NOAA FISHERIES

Constructing indexes of abundance

Fish 576, Week 2

What is an index?

- Indexes of abundance give information on relative changes in population abundance or biomass over time, especially when coupled with information on the lengths and ages that index tends to capture.
- **Fishery independent** indexes generally have consistent and well-documented effort, easy to calculate CPUE
  - Used for assessments in this class
- **Fishery dependent** indexes standardize fishery catch rates by information on effort from monitoring programs, logboks, etc.

Index evolution:

Design based → DeltaGLMM → Geospatial models

Design-based indexes

- Essentially a weighted average
- Stratification: partitioning the population so that samples within a **stratum** are more homogeneous than samples across strata.
  - Decreases sampling variance.
- What variables might we stratify by?
  - Possible ideas: Depth, state, habitat type, latitude (note continuous variables must be binned)

Calculation for a single year

- Calculate CPUE for tow i which sweeps over area a_i

$$CPUE_i = C_i / a_i$$

- Calculate average CPUE in stratum j which has n_j tows

$$\overline{CPUE}_j = \frac{1}{n_j} \sum_{i}^{n_j} CPUE_i$$

- Expand to total estimated biomass in stratum i, which has area A_j

$$\hat{b}_j = \overline{CPUE}_j \cdot A_j$$

Calculation, continued

- Sum over all strata to obtain index estimate

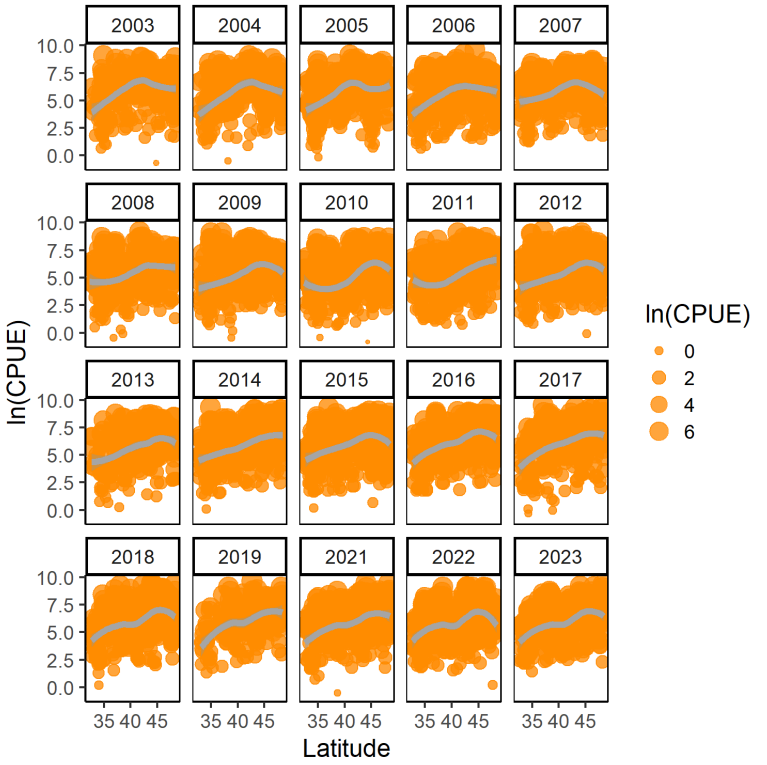
$$\hat{B} = \sum_{j} \hat{b}_j$$

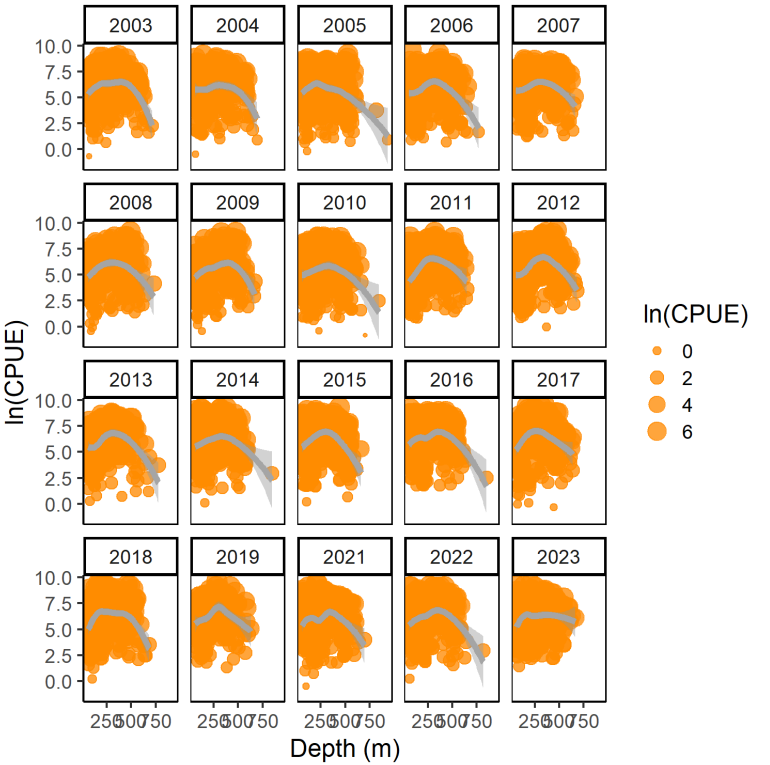
Design-based index in R

```
pak::pkg_install("pfmc-assessments/nwfscSurvey")
```

