

A century of Chinook salmon consumption by marine mammal predators in the Northeast Pacific Ocean

Jesse Adams et al.

08-01-2015

Loading in the data

```
#setwd("/users/eric.ward/dropbox/PSC project 2015/PeerJ paper")
# Try readxl package
if("readxl" %in% rownames(installed.packages()) == FALSE) {
  devtools::install_github("hadley/readxl")}
require(readxl)
```

```
## Loading required package: readxl
```

```
require(knitr)
```

```
## Loading required package: knitr
```

```
diets = read_excel("Consumption spreadsheet_noThomas.xlsx")
```

Meta-analysis

First we need to convert percentages to be in 0-1 range:

```
diets$Percent_chinook = as.numeric(diets$Percent_chinook)/100
diets$Percent_salmon = as.numeric(diets$Percent_salmon)/100
```

```
ilogit = function(x) {return(exp(x)/(1+exp(x)))}
# Harbor seal
totalSampleSize = diets$Sample_size
N = dim(diets)[1]
totalSalmon = round(totalSampleSize*diets$Percent_salmon)
totalChinook = round(totalSampleSize*diets$Percent_chinook)
indx = which(is.na(totalSalmon+totalSampleSize)==FALSE &
  diets$Species == "Harbor Seal" & diets$Count_type == "F0")
y = c(rep(rep(1,length(indx)), totalSalmon[indx]),rep(rep(0,length(indx)),
  totalSampleSize[indx]-totalSalmon[indx]))
g = try(glm(y ~ 1, family="binomial"),silent=T)
p.hs = rep(NA,10000)
if(class(g)[1]!="try-error") p.hs = rnorm(10000, coef(summary(g))[1], coef(summary(g))[2])

indx = which(is.na(totalChinook+totalSampleSize)==FALSE &
  diets$Species == "Harbor Seal" & diets$Count_type == "F0")
y = c(rep(rep(1,length(indx)), totalChinook[indx]),rep(rep(0,length(indx)),
```

```

    totalSampleSize[indx]-totalChinook[indx]))
g = try(glm(y ~ 1, family="binomial"),silent=T)
pch.hs = rep(NA,10000)
if(class(g)[1]!="try-error") pch.hs = rnorm(10000, coef(summary(g))[1], coef(summary(g))[2])

indx = which(is.na(totalSalmon+totalSampleSize)==FALSE &
  diets$Species == "Steller Sea lion" & diets$Count_type == "F0")
y = c(rep(rep(1,length(indx)), totalSalmon[indx]),rep(rep(0,length(indx)),
  totalSampleSize[indx]-totalSalmon[indx]))
g = try(glm(y ~ 1, family="binomial"),silent=T)
p.ssl = rep(NA,10000)
if(class(g)[1]!="try-error") p.ssl = rnorm(10000, coef(summary(g))[1], coef(summary(g))[2])

indx = which(is.na(totalChinook+totalSampleSize)==FALSE &
  diets$Species == "Steller Sea lion" & diets$Count_type == "F0")
y = c(rep(rep(1,length(indx)), totalChinook[indx]),rep(rep(0,length(indx)),
  totalSampleSize[indx]-totalChinook[indx]))
g = try(glm(y ~ 1, family="binomial"),silent=T)
pch.ssl = rep(NA,10000)
if(class(g)[1]!="try-error") pch.ssl = rnorm(10000, coef(summary(g))[1], coef(summary(g))[2])

indx = which(is.na(totalSalmon+totalSampleSize)==FALSE &
  diets$Species == "California Sea Lion" & diets$Count_type == "F0")
y = c(rep(rep(1,length(indx)), totalSalmon[indx]),rep(rep(0,length(indx)),
  totalSampleSize[indx]-totalSalmon[indx]))
g = try(glm(y ~ 1, family="binomial"),silent=T)
p.csl = rep(NA,10000)
if(class(g)[1]!="try-error") p.csl = rnorm(10000, coef(summary(g))[1], coef(summary(g))[2])

indx = which(is.na(totalChinook+totalSampleSize)==FALSE &
  diets$Species == "California Sea Lion" & diets$Count_type == "F0")
y = c(rep(rep(1,length(indx)), totalChinook[indx]),rep(rep(0,length(indx)),
  totalSampleSize[indx]-totalChinook[indx]))
g = try(glm(y ~ 1, family="binomial"),silent=T)
pch.csl = rep(NA,10000)
if(class(g)[1]!="try-error") pch.csl = rnorm(10000, coef(summary(g))[1], coef(summary(g))[2])

indx = which(is.na(totalSalmon+totalSampleSize)==FALSE &
  diets$Species == "Killer Whale" & diets$Count_type == "F0")
y = c(rep(rep(1,length(indx)), totalSalmon[indx]),rep(rep(0,length(indx)),
  totalSampleSize[indx]-totalSalmon[indx]))
g = try(glm(y ~ 1, family="binomial"),silent=T)
p.kw = rep(NA,10000)
if(class(g)[1]!="try-error") p.kw = rnorm(10000, coef(summary(g))[1], coef(summary(g))[2])

indx = which(is.na(totalChinook+totalSampleSize)==FALSE &
  diets$Species == "Killer Whale" & diets$Count_type == "F0")
y = c(rep(rep(1,length(indx)), totalChinook[indx]),rep(rep(0,length(indx)),
  totalSampleSize[indx]-totalChinook[indx]))
g = try(glm(y ~ 1, family="binomial"),silent=T)
pch.kw = rep(NA,10000)
if(class(g)[1]!="try-error") pch.kw = rnorm(10000, coef(summary(g))[1], coef(summary(g))[2])

