -3.9741     0.4282  -9.281  < 2e-16 ***
## factor(year)2014  -3.5168     0.3293 -10.679  < 2e-16 ***
## factor(year)2015  -1.4929     0.4148  -3.599 0.000330 ***
## factor(year)2016  -2.9221     0.3804  -7.682 2.76e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                    edf Ref.df      F p-value
## s(doy)           5.211  5.698 22.324  <2e-16 ***
## s(lon,lat)      27.525 28.861 41.713  <2e-16 ***
## s(bottom_depth)  1.000  1.001  1.332   0.249
## s(temperature)   8.017  8.756 63.418  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  -0.626   Deviance explained = 58.4%
## -REML = 8974.6  Scale est. = 2.0962    n = 1596
```

Then additive temperature and salinity, in individual additive terms. This is the second-best performing model.

```
lv.temp.sal<-readRDS("./GAM Models/pk_larvae_addtempsal.rds")
summary(lv.temp.sal)
```

```
##
## Family: Tweedie(p=1.99)
## Link function: log
##
## Formula:
## (Cper10m2 + 1) ~ factor(year) + s(doy, k = 7) + s(lon, lat) +
##     s(bottom_depth, k = 5) + s(temperature) + s(salinity)
##
## Parametric coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         7.0437     0.2874  24.505  < 2e-16 ***
## factor(year)1998  -2.9139     1.1074  -2.631 0.008594 **
## factor(year)1999  -1.3131     0.6147  -2.136 0.032817 *
## factor(year)2000   0.5760     0.4911   1.173 0.241046
## factor(year)2002  -1.2521     0.3605  -3.473 0.000528 ***
## factor(year)2003  -2.7543     0.3781  -7.284 5.16e-13 ***
## factor(year)2005  -3.0253     0.3570  -8.475  < 2e-16 ***
## factor(year)2006  -1.9622     0.3274  -5.993 2.57e-09 ***
## factor(year)2007  -1.6508     0.3420  -4.827 1.53e-06 ***
## factor(year)2008  -3.4033     0.5221  -6.518 9.64e-11 ***
## factor(year)2009  -4.4572     0.3316 -13.441  < 2e-16 ***
## factor(year)2010  -1.9827     0.3180  -6.236 5.80e-10 ***
## factor(year)2011  -4.2136     0.4821  -8.740  < 2e-16 ***
## factor(year)2012  -1.0241     0.3445  -2.973 0.002993 **
```

```
## factor(year)2013  -4.0379     0.4322  -9.342  < 2e-16 ***
## factor(year)2014  -3.2547     0.3327  -9.782  < 2e-16 ***
## factor(year)2015  -1.3099     0.4167  -3.143 0.001703 **
## factor(year)2016  -2.7322     0.3856  -7.086 2.09e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                     edf Ref.df      F  p-value
## s(doy)            5.269  5.734 24.675  < 2e-16 ***
## s(lon,lat)       27.302 28.807 37.501  < 2e-16 ***
## s(bottom_depth)   1.714  2.111  1.167    0.287
## s(temperature)    7.869  8.683 44.803  < 2e-16 ***
## s(salinity)       7.345  8.368  6.023 5.78e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =   0.25   Deviance explained = 59.3%
## -REML = 8966.1  Scale est. = 2.0657    n = 1596
```

```
AIC(lv.temp.sal)
```

```
## [1] 17904.07
```

And finally, the best performing model: the bivariate salinity-temperature additive term:

```
lv.2d<-readRDS("./GAM Models/pk_larvae_2d.rds")
summary(lv.2d)
```

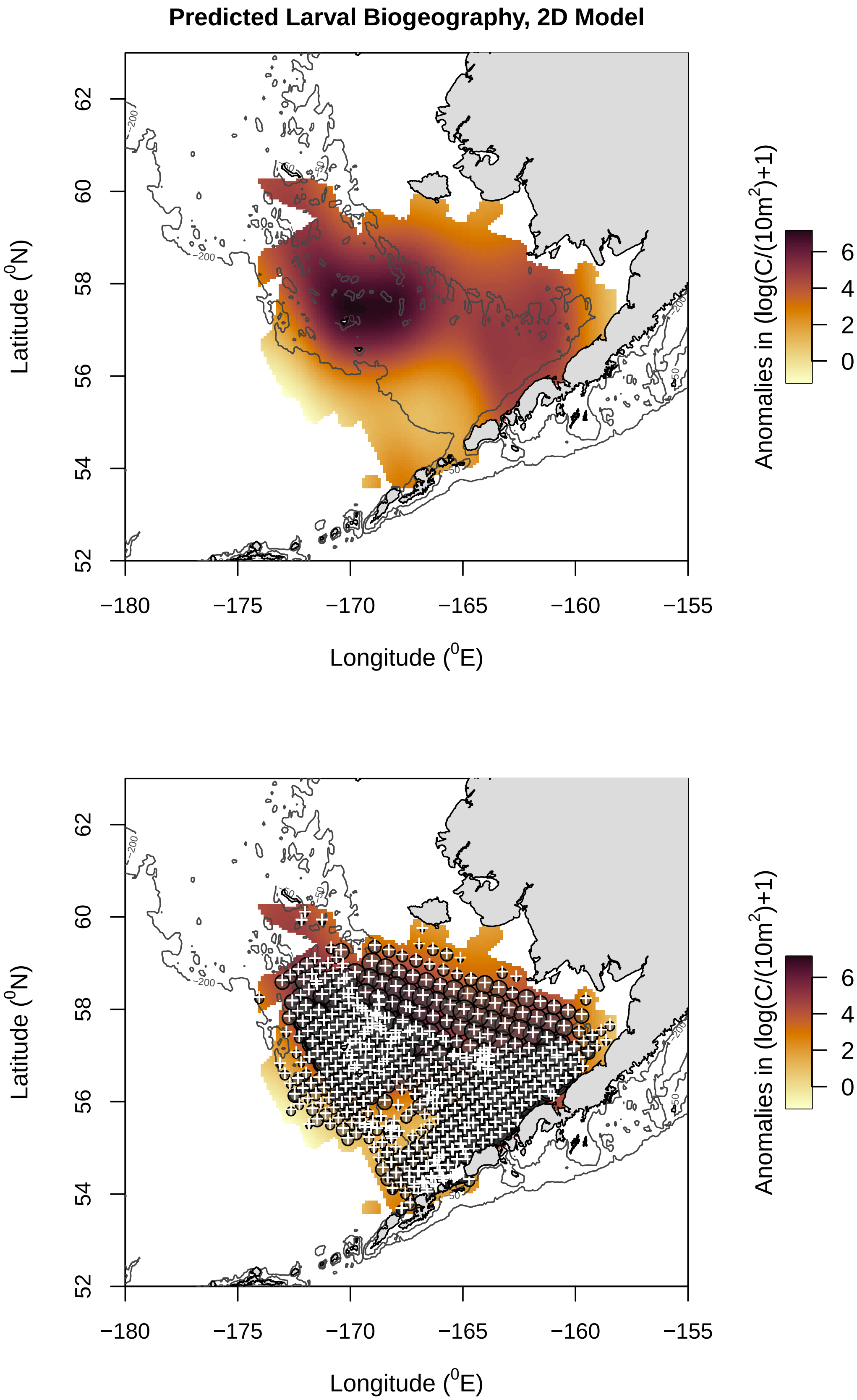
```
##
## Family: Tweedie(p=1.99)
## Link function: log
##
## Formula:
## (Cper10m2 + 1) ~ factor(year) + s(lon, lat) + s(doy, k = 7) +
##     s(bottom_depth) + te(salinity, temperature)
##
## Parametric coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         6.7589     0.2907  23.248  < 2e-16 ***
## factor(year)1998  -1.4468     0.9960  -1.453 0.146540
## factor(year)1999  -0.6398     0.6035  -1.060 0.289254
## factor(year)2000   0.6916     0.4898   1.412 0.158146
## factor(year)2002  -0.5680     0.3542  -1.604 0.108954
## factor(year)2003  -2.3241     0.3865  -6.013 2.27e-09 ***
## factor(year)2005  -2.3463     0.3573  -6.566 7.06e-11 ***
## factor(year)2006  -1.4834     0.3328  -4.457 8.90e-06 ***
## factor(year)2007  -1.2626     0.3500  -3.607 0.000319 ***
## factor(year)2008  -3.0206     0.5290  -5.710 1.36e-08 ***
## factor(year)2009  -3.9922     