See vignette on {nwfscSurvey} here

```
library(nwfscSurvey)
# Pull survey data from NWFSC data warehouse
catch <- pull_catch(common_name = "rex sole", # all lowercase
                    survey = "NWFSC.Combo",
                    dir = NULL) # this is the default
plot_cpue(dir = NULL, catch)
```


2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2021
2022
2023
10.0
7.5
5.0
2.5
0.0
ln(CPUE)
35 40 45
Latitude
ln(CPUE)
0
2
4
6


2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2021
2022
2023
10.0
7.5
5.0
2.5
0.0
ln(CPUE)
250 500 750
Depth (m)
ln(CPUE)
0
2
4
6

Defining strata

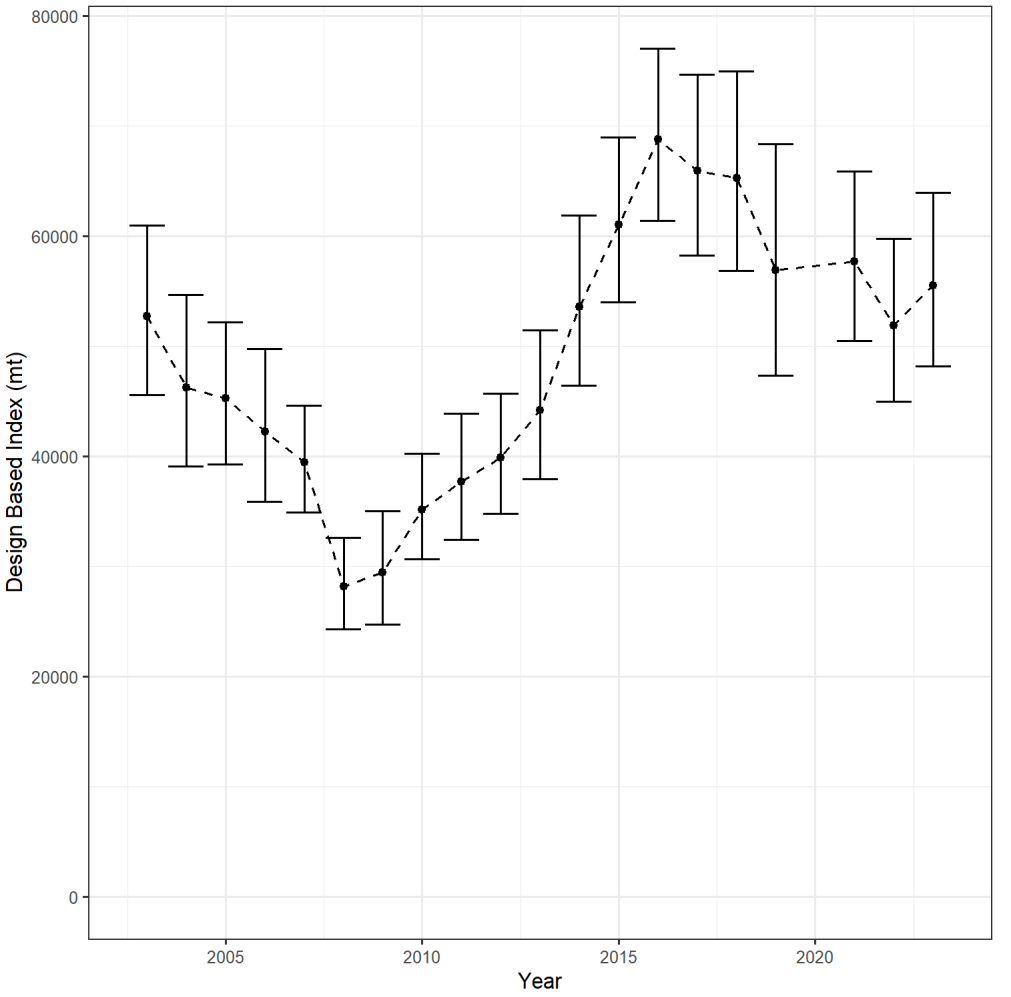
```
strata <- CreateStrataDF.fn(
  names = c("shallow_s", "mid_s", "deep_s", "shallow_n", "mid_
  depths.shallow = c( 55,   200, 300,    55, 200, 300),
  depths.deep    = c(200,   300, 400,   200, 300, 400),
  lats.south     = c( 32,    32,  32,    42,  42,  42),
  lats.north     = c( 42,    42,  42,    49,  49,  49))
strata
```

```
       name      area Depth_m.1 Depth_m.2 Latitude_dd.1 Latitude_dd.
1 shallow_s 19541.690        55       200            32           4
2     mid_s  4335.806       200       300            32           4
3    deep_s  4115.957       300       400            32           4
4 shallow_n 21852.124        55       200            42           4
5     mid_n  3271.327       200       300            42           4
6    deep_n  2608.131       300       400            42           4
```