```

```

m = matrix(NA, 4, 3)
m[1,] = round(c(mean(ilogit(p.hs)), quantile(ilogit(p.hs),c(0.025,0.975)))),3)
m[2,] = round(c(mean(ilogit(p.ssl)), quantile(ilogit(p.ssl),c(0.025,0.975)))),3)
m[3,] = round(c(mean(ilogit(p.csl)), quantile(ilogit(p.csl),c(0.025,0.975)))),3)
m[4,] = round(c(mean(ilogit(p.kw)), quantile(ilogit(p.kw),c(0.025,0.975)))),3)
rownames(m) = c("HS", "SSL", "CSL", "KW")
colnames(m) = c("Mean.salmon", "Lower95.salmon", "Upper95.salmon")
kable(m)

```

	Mean.salmon	Lower95.salmon	Upper95.salmon
HS	0.162	0.159	0.167
SSL	0.227	0.221	0.232
CSL	0.077	0.070	0.086
KW	0.977	0.973	0.981

```

m = matrix(NA, 4, 3)
m[1,] = round(c(mean(ilogit(pch.hs)), quantile(ilogit(pch.hs),c(0.025,0.975)))),3)
m[2,] = round(c(mean(ilogit(pch.ssl)), quantile(ilogit(pch.ssl),c(0.025,0.975)))),3)
m[3,] = round(c(mean(ilogit(pch.csl)), quantile(ilogit(pch.csl),c(0.025,0.975)))),3)
m[4,] = round(c(mean(ilogit(pch.kw)), quantile(ilogit(pch.kw),c(0.025,0.975)))),3)
rownames(m) = c("HS", "SSL", "CSL", "KW")
colnames(m) = c("Mean.chinook", "Lower95.chinook", "Upper95.chinook")
kable(m)

```

	Mean.chinook	Lower95.chinook	Upper95.chinook
HS	0.011	0.009	0.013
SSL	0.006	0.001	0.025
CSL	0.021	0.006	0.054
KW	0.631	0.617	0.645

Using the values simulated from the distributions above, we can then generate distributions of the global frequency of occurrence, by species. Note that these estimates are aggregating across space and time.