0.3334 -11.974  < 2e-16 ***
## factor(year)2010  -1.8354     0.3161  -5.806 7.77e-09 ***
## factor(year)2011  -4.7228     0.4941  -9.558  < 2e-16 ***
```

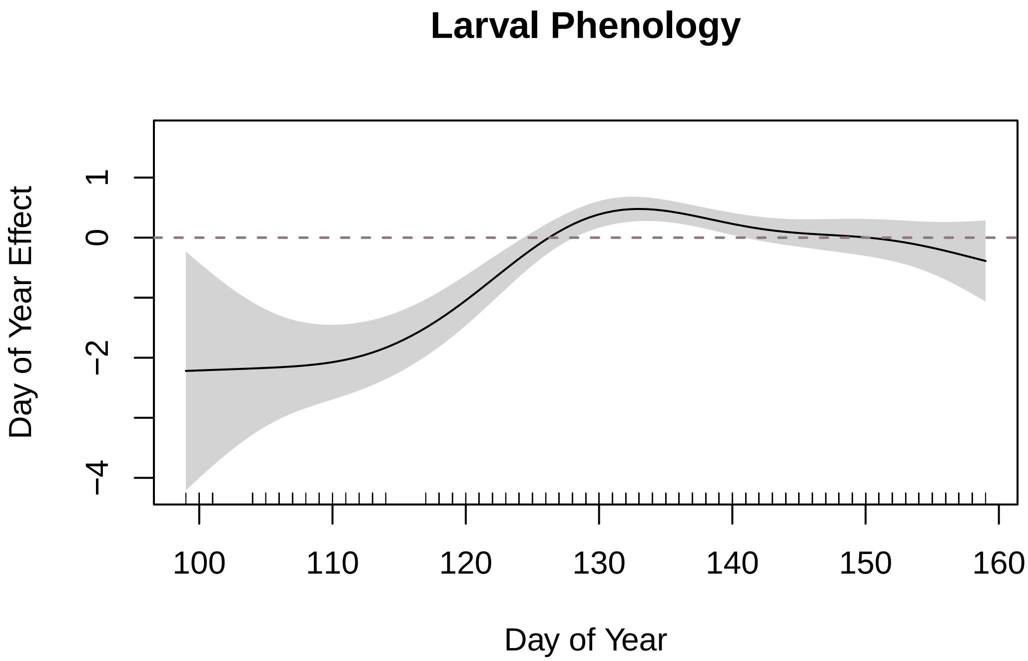
```
## factor(year)2012  -1.0937     0.3500  -3.125 0.001811 **
## factor(year)2013  -3.9406     0.4354  -9.051  < 2e-16 ***
## factor(year)2014  -3.0633     0.3373  -9.081  < 2e-16 ***
## factor(year)2015  -1.6365     0.4234  -3.865 0.000116 ***
## factor(year)2016  -2.4504     0.3864  -6.341 2.99e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                              edf Ref.df      F p-value
## s(lon,lat)                27.263 28.795 37.239  <2e-16 ***
## s(doy)                     4.689  5.326 14.609  <2e-16 ***
## s(bottom_depth)            1.473  1.801  0.986   0.474
## te(salinity,temperature) 18.511 20.580 31.736  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.0341   Deviance explained = 59.7%
## -REML = 8956.8  Scale est. = 2.052     n = 1596
```