Calculating index

```
biomass <- get_design_based(data = catch,
                            strata = strata)

plot_index(data = biomass, plot = 1)
```


80000
60000
40000
20000
0
Design Based Index (mt)
2005
2010
2015
2020
Year

Exercise

Calculate design-based index for a survey for one of the two species. Options are:

- Yelloweye rockfish or Widow rockfish
- Triennial or WCGBTS (`"NWFSC.Combo"`)

Problems with design-based index

- High year-to-year variance for species that cluster or school, like many rockfish
- Hard to account for auxiliary information such as vessel effects
- Does not handle unbalanced sampling designs well
- Observations near each other should be more similar, information sharing could be used to decrease variance

DeltaGLMM

- Encounter model

$$\Pr(C_{i,j,y,v} > 0) = \text{logit}^{-1}(Y_{p,y} + S_{p,j} + \gamma_{p,v}), \gamma_{p,v} \sim N(0, \sigma_p)$$

- Positive catch rate model

$$\mathrm{E}[C_{i,j,y,v} | C_{i,j,y,v} > 0] = \exp(Y_{r,y} + S_{r,j} + \gamma_{r,v})$$
$$\gamma_{r,v} \sim N(0, \sigma_r)$$

- $C_{i,j,y,v} | C_{i,j,y,v} > 0$ follows continuous positive probability distribution (e.g., lognormal, gamma)
- y is year, v is vessel, Y, S, γ are year, stratum, and vessel effects, respectively

DeltaGLMM index

$$\hat{B}_y = \sum_j A_j \text{logit}^{-1}(Y_{p,y} + S_{p,j}) \cdot \exp(Y_{r,y} + S_{r,j})$$

- Note that we dropped the vessel effect
- This is a simple example. What else could you include?
  - Year-stratum interactions, pass (early/late), ?

Geostatistical models

- deltaGLMMs solve many problems with design-based indexes, but do not permit information sharing among locations close to each other in space and time
- Geostatistical encounter model

$$\Pr(C_{i,j,y,v,s} > 0) = \text{logit}^{-1}(Y_{p,y} + S_{p,j} + \gamma_{p,v} + \omega_{p,s} + \psi_{p,s,y})$$

$$\gamma_{p,v} \sim N(0, \sigma_p)$$

$$\bar{\omega}_p \sim \text{MVN}(0, \Sigma_{p,\omega}), \bar{\psi}_{p,y} \sim \text{MVN}(0, \Sigma_{p,\psi})$$

- Positive catch rate model adjusted similarly

Geostatistical index

$$\hat{B}_y = \int_s \text{logit}^{-1}(Y_{p,y} + S_{p,j} + \omega_{p,s} + \psi_{p,s,y}) \cdot \exp(Y_{r,y} + S_{r,j} + \omega_{r,s} + \psi_{r,s,y}) ds$$

$$\approx \sum_{s_k \in \text{grid}} \left[\text{logit}^{-1}(Y_{p,y} + S_{p,j} + \omega_{p,s} + \psi_{p,s,y}) \cdot \exp(Y_{r,y} + S_{r,j} + \omega_{r,s} + \psi_{r,s,y}) \right]$$

- NWFSC staff will be producing all indices for the WCGBTS and Triennial survey using {sdmTMB}, which fits geostatistical models
- IPHC survey may be provided directly by WDFW.
- You will need to work up the ORBS index in the yelloweye assessment.