As a second analysis, we can examine the support for a time trend in the reporting of Chinook in marine mammal diets. This is potentially confounded with both (1) and relative changes in Chinook and other potential prey items, and (2) any diet switching. Each of the studies in our database can be scored as “0” (didn’t report Chinook) or “1” (reported Chinook) and we can fit Binomial GLMs to those data to examine support for temporal trends.

```

diets$reportChinook = ceiling(diets$Percent_chinook)
diets$reportChinook[which(is.na(diets$reportChinook))] = 0

```

We’ll use the mid-point (years) of each study as the temporal predictor.

```

diets$midPoint = (diets$Year_End + diets$Year_Start)/2

```

We’ll restrict the analysis to only studies that include a single predator species, but our model will allow for species-specific intercepts (allowing reporting by species to vary, in part because Chinook plays a different role in the diet of each).

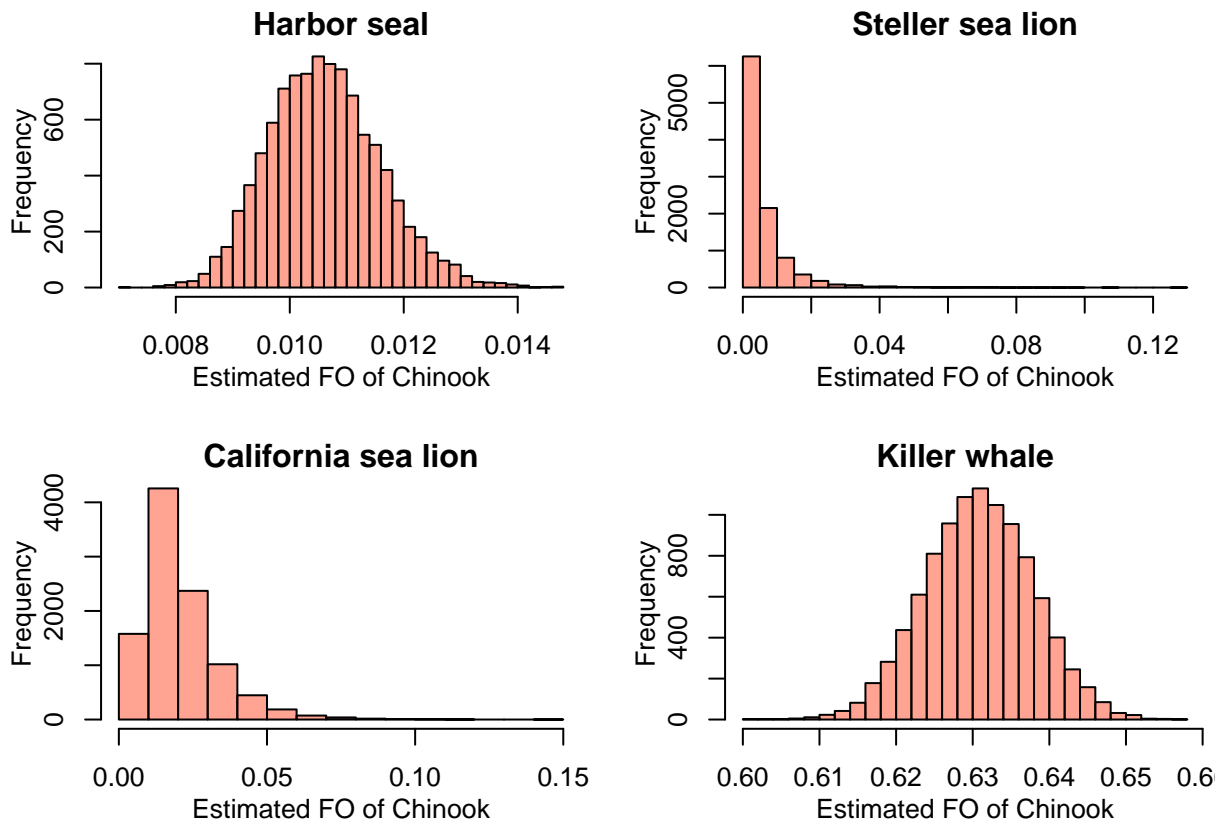


Figure 1: Estimated FO of total salmon by species, across all regions and years

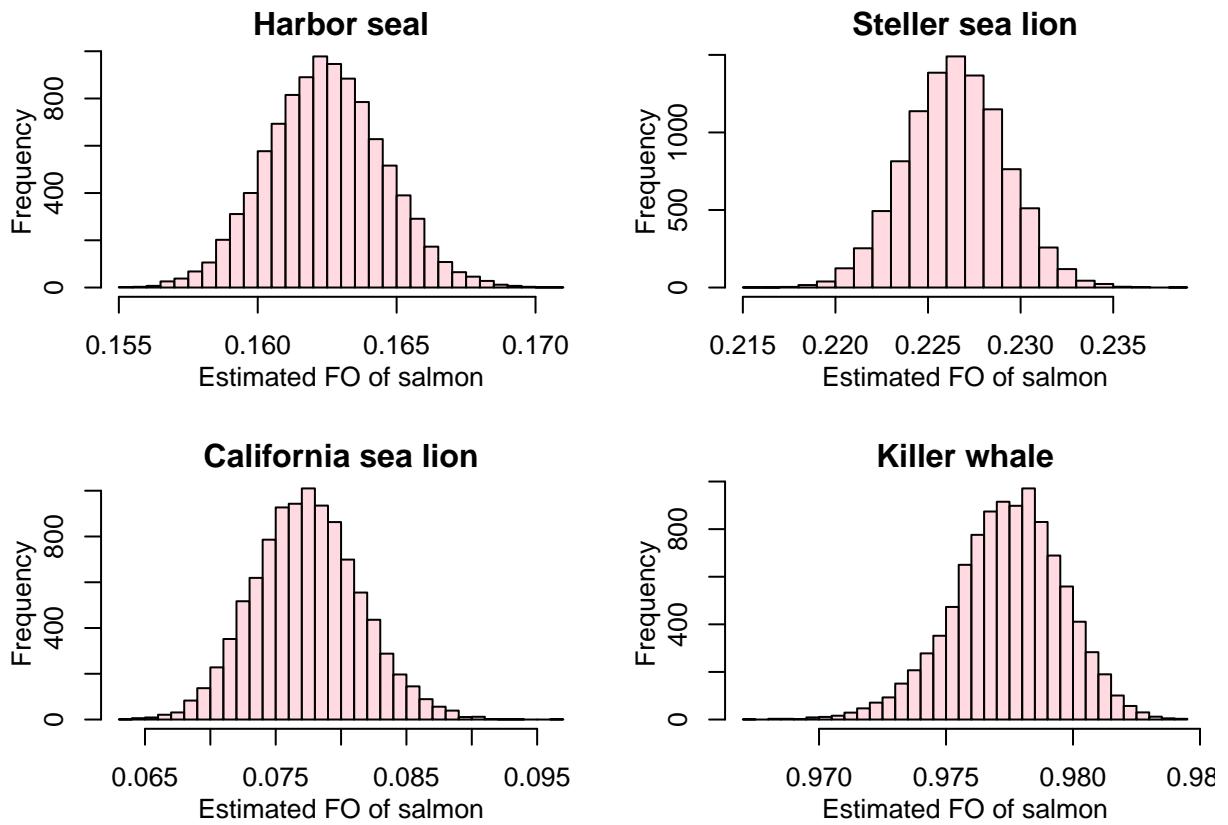


Figure 2: Estimated FO of Chinook salmon by species, across all regions and years

```