```
AIC(lv.2d)
```

```
## [1] 17890.27
```

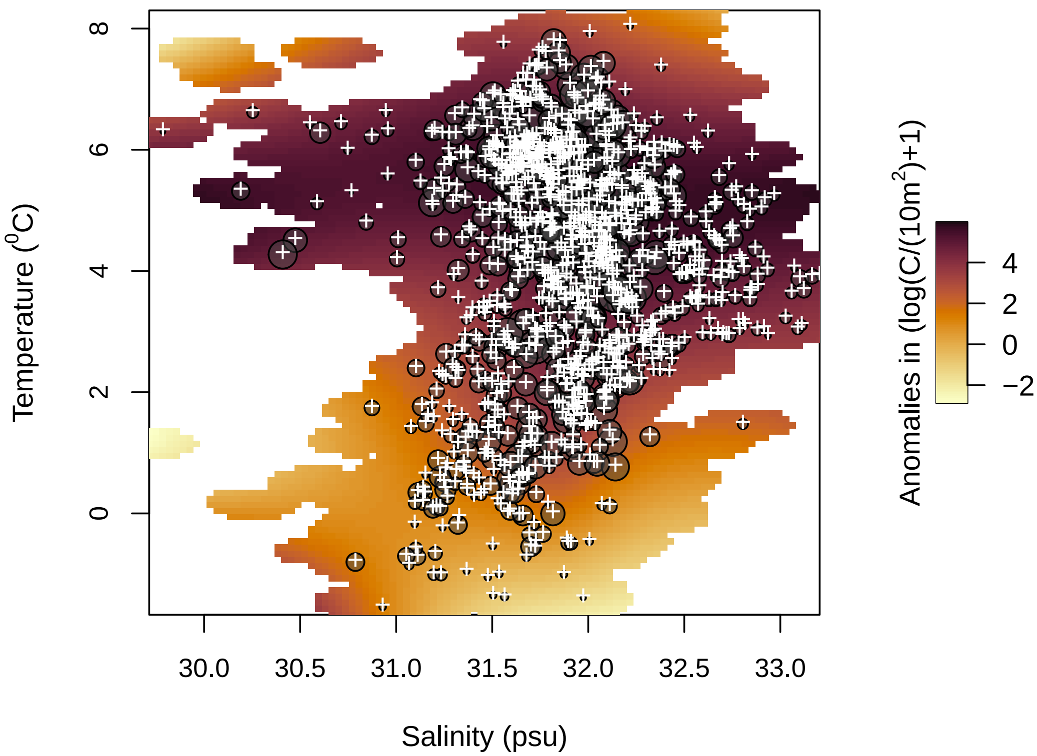
The following plot is the predicted pollock larval biogeography based on the best performing model, the bivariate salinity-temperature GAM. Observations (log transformed, n+1) are shown as well.

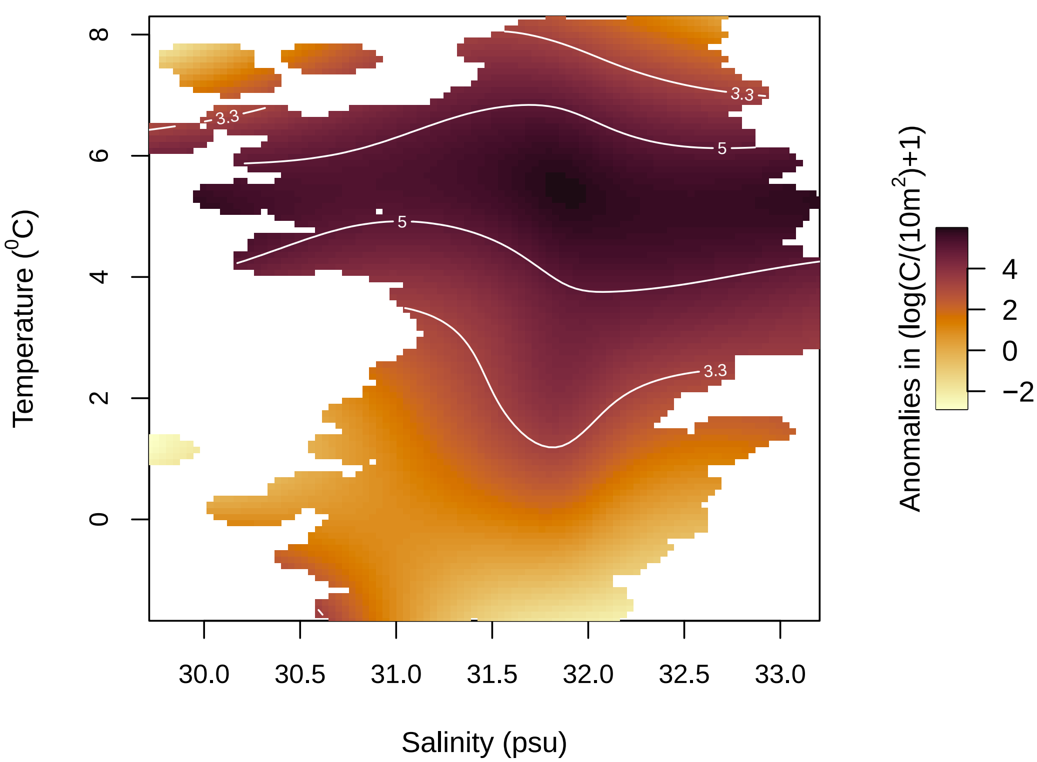

Predicted Larval Biogeography, 2D Model
Latitude (0N)
Longitude (0E)
Anomalies in (log(C/(10m2)+1)
62
60
58
56
54
52
−180
−175
−170
−165
−160
−155
6
4
2
0
−200
−50
Latitude (0N)
Longitude (0E)
Anomalies in (log(C/(10m2)+1)