g = glm(reportChinook ~ -1+midPoint + Species, family = "binomial",
data = diets[which(diets$Species%in%c("Steller Sea lion","California Sea Lion",
"Killer Whale","Harbor Seal")),])
summary(g)
```

```
##
## Call:
## glm(formula = reportChinook ~ -1 + midPoint + Species, family = "binomial",
##      data = diets[which(diets$Species %in% c("Steller Sea lion",
##        "California Sea Lion", "Killer Whale", "Harbor Seal")),
##      ])
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1921  -0.6224  -0.4631   0.2616   2.5741
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## midPoint           0.08399    0.01958   4.289 1.80e-05 ***
## SpeciesCalifornia Sea Lion -168.96829    39.22651  -4.308 1.65e-05 ***
## SpeciesHarbor Seal        -169.35863    39.14951  -4.326 1.52e-05 ***
## SpeciesKiller Whale       -165.11860    39.00049  -4.234 2.30e-05 ***
## SpeciesSteller Sea lion   -169.76734    39.22000  -4.329 1.50e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 438.07  on 316  degrees of freedom
## Residual deviance: 244.38  on 311  degrees of freedom
## AIC: 254.38
##
## Number of Fisher Scoring iterations: 6
```

```
m = matrix(NA, length(coef(g)), 2)
m = round(coef(summary(g))[1:2],4)
colnames(m) = c("Estimate","SE")
kable(m)
```

	Estimate	SE
midPoint	0.0840	0.0196
SpeciesCalifornia Sea Lion	-168.9683	39.2265
SpeciesHarbor Seal	-169.3586	39.1495
SpeciesKiller Whale	-165.1186	39.0005
SpeciesSteller Sea lion	-169.7673	39.2200

We can create a new data frame and make predictions from this to make plots of the temporal patterns by species. To do this,

```
newdata = expand.grid(midPoint = seq(1970,2015),
Species=c("Steller Sea lion","California Sea Lion","Killer Whale","Harbor Seal"))
```

To predict with this new data frame,

```
newdata$pred = predict(g, newdata = newdata, type="response")
```

And we can plot these trends on the same panel

```
#pdf("Temporal trend.pdf")
Yrs = seq(1970,2015)
plot(Yrs, newdata$pred[which(newdata$Species=="Harbor Seal")], type="l",
     col= "red",ylim=c(0, 1), xlab="", ylab="Probability of studies reporting Chinook",lwd=2)
lines(Yrs, newdata$pred[which(newdata$Species=="Steller Sea lion")], col= "brown",lwd=2)
lines(Yrs, newdata$pred[which(newdata$Species=="California Sea Lion")], col= "gold",lwd=2)
lines(Yrs, newdata$pred[which(newdata$Species=="Killer Whale")], col= "black",lwd=2)
legend("topleft",c("Harbor Seal","Steller Sea lion","California Sea Lion","Killer Whale"),
lwd=2,col=c("red","brown","gold","black"),cex=0.8,bty='n')
```