With this bivariate model, we can also calculate the predicted anomalous larval catch (more or less than expected) on a salinity-temperature plot. This figure shows that prediction, with observed larval catch (log(n=1)) overlaid.

Larval Biogeography by Temperature and Salinity

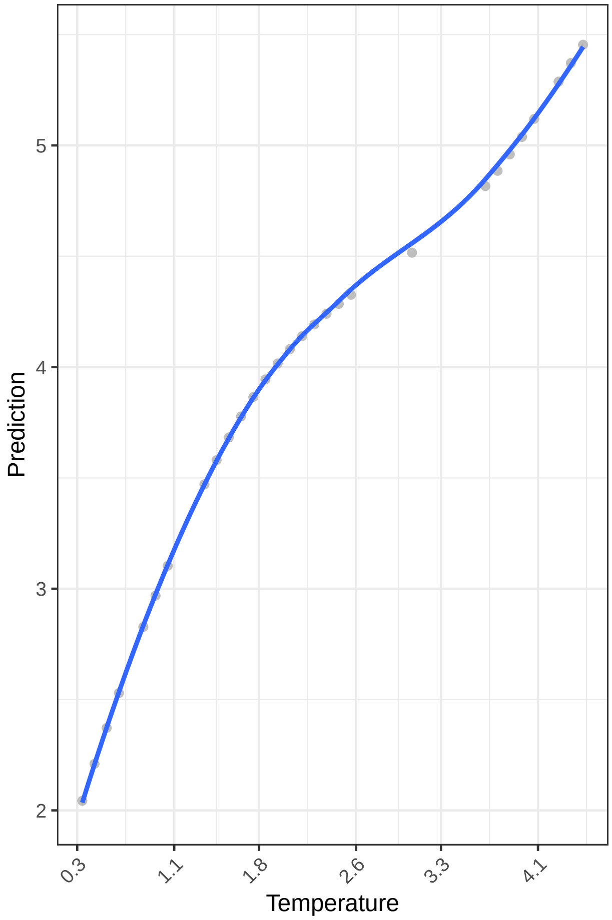


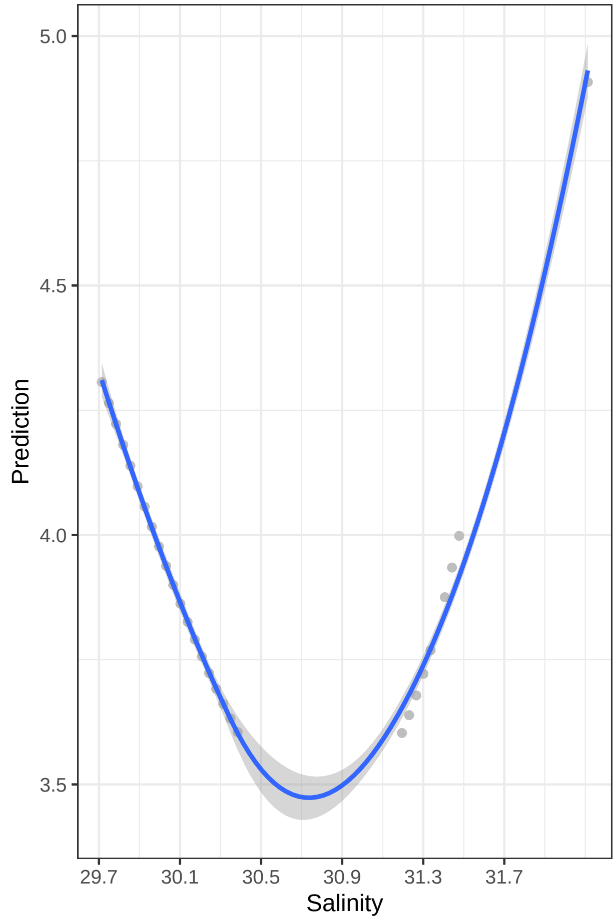


Now I'll calculate a specific range of temperature, salinity, and both temperature and salinity to evaluate breadth of environmental tolerances.

For the univariate predictions, make a grid that holds either temperature or salinity constant depending on the variable of interest.

```
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
```

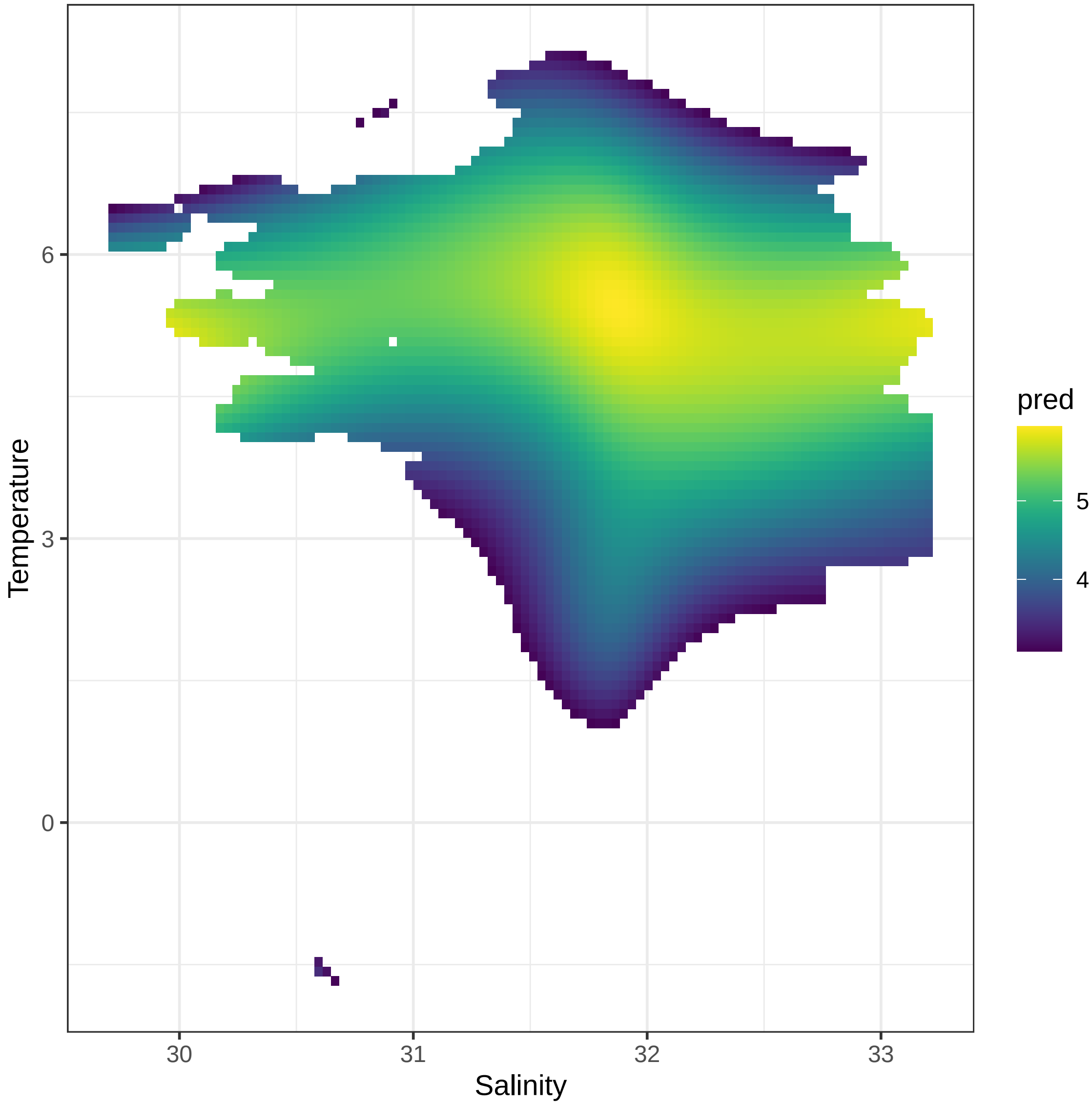




For the bivariate analysis:

```
##   salinity temperature       dist year      lon      lat doy bottom_depth
## 1 31.86506    5.479385 0.02266126 2005 -166.034 56.30985 137           89
## 2 31.86506    5.378647 0.05330482 2005 -166.034 56.30985 137           89
## 3 31.90032    5.378647 0.03513595 2005 -166.034 56.30985 137           89
## 4 31.90032    5.479385 0.04646564 2005 -166.034 56.30985 137           89
## 5 31.82980    5.479385 0.01059603 2005 -166.034 56.30985 137           89
## 6 31.86506    5.580123 0.04638489 2005 -166.034 56.30985 137           89
##       pred
## 1 5.942455
## 2 5.939563
## 3 5.938218
## 4 5.937370
```

```
## 5 5.935396
## 6 5.932506
```



To again share the improvements of the best performing models from the base models, we can look at the AIC division produces.

Table 3: Model Power through AIC Comparisons, Pollock

	Best Divided By Base	Best Divided By Second Best
Eggs	0.9798668	0.9974769
Larvae	0.9786646	0.9992294

Reduction in MSE (%):

```
## [1] 15.1811
```