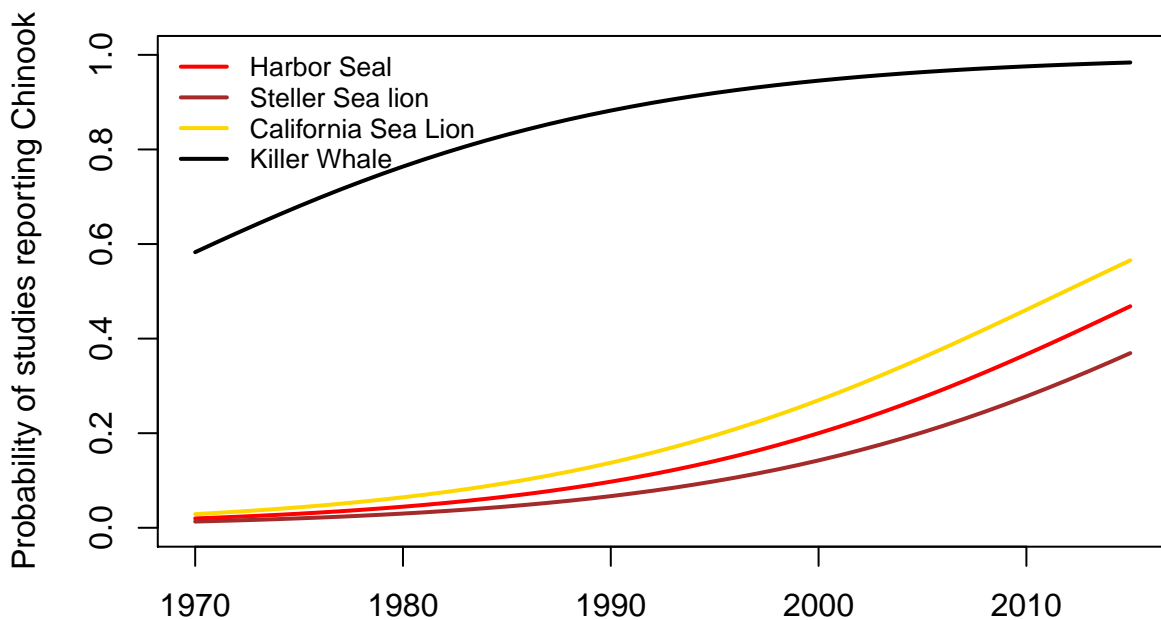


Figure 3: Estimated time trend of studies reporting Chinook salmon in marine mammal diets, across all regions and seasons

```
#dev.off()
```

To examine whether this temporal trend is in part due to shifts in scientific technology (e.g. genetic or molecular methods), we can also try to split out whether or not studies include DNA, and use that as predictors.

```
diets$DNA = 0
for(i in 1:dim(diets)[1]) {
  if(length(grep("DNA",diets$Diet_data[i]))>0) diets$DNA[i] = 1
}
```

Now we can re-fit the GLM with DNA as a predictor,

```
g = glm(reportChinook ~ -1+DNA + midPoint + Species, family = "binomial",
data = diets[which(diets$Species%in%c("Steller Sea lion","California Sea Lion",
"Killer Whale","Harbor Seal")),,])
```

The AIC value of this model actually appears worse than the model above without DNA as a predictor, and the coefficient isn't significant, so the simpler model seems to be more supported.

```
x = read.table("areaCoords.dat", header=TRUE)
sx = read.table("speciesdistribution.dat", header=TRUE)
sxnew = sx[sx$SP!="x",]
library(maps)
library(mapproj)
library(mapdata)
par(mai=c(0,0,0,0), fig=c(0,1,0,1))
map('world', 'usa', xlim=c(-150,-115), ylim=c(35,60), xaxs="i", yaxs="i", fill=TRUE, col='white')

icnt = 1
for(i in levels(x$Area)[1:6])
{
  polygon(x[x$Area==i,2:3], col=icnt)
  icnt = icnt + 1
}

map('world', 'usa', xlim=c(-150,-115), ylim=c(30,75), xaxs="i", yaxs="i", fill=TRUE, add=TRUE, col='white')
map('world', 'canada', add=TRUE, xlim=c(-150,-110), ylim=c(30,60), xaxs="i", yaxs="i", col='white', fill=TRUE)

map('state', region=c('washington'), add=TRUE, xlim=c(-150,-115), ylim=c(30,65), fill=TRUE, col=c('white'))
map('state', region=c('oregon'), add=TRUE, xlim=c(-150,-115), ylim=c(30,65), fill=TRUE, col=c('white'))
map('state', region=c('california'), add=TRUE, xlim=c(-150,-115), ylim=c(30,65), fill=TRUE, col=c('white'))
#map('state', region=c('idaho', 'nevada', 'montana', #'arizona', 'utah', 'wyoming'), add=TRUE, xlim=c(-150,-115), ylim=c(30,65), fill=TRUE, col=c('white'))

icnt = 1
par(mgp=c(2,0.25,0))
for(i in levels(x$Area)[1:6])
{
  par(fig=c(0.1,0.25,0.1+(icnt-1)/7,0.1+(icnt)/7-0.05), mai=c(0,0,0,0), new=TRUE)
  barplot(table(sxnew$SP[sxnew$Region==i])[1:4],
           horiz = TRUE, xlim=c(0,75), las=1, axes=FALSE, space=0, cex.names=0.9, cex.axis=0.65, col=icnt)
  axis(1, cex.axis=0.75)
  icnt = icnt + 1
}
par(fig=c(0,1,0,1))
box()

par(fig=c(0.5,1,0.5,1),mai=c(0,0,0,0), new=TRUE, bg="transparent")
map(xlim=c(-180,-105), ylim=c(20,80), xaxs="i", yaxs="i", lwd=0.5, col="grey")
rect(-135,30,-115,60, lwd=2)
```

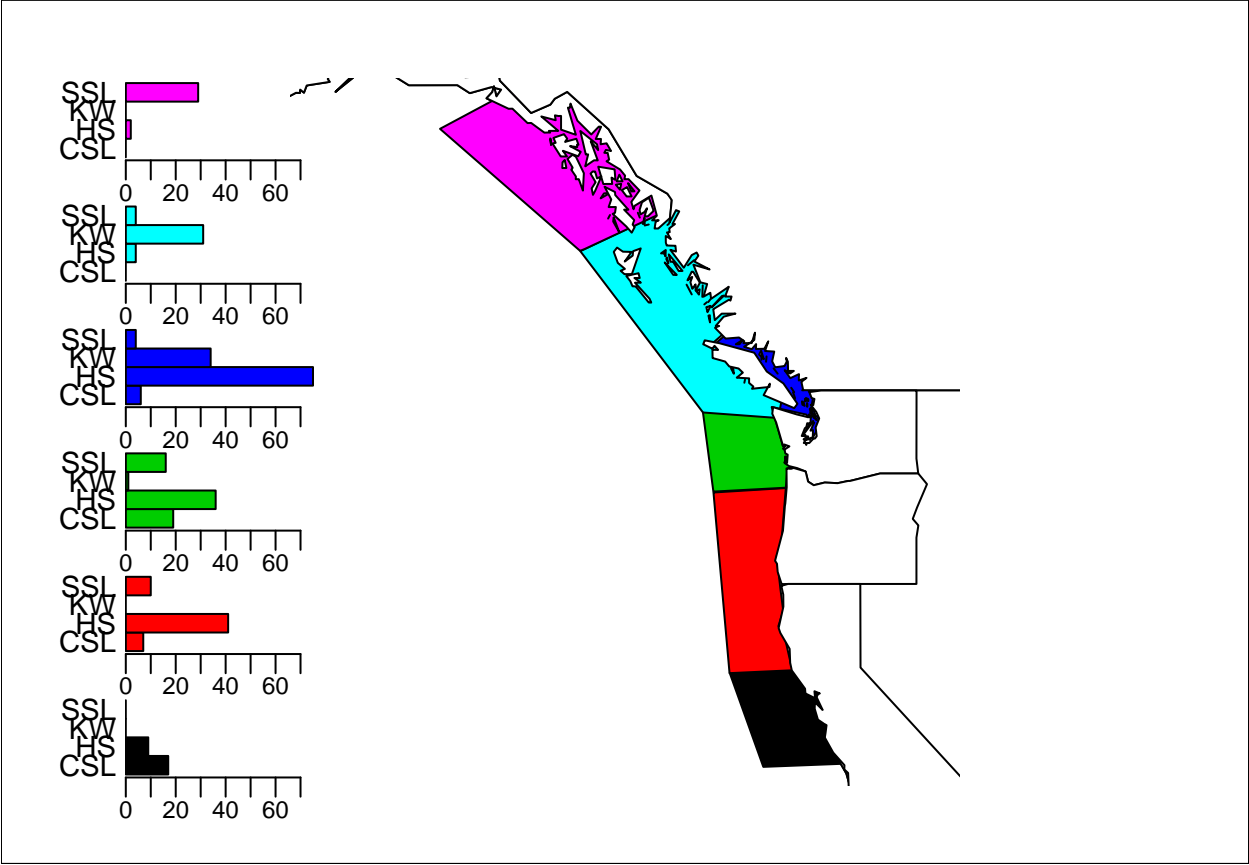



Figure 4:



Figure 5:

```
par(fig=c(0.5,1,0.5,1),mai=c(0,0,0,0), new=TRUE, bg="transparent")
```

```
## Warning in par(fig = c(0.5, 1, 0.5, 1), mai = c(0, 0, 0, 0), new = TRUE, :  
## calling par(new=TRUE) with no plot
```

```
map(xlim=c(-180,-105), ylim=c(20,80), xaxs="i", yaxs="i", lwd=0.5, col="grey")  
rect(-135,30,-115,60, lwd=2)
```



Figure 6: