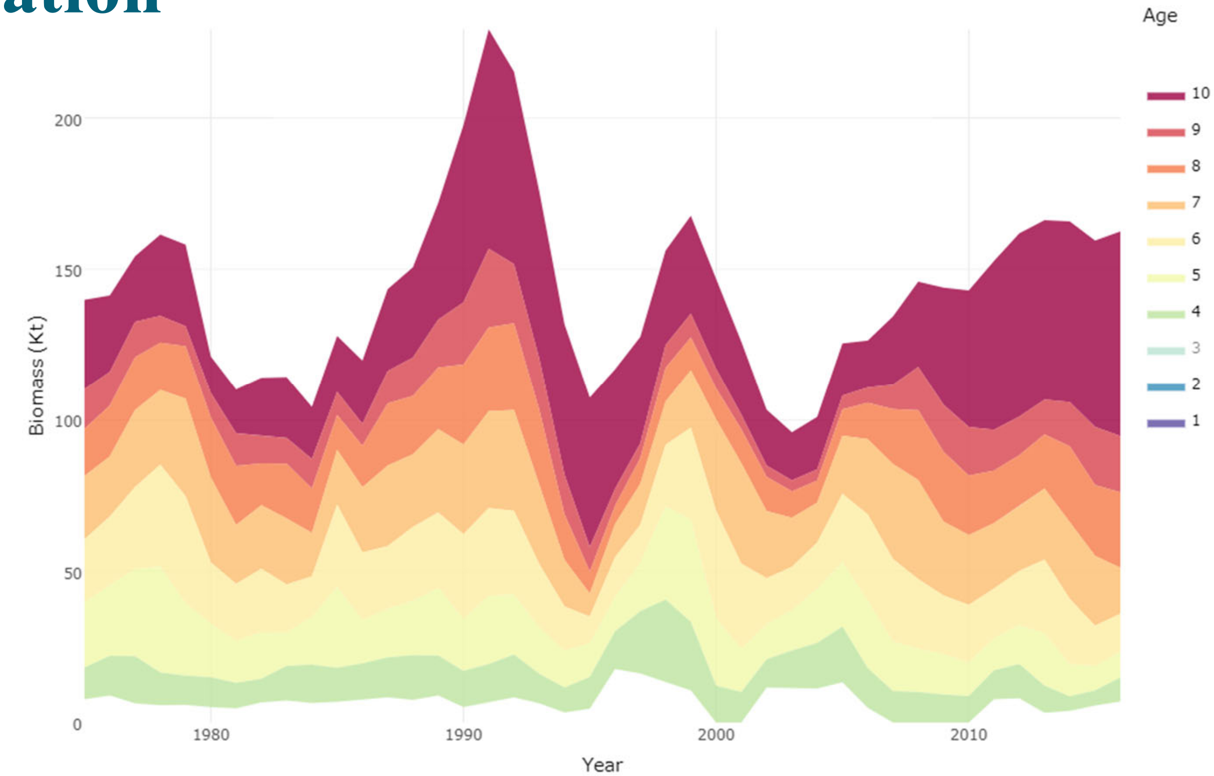
Marine Institute

Lecture 9: Age Structured Population Models

Noel Cadigan



CFER

Centre for Fisheries Ecosystems Research

MARINE INSTITUTE

MEMORIAL UNIVERSITY

F6004 Lecture 9 Outline

Age Structured Population Models (Haddon Chapter 11)

1) Why is Age Important?
2) Cohort Strength Model for pre-exploitation ages
    a. Estimation
    b. Case Study: 3NO cod
3) Cohort Strength Model with Random year Effects
4) SURBA: Survey-based cohort model
    a. Estimation
    b. Case Study: 3NO cod
5) Retrospective Analysis
6) Harvest Advice
    a. Catch multiplier projections

Why is Age so Important?

- We saw for surplus production models that we need a long time-series of varying population size and catch levels to get reliable parameter estimates.
- These models do not know about the size or age-structure of the population and are less accurate for short-term fisheries management
- For example, if a population with total biomass B primarily consists of young fish then we can expect biomass to increase more (due to body growth) than if the population consists of old fish, in which case body growth has largely concluded and mortality is the only outcome

Why is Age so Important?

- The production function of a stock depends on the stocks size or age composition
- But production models assume this is constant
- Production models try to describe stock productivity dynamics based on little information
- Age-based models have more specific information on reproduction which helps untangle r and K
- Tracking how survey catches change with age for a cohort gives information on total mortality rates (Z's)
- But there may still be issues, as we shall see

A simple cohort strength model for young pre-exploitation (i.e. not fished) ages

- A very simple model that is sometimes used in stock assessment.
- Involves only pre-exploitation ages in which fishing has no or little direct impact on abundance
- The cohort population dynamics model is

$$N_{ay} = N_{a-1,y-1} \exp(-M)$$

- N_{ay} is the abundance (i.e. total stock numbers) at age a in year y, and M is the annual natural mortality rate.

$$M = -log(N_{ay}/N_{a-1,y-1}) \geq 0$$

- Fraction that survive: $S = exp(-M)$

A first step for cohort population dynamics model

A Simple Cohort Strength Model

- For example, if there are 5 ages and 8 years, the population dynamics model for N_{ay} can be depicted as

	Age				
Year	1	2	3	4	5
1	N_{11} S	N_{21} S	N_{31} S	N_{41} S	N_{51}
2	N_{12}	N_{22} S	N_{32} S	N_{42} S	N_{52}
3	N_{13}	N_{23}	N_{33} S	N_{43} S	N_{53}
4	N_{14}	N_{24}	N_{34}	N_{44} S	N_{54}
5	N_{15}	N_{25}	N_{35}	N_{45}	N_{55}
6	N_{16}	N_{26}	N_{36}	N_{46}	N_{56}
7	N_{17}	N_{27}	N_{37}	N_{47}	N_{57}
8	N_{18}	N_{28}	N_{38}	N_{48}	N_{58}

A Simple Cohort Strength Model

Year	Age 1	Age 2	Age 3	Age 4	Age 5
1	N_{11}	N_{21}	N_{31}	N_{41}	N_{51}
2	N_{12}	N_{22}	N_{32}	N_{42}	N_{52}
3	N_{13}	N_{23}	N_{33}	N_{43}	N_{53}
4	N_{14}	N_{24}	N_{34}	N_{44}	N_{54}
5	N_{15}	N_{25}	N_{35}	N_{45}	N_{55}
6	N_{16}	N_{26}	N_{36}	N_{46}	N_{56}
7	N_{17}	N_{27}	N_{37}	N_{47}	N_{57}
8	N_{18}	N_{28}	N_{38}	N_{48}	N_{58}

Age composition in year 1

recruitment

Year	Age 1	Age 2	Age 3	Age 4	Age 5
1	N_{11}	N_{21}	N_{31}	N_{41}	N_{51}
2	N_{12}	N_{22}	N_{32}	N_{42}	N_{52}
3	N_{13}	N_{23}	N_{33}	N_{43}	N_{53}
4	N_{14}	N_{24}	N_{34}	N_{44}	N_{54}
5	N_{15}	N_{25}	N_{35}	N_{45}	N_{55}
6	N_{16}	N_{26}	N_{36}	N_{46}	N_{56}
7	N_{17}	N_{27}	N_{37}	N_{47}	N_{57}
8	N_{18}	N_{28}	N_{38}	N_{48}	N_{58}

A Simple Cohort Strength Model

	Age				
Year	1	2	3	4	5
1	N_{11}	N_{21}	N_{31}	N_{41}	N_{51}
2	N_{12}	N_{22}	N_{32}	N_{42}	N_{52}
3	N_{13}	N_{23}	N_{33}	N_{43}	N_{53}
4	N_{14}	N_{24}	N_{34}	N_{44}	N_{54}
5	N_{15}	N_{25}	N_{35}	N_{45}	N_{55}
6	N_{16}	N_{26}	N_{36}	N_{46}	N_{56}
7	N_{17}	N_{27}	N_{37}	N_{47}	N_{57}
8	N_{18}	N_{28}	N_{38}	N_{48}	N_{58}

Can estimate all N_{ay} if we know recruits, initial numbers at age, and S

Similarly, we can estimate all N_{ay} if we know recruits, initial numbers at age, S, and work backwards

	Age				
Year	1	2	3	4	5
1	N_{11}	N_{21}	N_{31}	N_{41}	N_{51}
2	N_{12}	N_{22}	N_{32}	N_{42}	N_{52}
3	N_{13}	N_{23}	N_{33}	N_{43}	N_{53}
4	N_{14}	N_{24}	N_{34}	N_{44}	N_{54}
5	N_{15}	N_{25}	N_{35}	N_{45}	N_{55}
6	N_{16}	N_{26}	N_{36}	N_{46}	N_{56}
7	N_{17}	N_{27}	N_{37}	N_{47}	N_{57}
8	N_{18}	N_{28}	N_{38}	N_{48}	N_{58}

A Simple Cohort Strength Model

	Age				
Year	1	2	3	4	5
1	N_{11}	N_{21}	N_{31}	N_{41}	N_{51}
2	N_{12}	N_{22}	N_{32}	N_{42}	N_{52}
3	N_{13}	N_{23}	N_{33}	N_{43}	N_{53}
4	N_{14}	N_{24}	N_{34}	N_{44}	N_{54}
5	N_{15}	N_{25}	N_{35}	N_{45}	N_{55}
6	N_{16}	N_{26}	N_{36}	N_{46}	N_{56}
7	N_{17}	N_{27}	N_{37}	N_{47}	N_{57}
8	N_{18}	N_{28}	N_{38}	N_{48}	N_{58}

Can estimate all N_{ay} if we know recruits, initial numbers at age, and S

Similarly, we can estimate all N_{ay} if we know recruits, initial numbers at age, S, and work backwards

Age				
1	2	3	4	5
N_{11}	N_{21}	N_{31}	N_{41}	N_{51}
N_{12}	N_{22}	N_{32}	N_{42}	N_{52}
N_{13}	N_{23}	N_{33}	N_{43}	N_{53}
N_{14}	N_{24}	N_{34}	N_{44}	N_{54}
N_{15}	N_{25}	N_{35}	N_{45}	N_{55}
N_{16}	N_{26}	N_{36}	N_{46}	N_{56}
N_{17}	N_{27}	N_{37}	N_{47}	N_{57}
N_{18}	N_{28}	N_{38}	N_{48}	N_{58}

- The N's are all unknown and *practically* unobservable
- More specifically, the N's as hidden (i.e. latent) from observation
- The cohort model describes the basic population dynamics
- More specifically, the cohort model gives the "state" of the population
- Cohort model is the <u>state equation</u>

A Simple Cohort Strength Model

- Let *c = y - a* indicate the cohort and
- let $n_{ay} = \log\left(N_{ay}\right)$.
- One can show through recursive applications of $N_{ay} = N_{a-1,y-1} \exp(-M)$
- that

$$N_{ac} = N_{oc} \exp(-aM)$$

- and that

$$n_{ac} = n_{0c} - aM, a > 0$$

A Simple Cohort Strength Model

- Let R_{ac} denote the survey or fishery CPUE index for age a, and cohort c. [F6004 rarely uses fishery CPUE]
- Each age-class has a possibly different catchability, Q_a, which is a product of gear selectivity and availability of fish to the survey – and Q_a's are unknown parameters to estimate
- The index observation equation is

t_f - fraction of year that survey occurs in

$$R_{ac} = Q_a N_{ac} \exp\{-t_f M\} exp(\varepsilon_{ay})$$

- If $r_{ac} = \log(R_{ac})$ and $q_a = \log(Q_a)$
- Then $r_{ac} = q_a + n_{ac} - t_f M + \varepsilon_{ay}$

Log-index observation equation

$$r_{ac} = q_a + n_{0c} - aM - t_f M + \varepsilon_{ay}$$

Random observation errors

A Simple Cohort Strength Model

- We don't know the values of q_a or M
- These are completely confounded
- We simply estimate $\mathrm{q}_\mathrm{a}^* = \mathrm{q}_\mathrm{a} - \mathrm{aM} - t_f \mathrm{M}$

$$r_{ac} = \mathrm{q}_\mathrm{a}^* + n_{0c} + \varepsilon_{ay}$$

- That's why in a cohort strength model we say that q_a^* represents survey catchability (i.e. selectivity and availability) plus cumulative mortality rates
- M can also be age-specific without loss of generality 'wlog' but cannot depend on year
- The statistical expected value of r_{ac} is

$$E(r_{ac}) = \mu_{ac} = \mathrm{q}_\mathrm{a}^* + n_{0c}$$

Age-specific M

- We assumed that *M*'s were the same at each age, but that will not be true for the young ages that we consider in a cohort-strength model.
- For example, we expect $M_0 >> M_1 >> M_2$.
- This does not affect the model:
- $r_{ac} = q_a + n_{ac} - t_f M_a + \varepsilon_{ay}$

$$r_{ac} = q_a + n_{0c} - \sum_{i=0}^{a-1} M_a - t_f M_a + \varepsilon_{ay}$$

- We simply estimate $\mathrm{q}_\mathrm{a}^* = \mathrm{q}_\mathrm{a} - \sum_{i=0}^{a-1} M_a - t_f M_a$

$$r_{ac} = \mathrm{q}_\mathrm{a}^* + n_{0c} + \varepsilon_{ay}$$

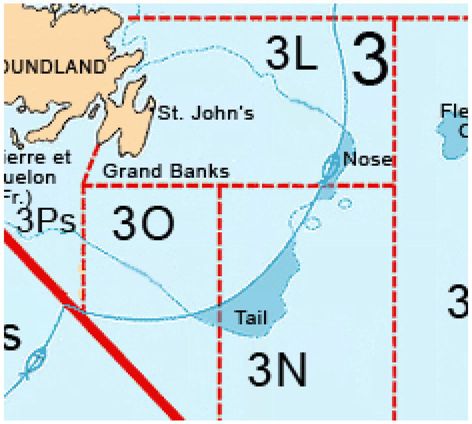
Same model as previous slide

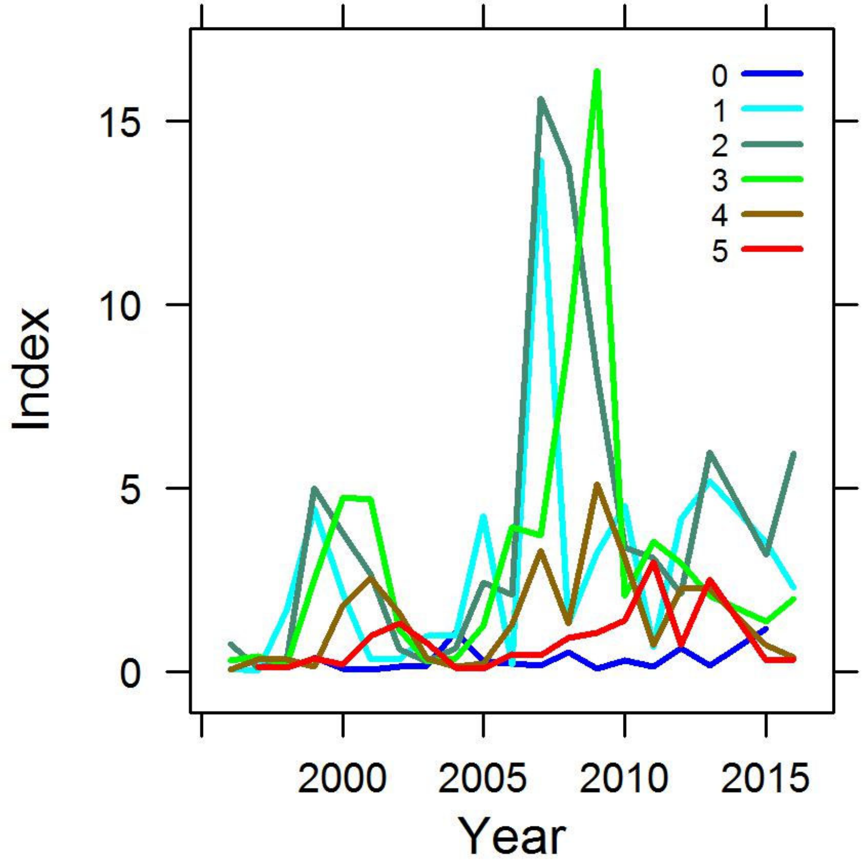
A Simple Cohort Strength Model

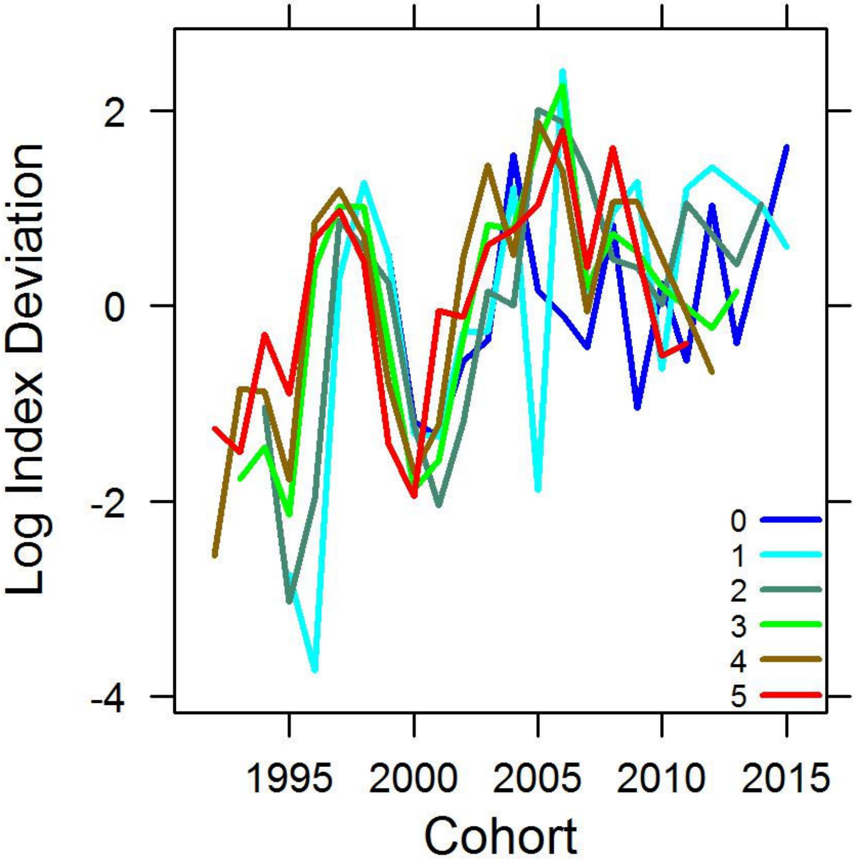
- The scale of $\mathrm{q}_{\mathrm{a}}^{*}$ and n_{0c} are still confounded
- i.e. $\mathrm{q}_{\mathrm{a}}^{**} = \mathrm{q}_{\mathrm{a}}^{*} - 10$ and $n_{oc}^{**} = n_{0c} + 10$ give the same model fit as $\mathrm{q}_{\mathrm{a}}^{*}$ and n_{0c}
- To fix this problem we constrain $\mathrm{q}_{\mathrm{a}}^{*} = 0$ for one reference age (usually R does the oldest age, A)
- in which case the scale of n_{0c} is the same as the scale of $r_{a,c,\ldots}$ for the reference age
- Hence a cohort-strength model estimated using a survey index only gives estimates of population <u>trends</u>
- Which is really useful information anyway!

Case Study: 3NO cod

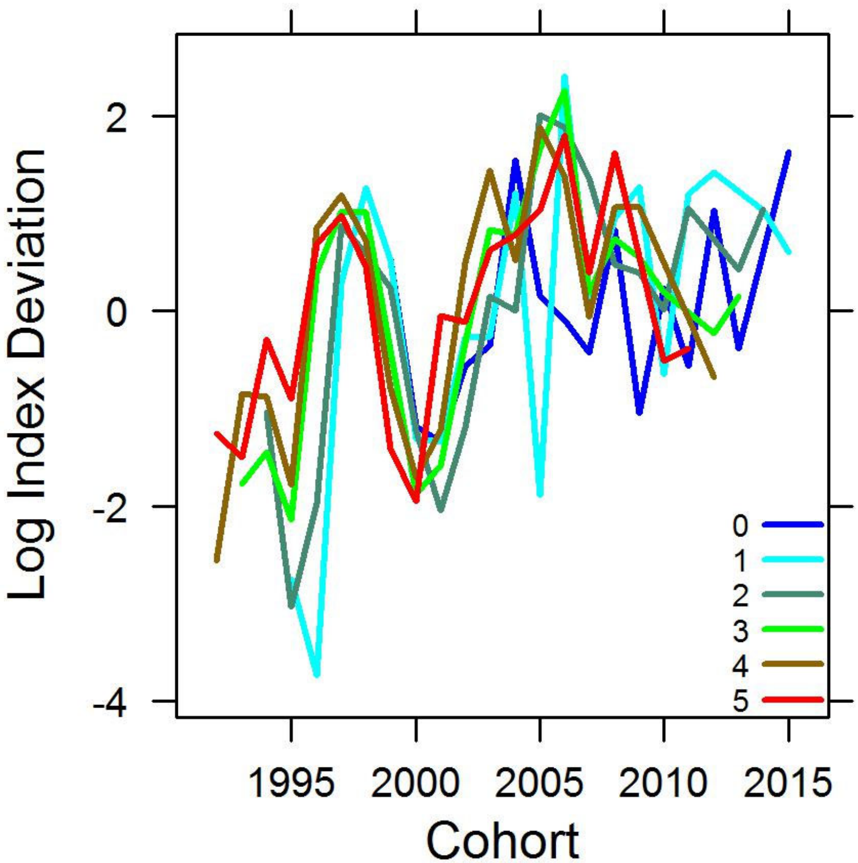
- Data: Fall survey indices of abundance, ages 0-5

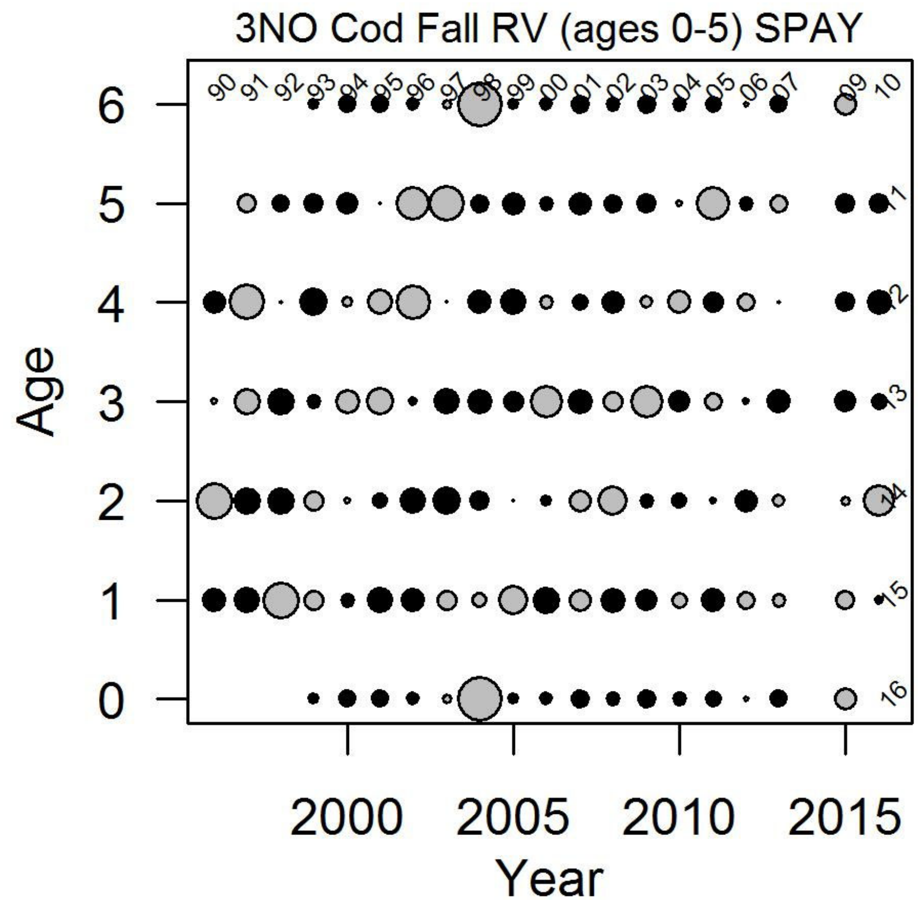






Case Study: 3NO cod





3NO cod Cohort Strength GLM

```
> sdat = read.table(file='F3LNO.dat',header=TRUE)
> head(sdat)
   YC Year Age index  log.index
1 1992 1996   4  0.06 -2.8134107
2 1992 1997   5  0.14 -1.9661129
3 1993 1996   3  0.29 -1.2378744
4 1993 1997   4  0.33 -1.1086626
5 1993 1998   5  0.11 -2.2072749
6 1994 1996   2  0.74 -0.3011051
```

step1

step2

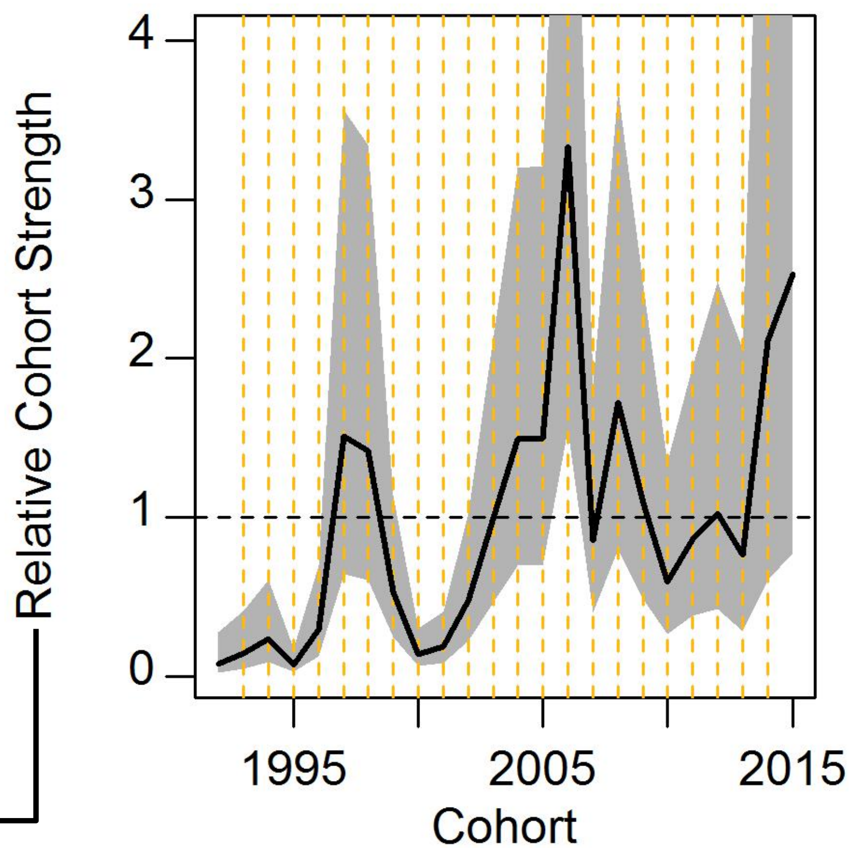
```
sdat$fYC = factor(sdat$YC)
sdat$fAge = factor(sdat$Age)
glmfit = glm(log.index ~ fYC + fAge -1,data=sdat)

sdat$pred = predict(glmfit)
sdat$resid = resid(glmfit)
sdat$sresid = rstandard(glmfit)

ttext="Cohort Strength, Fall Survey 3LNO cod"
stargazer(glmfit, type = "html",  out="out.doc",
align = TRUE, title=ttext, single.row=TRUE)
```

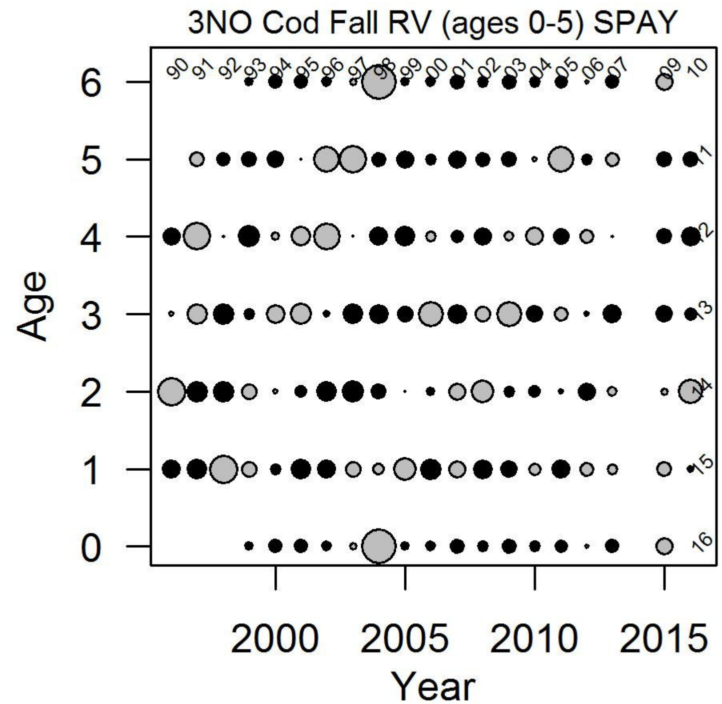
results

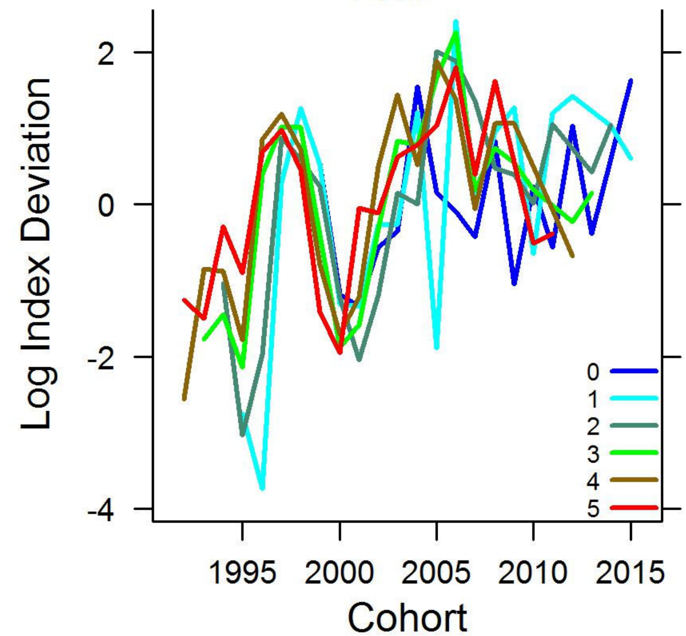
Cohort Strength, Fall Survey 3LNO cod

	Dependent variable:
	log.index
fYC1992	-3.835*** (0.639)
fYC1993	-3.259*** (0.536)
fYC1994	-2.747*** (0.477)
fYC1995	-3.899*** (0.436)
fYC1996	-2.530*** (0.436)
fYC1997	-0.903** (0.436)
fYC1998	-0.966** (0.436)
fYC1999	-1.946*** (0.389)
fYC2000	-3.268*** (0.389)
fYC2001	-2.985*** (0.389)
fYC2002	-2.053*** (0.389)
fYC2003	-1.318*** (0.389)
fYC2004	-0.917** (0.389)
fYC2005	-0.914** (0.389)
fYC2006	-0.115 (0.389)
fYC2007	-1.473*** (0.389)
fYC2008	-0.775** (0.389)
fYC2009	-1.229*** (0.414)
fYC2010	-1.833*** (0.415)
fYC2011	-1.463*** (0.415)
fYC2012	-1.291*** (0.451)
fYC2013	-1.587*** (0.506)
fYC2014	-0.570 (0.636)
fYC2015	-0.390 (0.601)
fAge1	1.771*** (0.284)
fAge2	2.410*** (0.287)
fAge3	2.334*** (0.286)
fAge4	1.639*** (0.289)
fAge5	1.250*** (0.294)
Observations	115
Log Likelihood	-125.485
Akaike Inf. Crit.	308.971
Note:	*p**p***p<0.01

3NO cod Cohort Strength

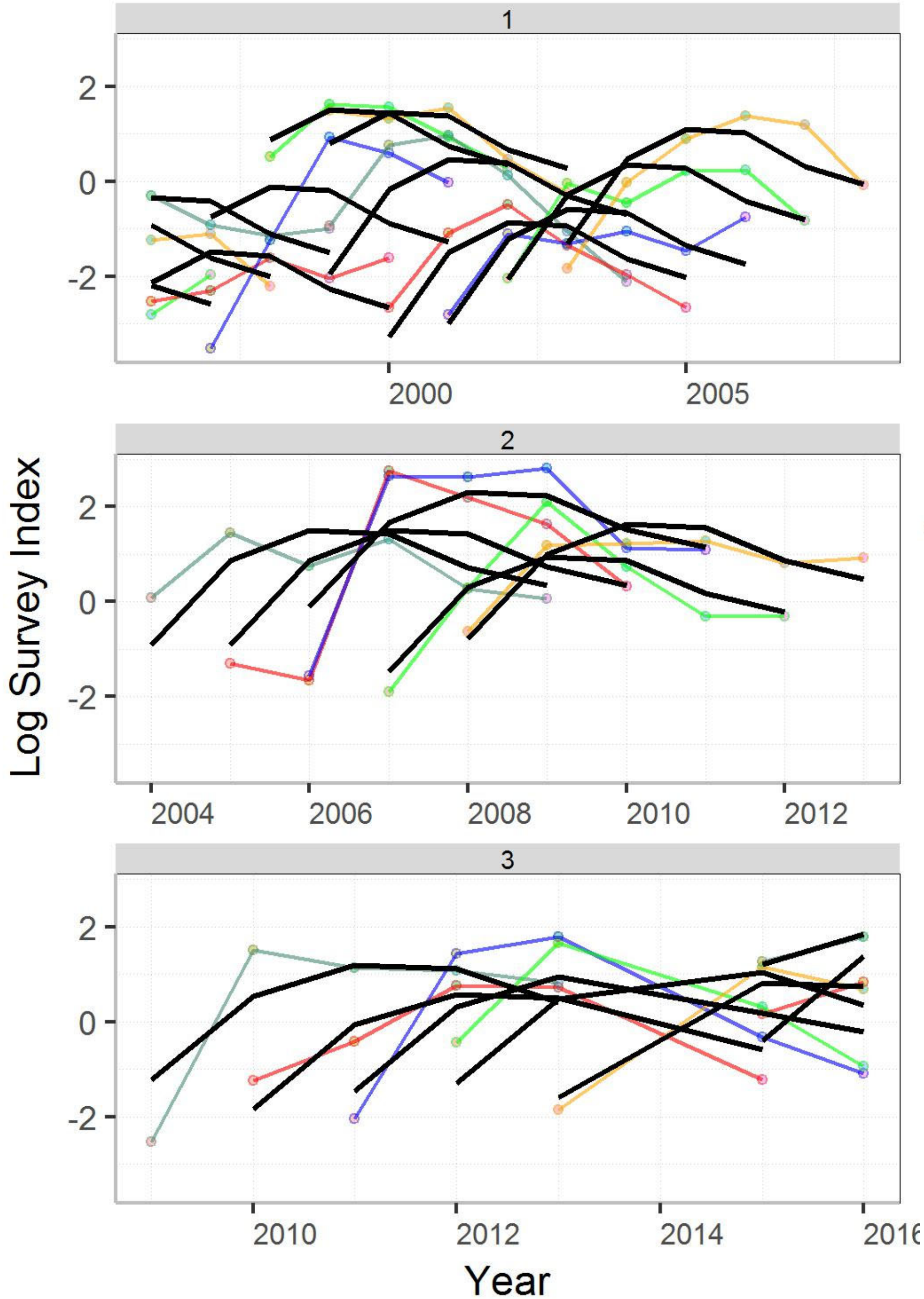


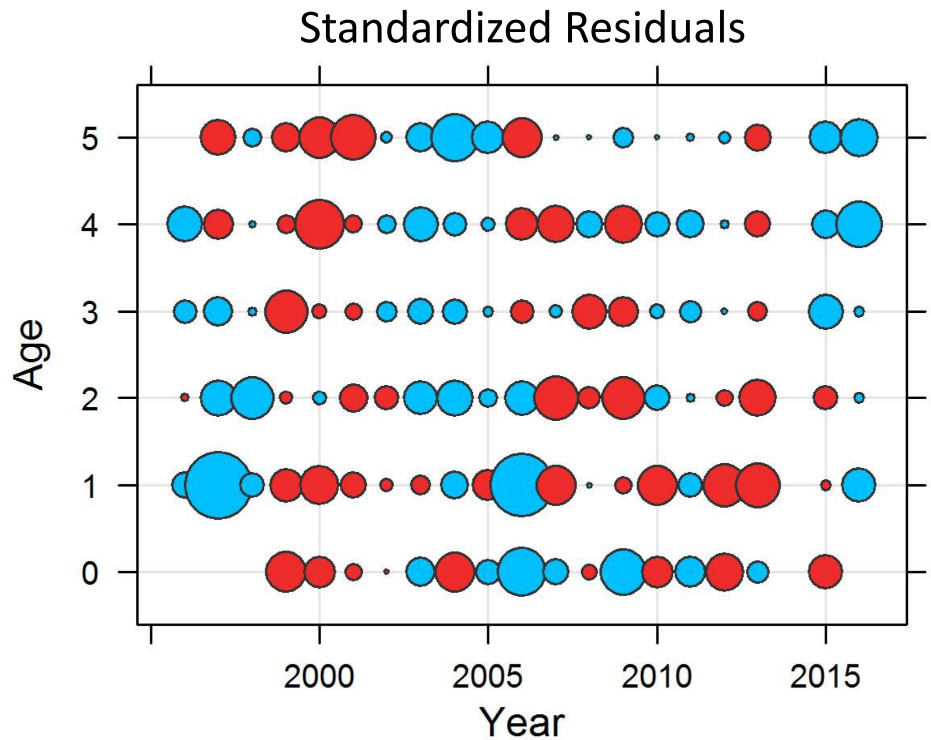




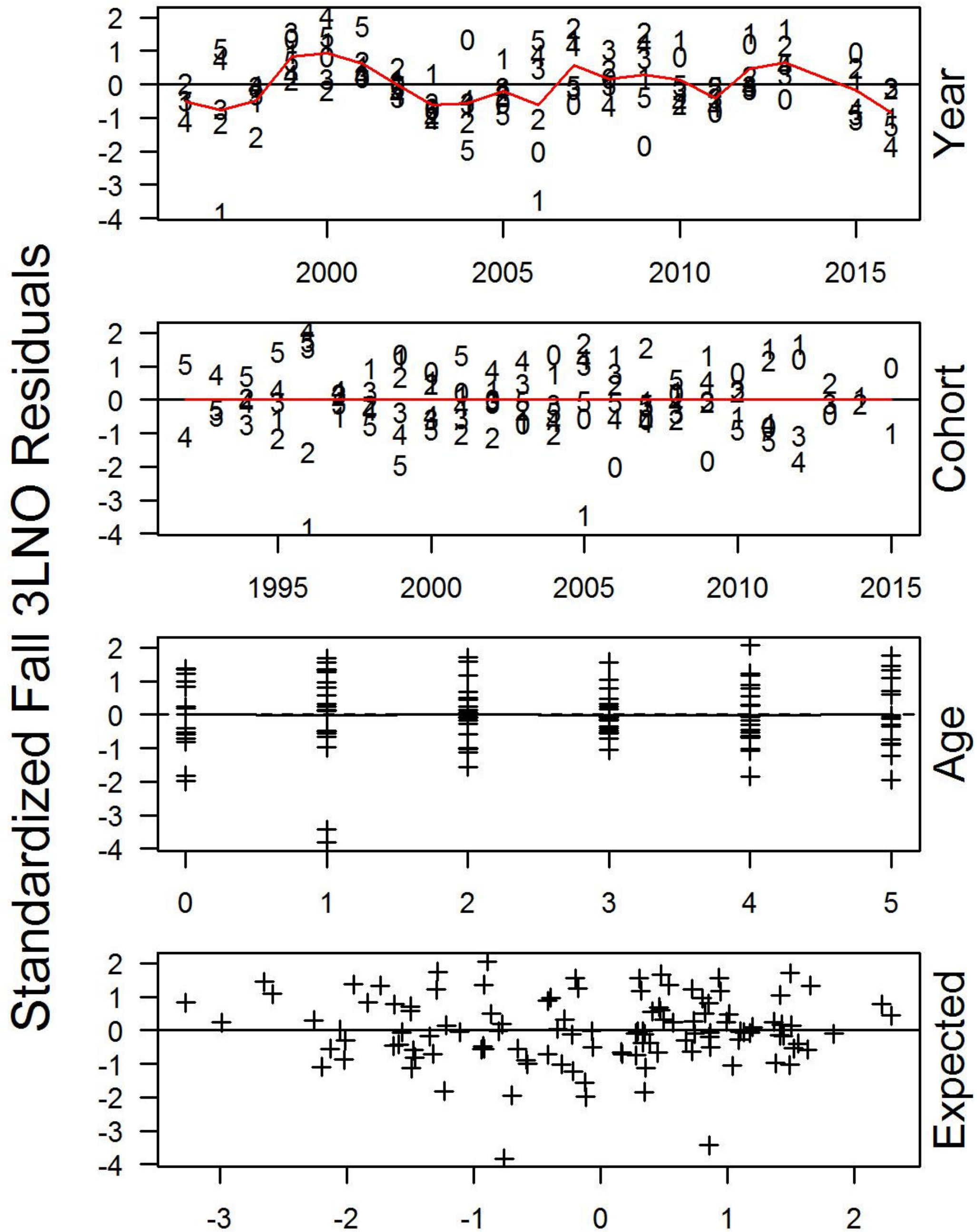

```
out = data.frame(log_est=coef(glmfit),ci=confint(glmfit))
rec.dev = subset(out,substr(rownames(out),1,3)=='fYC')
rnames = rownames(rec.dev)
rec.dev$YC = as.numeric(substring(rnames,4,7))
rec.dev$est = exp(rec.dev$log_est)
mean.est = mean(rec.dev$est)
rec.dev$est = rec.dev$est/mean.est
rec.dev$L95 = exp(rec.dev$ci.2.5..)/mean.est
rec.dev$U95 = exp(rec.dev$ci.97.5..)/mean.est
```

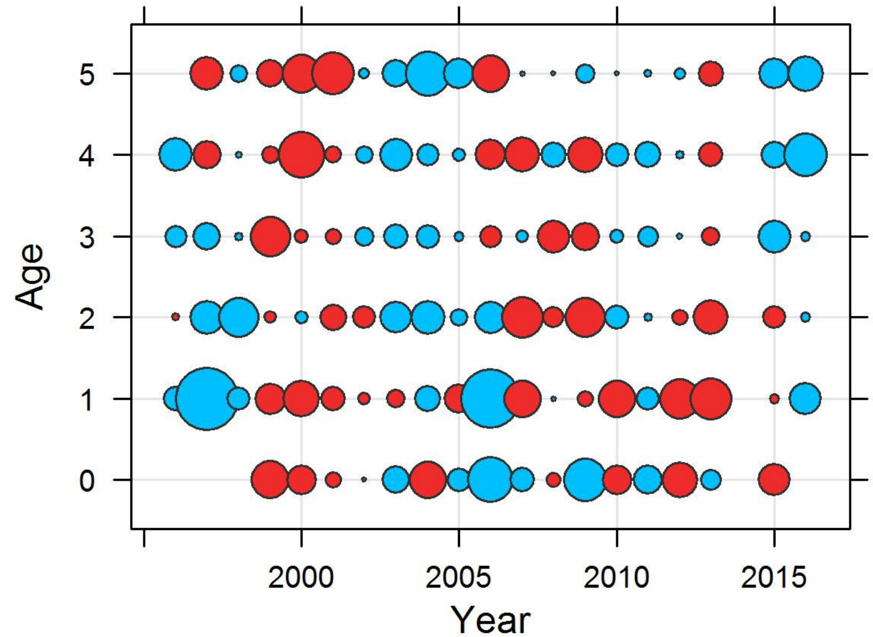
3NO cod GLM Fit





3NO cod GLM Residuals





Cohort Strength Model with Random year Effects

- Random effects are “parameters” that have a statistical distribution
- Very useful for more complete modelling of the stochastics of the data and system
- Will be taught more in F6005
- Mixed-effects Cohort Strength Model

$$r_{ac} = \mathrm{q}_{\mathrm{a}}^{*} + n_{0c} + \varepsilon_{ay} + \tau_y, \qquad \tau_y \sim N(0, \sigma_\tau^2),\ \varepsilon_{ay} \sim N(0, \sigma_\varepsilon^2),$$

- τ_y are common to all ages in a year
- Estimate this using lmer(), to be described in F6005

3NO cod Cohort Strength GLMM

```
sdat$fYear = factor(sdat$Year)
mixfit <- lmer(log.index ~ fYC + fAge -1 + (1|fYear),REML = FALSE, data=sdat)

sdat$mix.pred = predict(mixfit)
sdat$mix.resid = residuals(mixfit,type="response")
sdat$mix.sresid = residuals(mixfit,type="pearson")
```

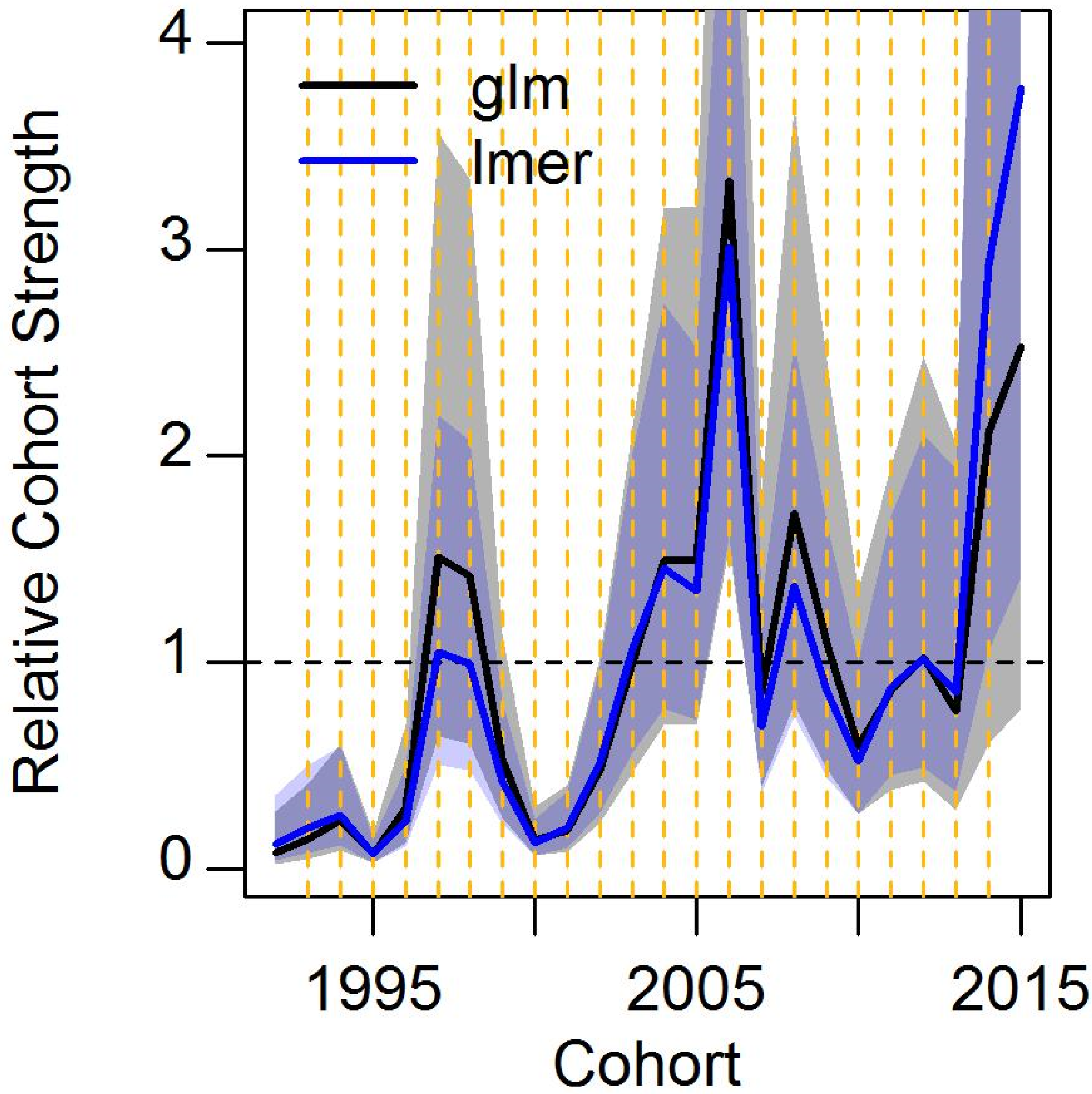
Cohort Strength, Fall Survey 3LNO

	Dependent variable: log.index
fYC1992	-3.487*** (0.551)
fYC1993	-2.991*** (0.467)
fYC1994	-2.725*** (0.417)
fYC1995	-4.039*** (0.382)
fYC1996	-2.841*** (0.378)
fYC1997	-1.337*** (0.376)
fYC1998	-1.391*** (0.374)
fYC1999	-2.247*** (0.335)
fYC2000	-3.453*** (0.332)
fYC2001	-3.001*** (0.329)
fYC2002	-2.044*** (0.326)
fYC2003	-1.320*** (0.323)
fYC2004	-1.009*** (0.321)
fYC2005	-1.087*** (0.320)
fYC2006	-0.286 (0.319)
fYC2007	-1.757*** (0.318)
fYC2008	-1.069*** (0.317)
fYC2009	-1.533*** (0.338)
fYC2010	-2.037*** (0.337)
fYC2011	-1.516*** (0.338)
fYC2012	-1.366*** (0.369)
fYC2013	-1.543*** (0.417)
fYC2014	-0.313 (0.519)
fYC2015	-0.056 (0.502)
fAge1	1.915*** (0.212)
fAge2	2.563*** (0.216)
fAge3	2.487*** (0.218)
fAge4	1.770*** (0.224)
fAge5	1.358*** (0.232)
Observations	115
Log Likelihood	-119.162
Akaike Inf. Crit.	300.324
Bayesian Inf. Crit.	385.417
Note:	*p **p ***p<0.01

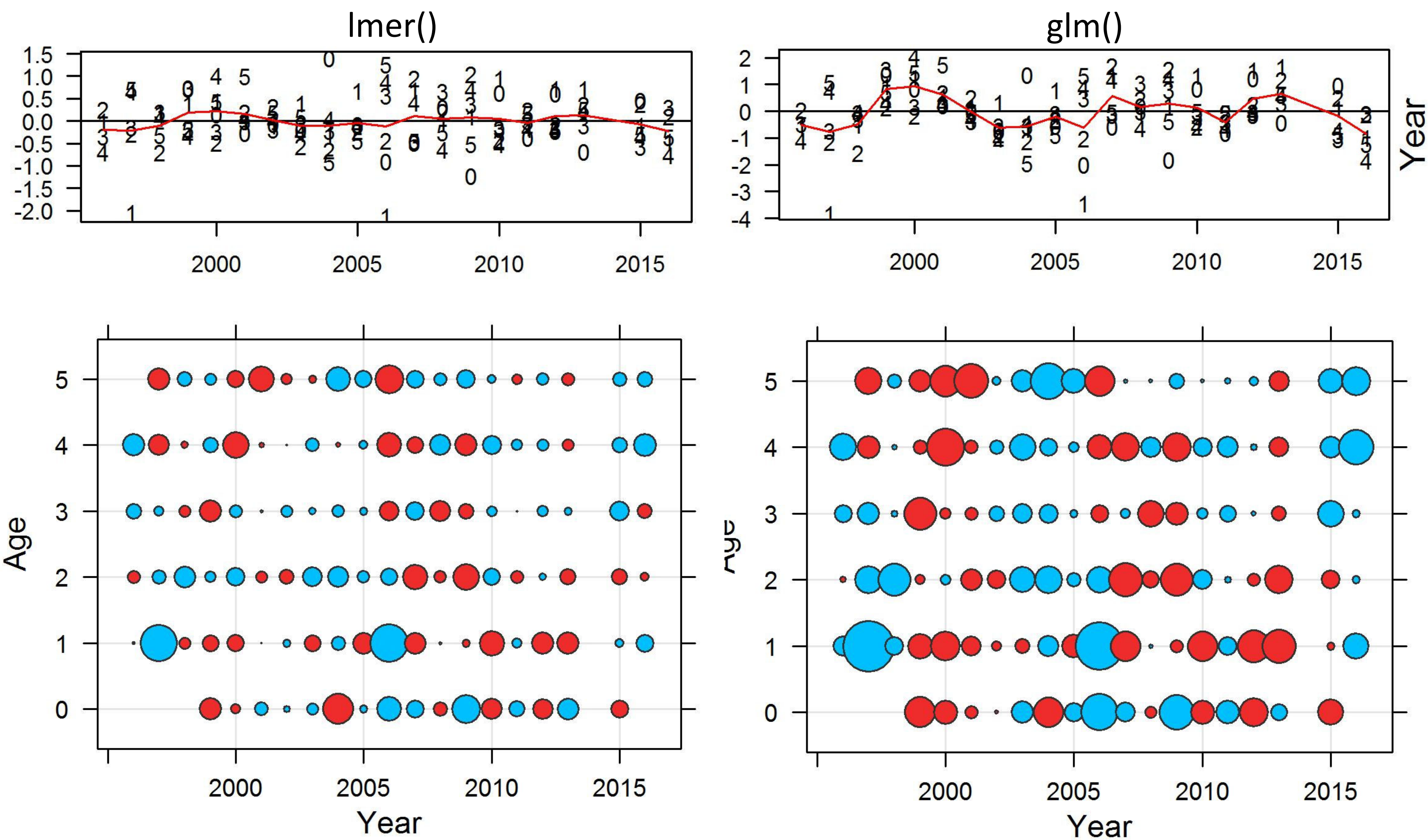
Some R code for relative cohort strength CI's

```
ci=confint(mixfit,parm=names(fixef(mixfit)),method = "Wald")
out = data.frame(log_est=fixef(mixfit),ci=ci)
mix.rec.dev = subset(out,substr(rownames(out),1,3)=='fYC')
rnames = rownames(mix.rec.dev)
mix.rec.dev$YC = as.numeric(substring(rnames,4,7))
mix.rec.dev$est = exp(mix.rec.dev$log_est)
mean.est = mean(mix.rec.dev$est)
mix.rec.dev$est = mix.rec.dev$est/mean.est
mix.rec.dev$L95 = exp(mix.rec.dev$ci.2.5..)/mean.est
mix.rec.dev$U95 = exp(mix.rec.dev$ci.97.5..)/mean.est
```

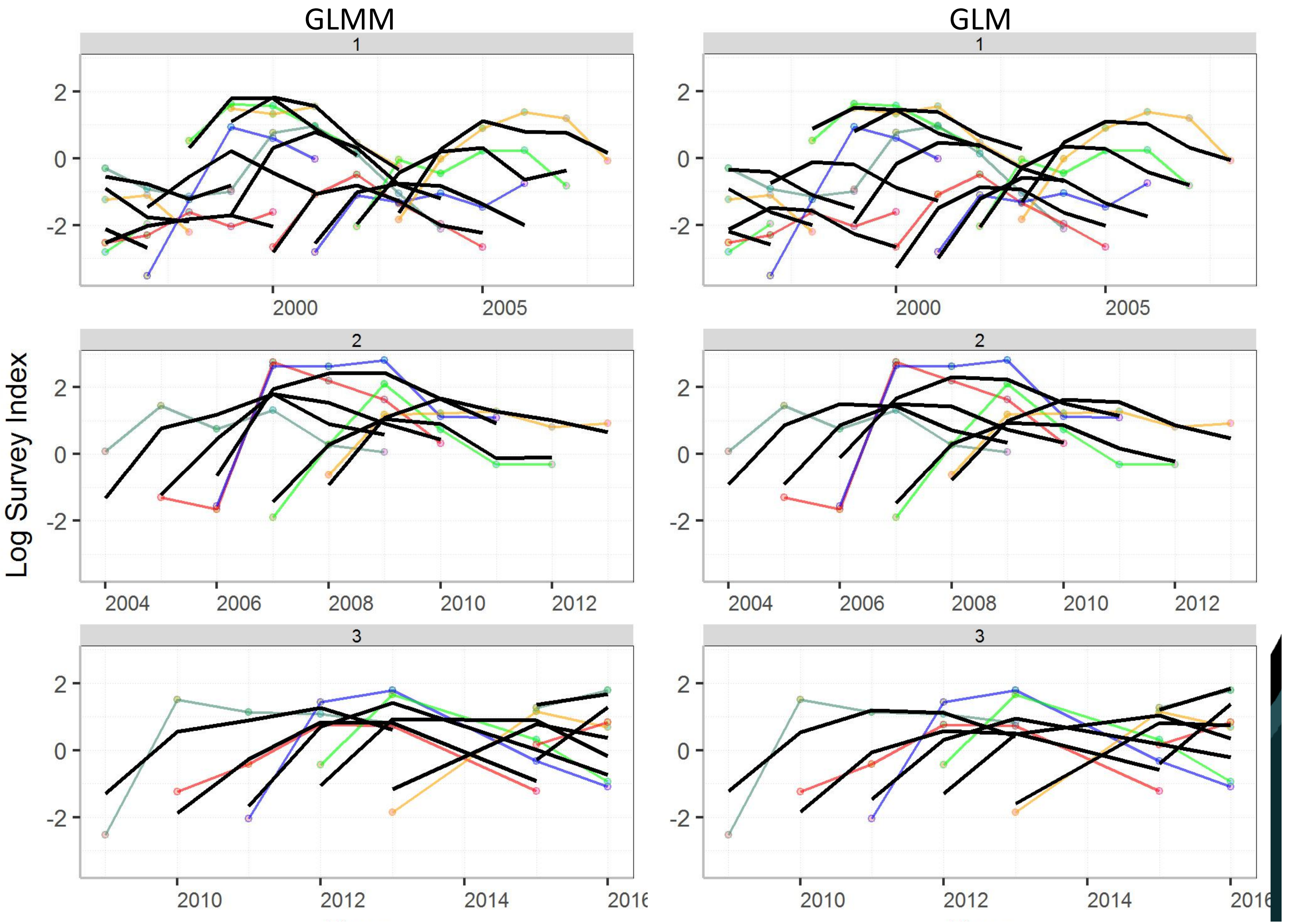
3NO cod GLMM Cohort Strength



3NO cod Residuals



3NO cod GLM Fit



Cohort Strength Model with Random year Effects

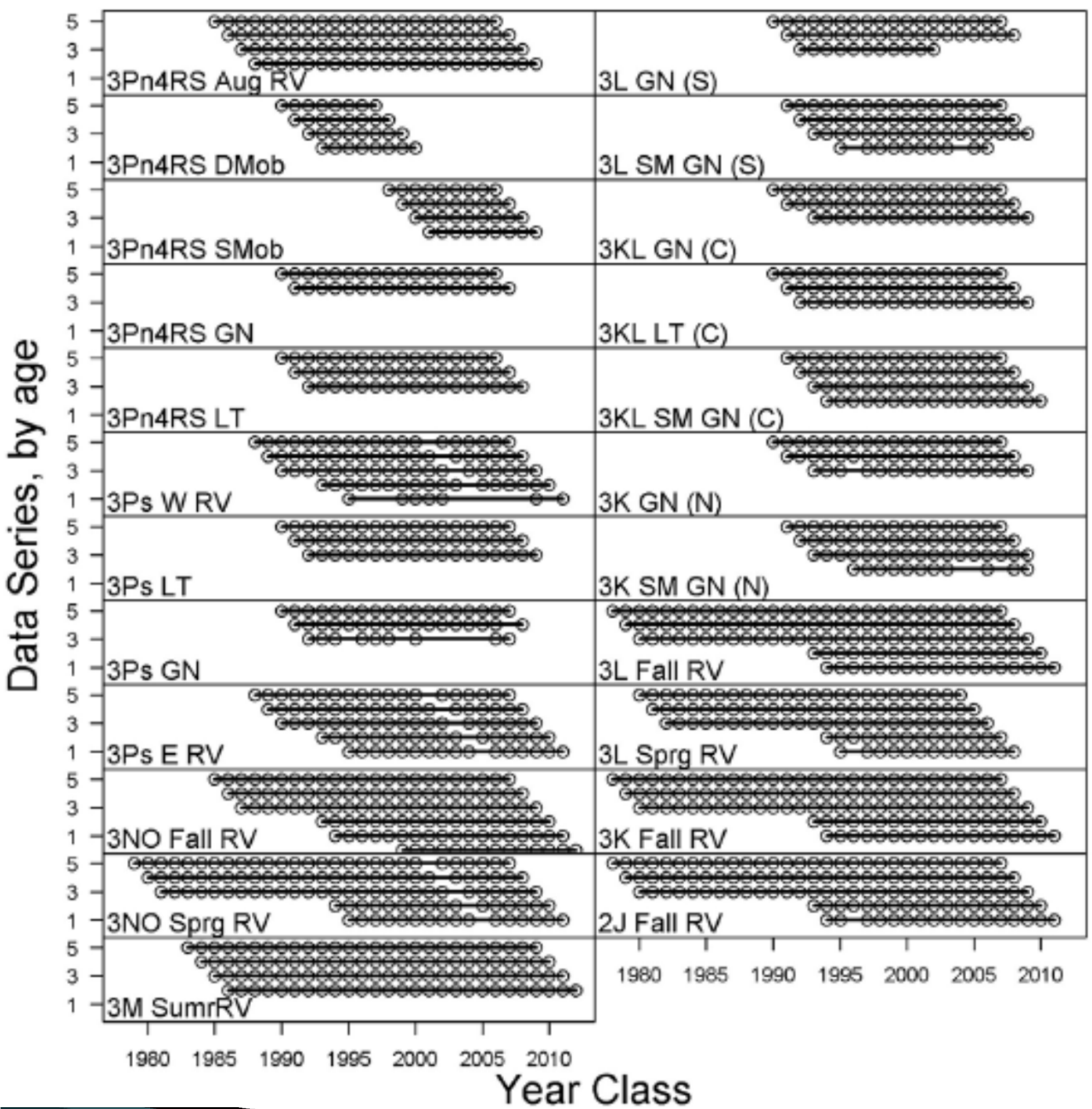
- The random year effects model seemed somewhat improved
- But there still seems to be some cohort patterns in residuals
- This may be related to changes in juvenile mortality rates that we have assumed to be constant
- Or could be changes in growth rates and concomitant changes in q (via length selectivity)
- We are currently working on this …..

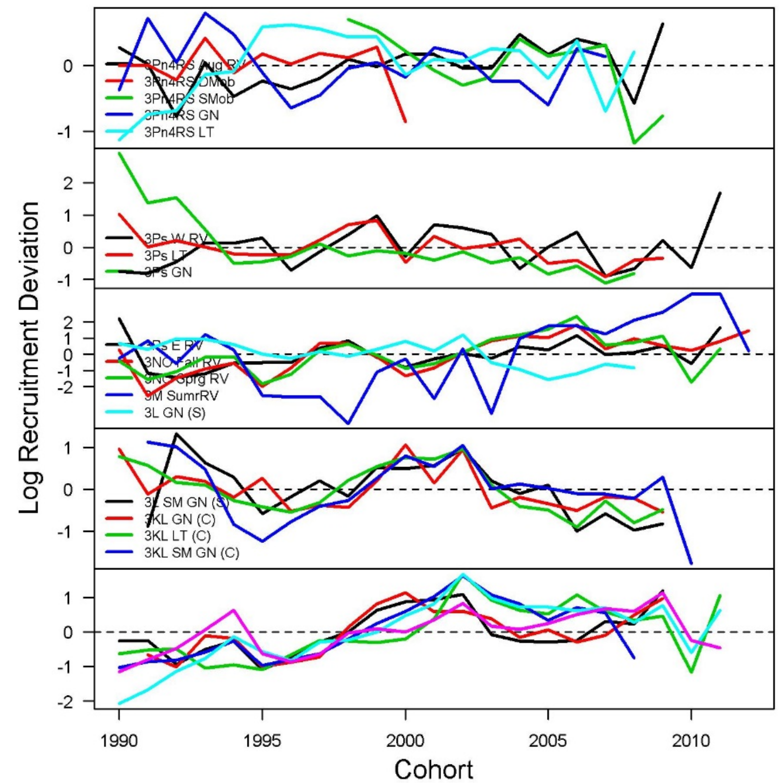
Zhang, F., Rideout, R.M. and Cadigan, N.G., 2020. Spatiotemporal variations in juvenile mortality and cohort strength of Atlantic cod (*Gadus morhua*) off Newfoundland and Labrador. *Canadian Journal of Fisheries and Aquatic Sciences*, *77*(3), pp.625-635.

Cohort Strength Model Application

Tulk, F.J., Cadigan, N.G., Brattey, J. and Robert, D., 2017. Spatial synchronicity in recruitment of Atlantic cod (Gadus morhua) stocks off Newfoundland and Labrador and the Flemish Cap. *Fisheries Research*, *191*, pp.49-59.

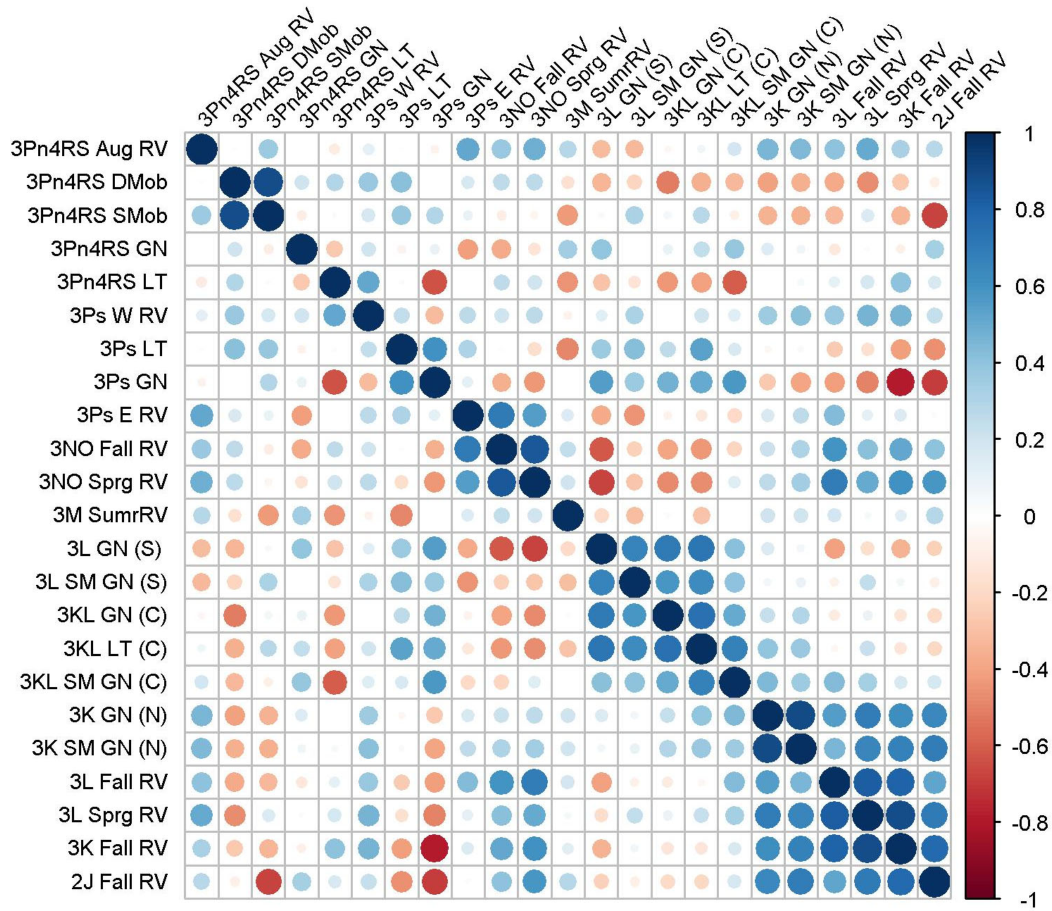
$$I_{say} = q_{sa} R_{sy} \exp(\varepsilon_{say}),$$



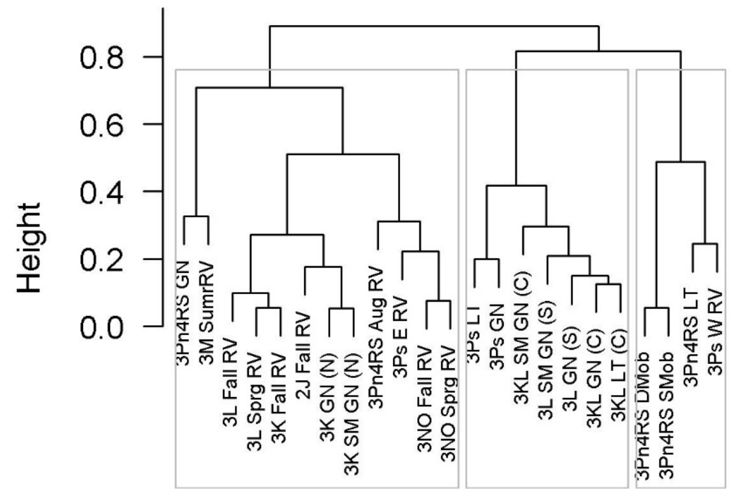


Estimates of log recruitment deviations for each survey. Each panel contains a unique survey grouping based on geographical location. The dashed lines indicate the horizontal reference line at zero, with line color indicating a specific survey.

Tulk et al Application



Pairwise correlations log recruitment estimates. The area and shading of the circles are proportional to the absolute value of the correlation. The blue and red scales define positive and negative correlations, respectively.



Cluster dendrogram based on the complete agglomeration method.

Extension: SURBA

- Add more ages and more information
- SURBA: Survey-based cohort model

Beare, D.J., Needle, C.L., Burns, F. and Reid, D.G., 2005. Using survey data independently from commercial data in stock assessment: an example using haddock in ICES Division VIa. *ICES Journal of Marine Science*, *62*(5), pp.996-1005.

Mesnil, B., Cotter, A.J.R., Fryer, R.J., Needle, C.L. and Trenkel, V.M. 2009. A review of fishery independent assessment models, and initial evaluation based on simulated data, Aquat. Living Resour. 22: 207–216.

Cook, R.M. 1997. Stock trends in six North Sea stocks as revealed by an analysis of research vessel surveys. ICES J. Mar. Sci. 54: 924–933.

Cook, R.M., 2013. A fish stock assessment model using survey data when estimates of catch are unreliable. *Fisheries research*, *143*, pp.1-11.

SURBA

- A problem with adding more ages is that older fish are exploited by the fishery
- And fishing mortality rates (F) are not constant
- (maybe M is not constant either, but if $F >> M$ then assuming M is constant may not be a bad thing)
- Many stock assessments (SA) assume $M = 0.2$
- And infer F using the ‘fundamental equation of fisheries science’, $F = Z – 0.2$

SURBA

- SURBA provides estimates of *Z*, and we can estimate *F* with some assumption about *M*
- We can provide some catch advice, without any catch statistics
- Cook (2013) outlines how this can be done
- For now we will focus on implementing a SURBA-type model in R

SURBA

- Abundance is modeled as cohort recruitment (N_0) times cumulative total mortality:

$$N_{a+1,y+1} = N_{ay}\exp(-Z_{ay})$$

$$\downarrow$$

$$N_{a+1,y+1} = N_{0,y-a}\exp(-Z_{ay}-Z_{a-1,y-1}-, \dots, -Z_{0,y-a})$$

$$N_{a+1,y+1} = N_{0,y-a}\exp\left(-\sum_{i=0}^{a} Z_{a-i,y-i}\right)$$

- Estimation is based on a time (y) series of age (a) based survey indices, R_{ay}.

- Model survey index prediction

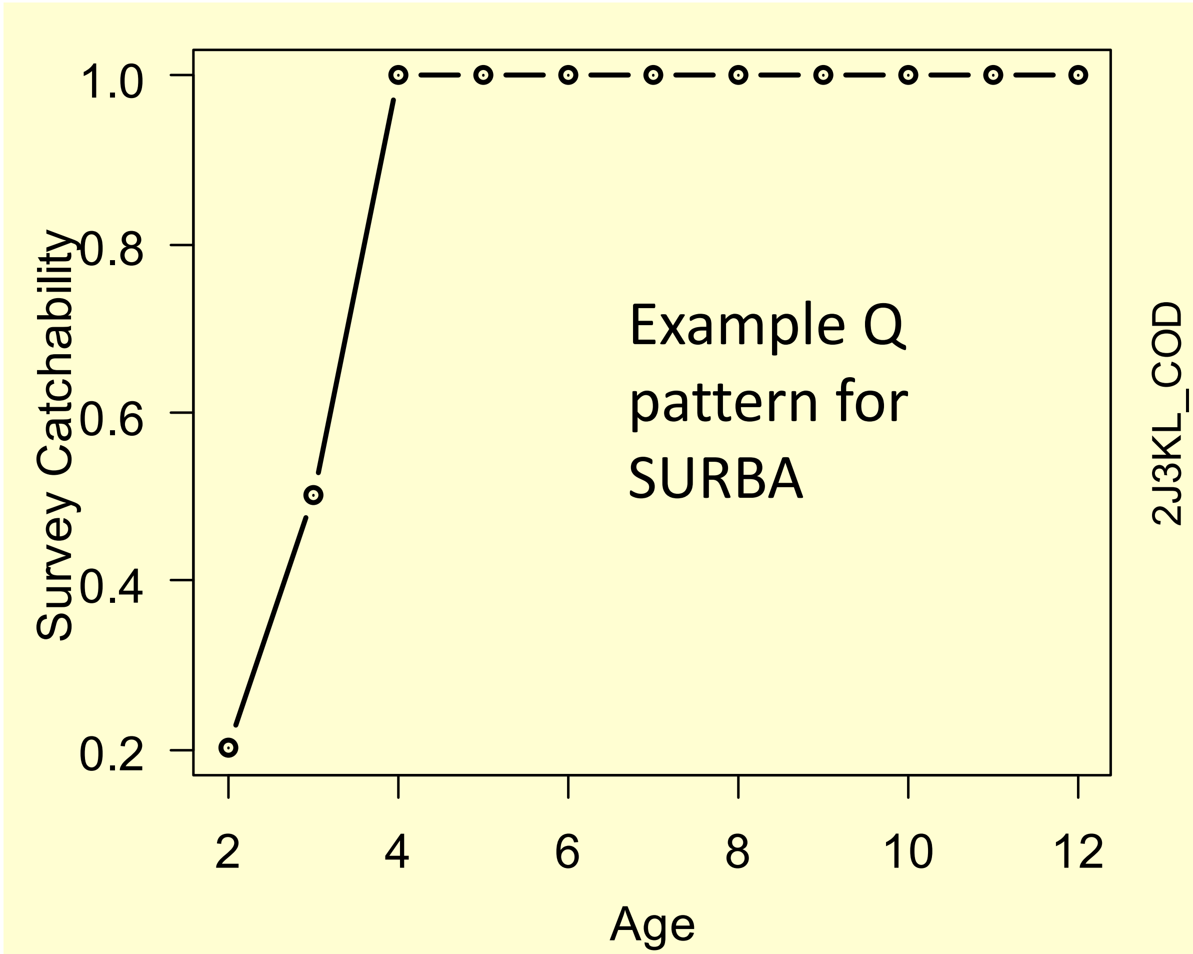
- $E(R_{ay}) = Q_a N_{ay}\exp(-f_y Z_{ay})$

Fraction of year which survey occurs in (could change over time)

SURBA Relativity

- Catchabilities (Q_a's) and Z's are highly confounded.
- The Z information comes from the slope of the survey cohort-catch curve, which is confounded overall with the Q pattern
- Need to specify (i.e. assume) the Q age pattern.
- SURBA estimates of stock size are relative to these assumptions.
- Fully recruited Q is fixed at $Q=1$, so that the scale of SURBA estimates is the same as the index.

SURBA Relativity



Separable Mortality

- A separable model for total mortality, $Z_{a,y} = s_a f_y$, is used in the standard ICES SURBA model.
- The age (s_a) and year (f_y) effects are not fully identifiable
- We constrain $s_a = 1$ at some age we think the fishery fully selects
  - E.g. age 6 for 3Ps cod

Separable Mortality

- We will use the separable model, but there are other possibilities

NL ad hoc study group on methods for stock assessment

- SGAM_NL – TMB SURBA

- 2J3KL cod: A correlated (over ages) random walk (over years) was used to model *Z*'s.

  - Details in Cadigan (2013); DFO Can. Sci. Advis. Sec. Res. Doc. 2013/054

  - An integrated state-space model now used

SURBA added value

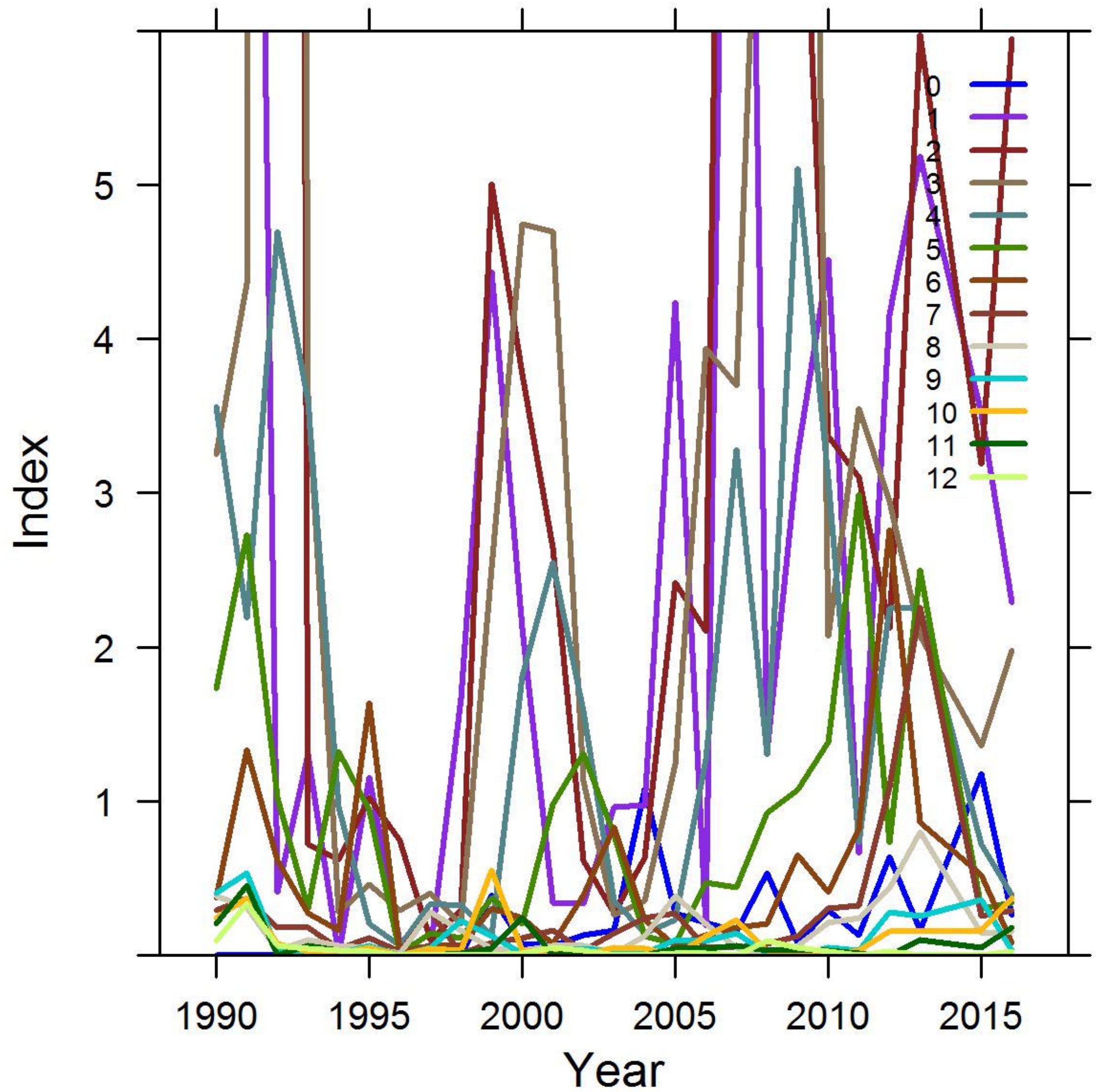
- With SURBA we get:

  1. **estimates of relative cohort strength**
  2. **estimates of total mortality rates (Z)**

- But we need to make an assumption about the pattern in Q_a to estimate Z's
- Q assumptions: we may have a good understanding of the Q age-pattern for older ages in a survey
- We did not assume an age-pattern in Q_a in the cohort-strength model which is why we could not estimate M at these young ages

3NO cod

Data: Fall survey indices of abundance, ages 0-12

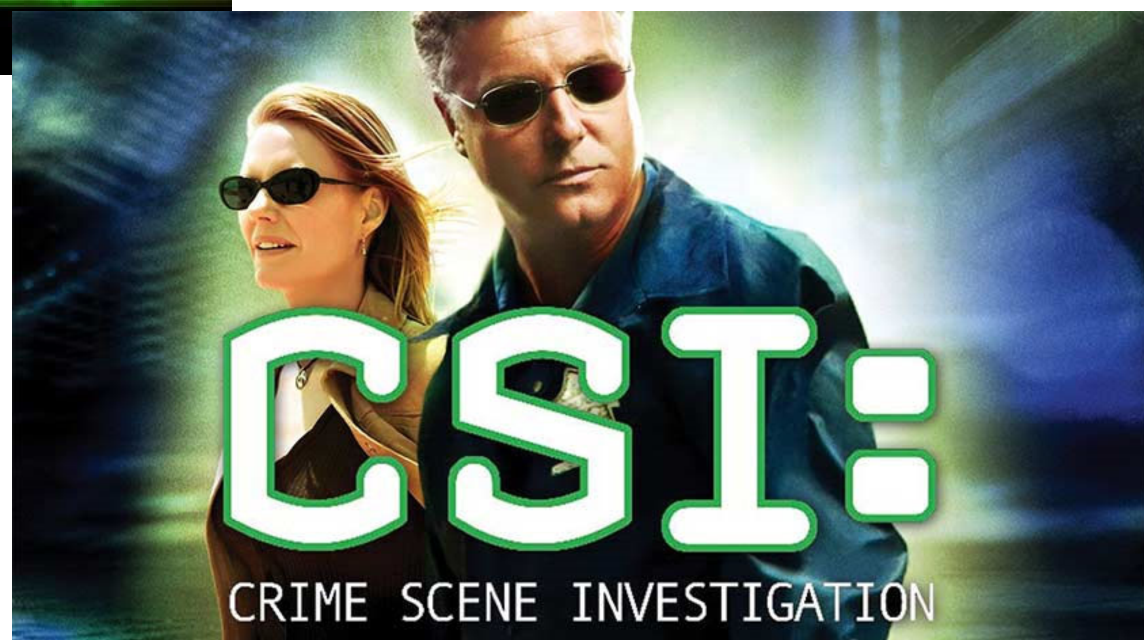


messy

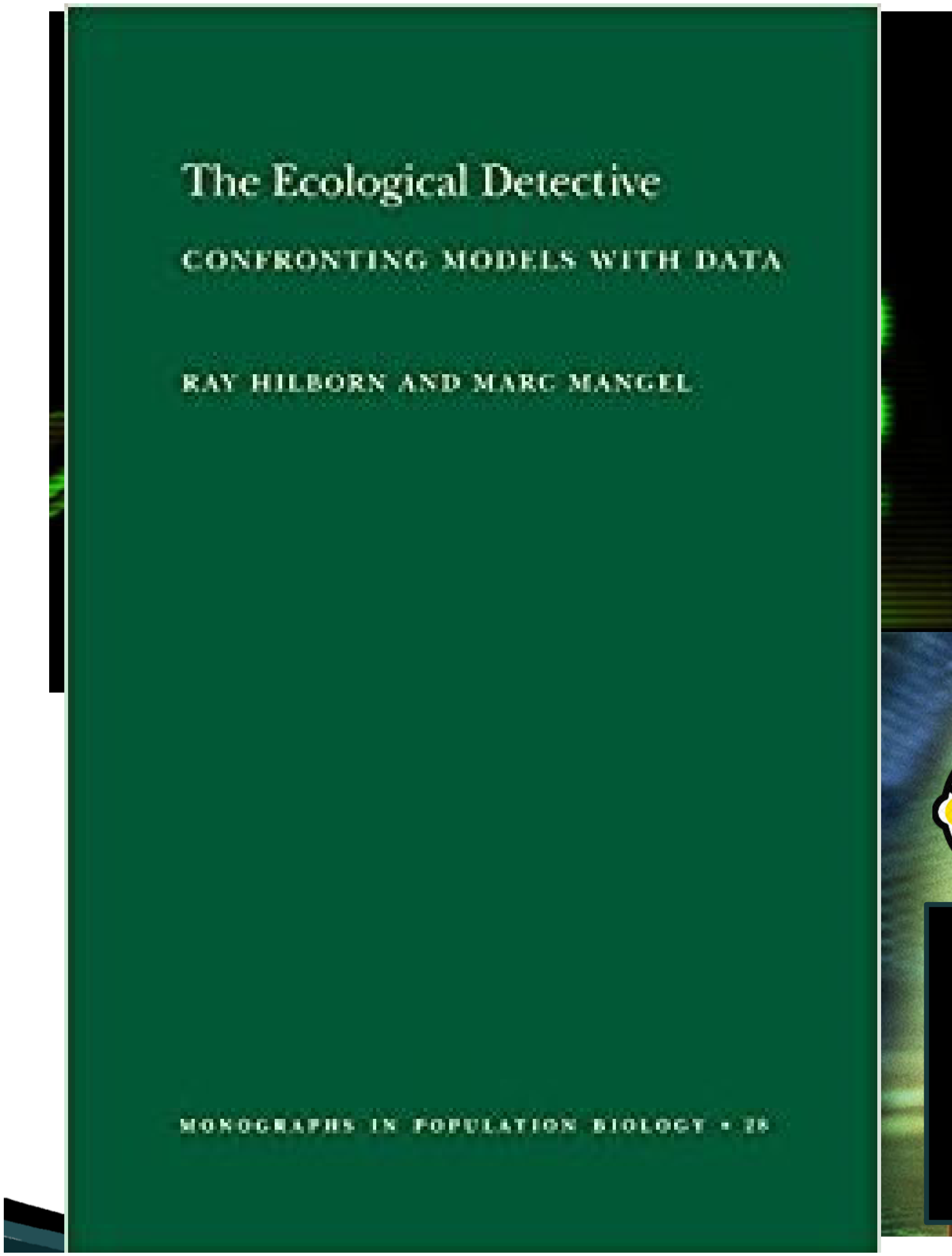
Stock Assessment Science

Our job is to figure out what happened to the stock based on the clues (i.e. data) available

Sometimes the data are limited and messy!



Stock Assessment Science



Our job is to figure out what happened to the stock based on the clues (i.e. data) available

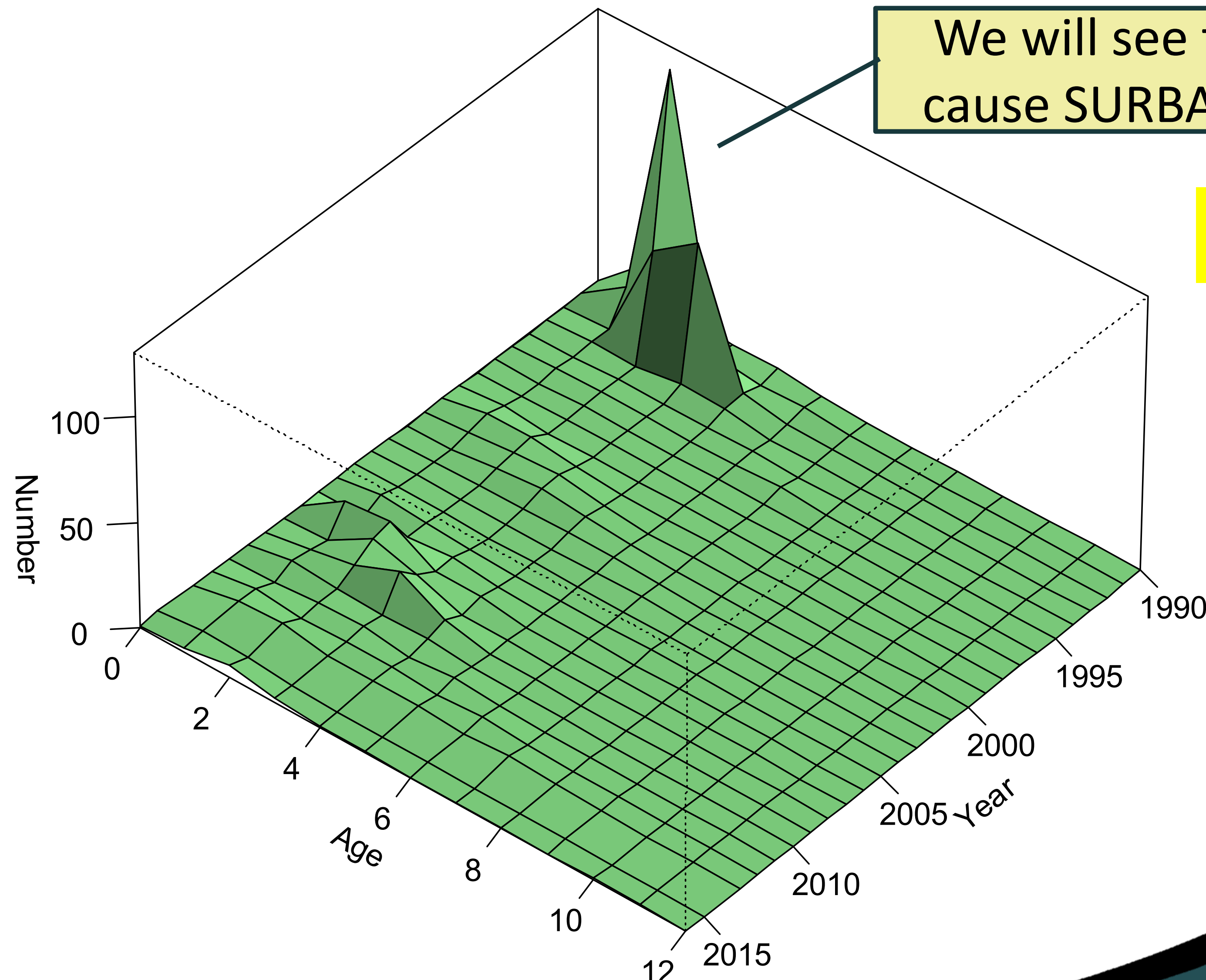
Sometimes the data are limited and messy!

3NO cod

Data: Fall survey indices of abundance, ages 0-12

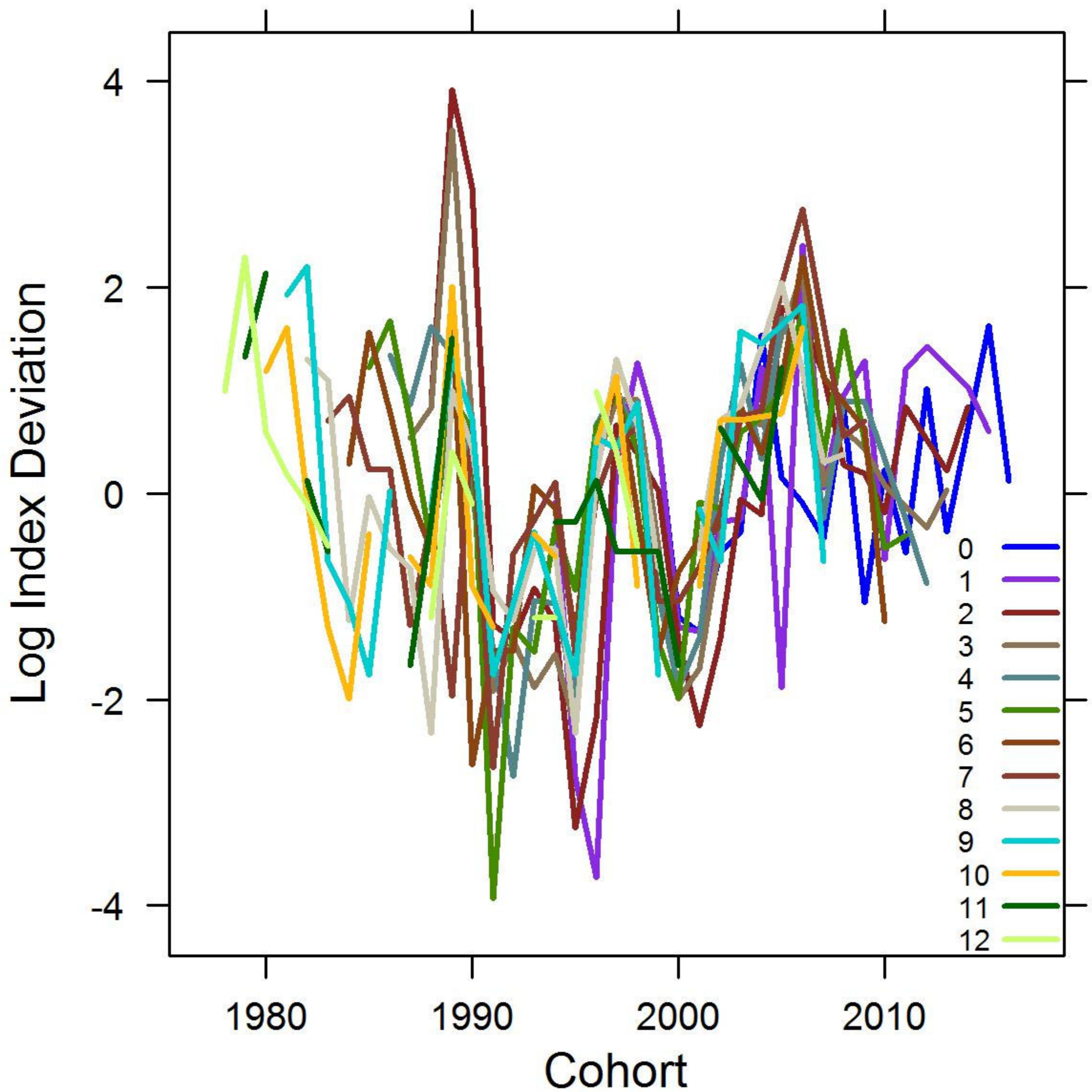
We will see that these cause SURBA problems

messy



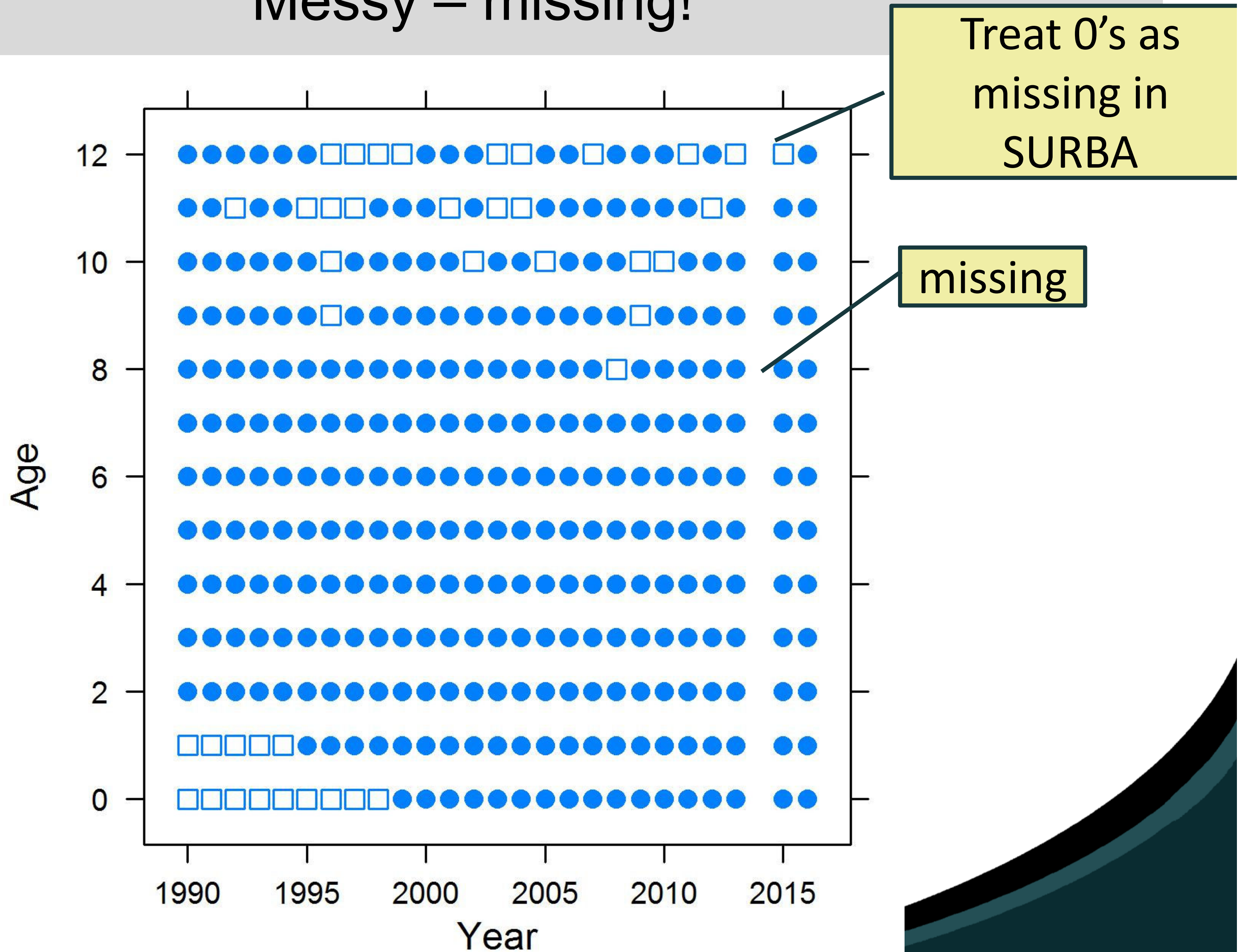
3NO cod

There is some signal! But what is it?

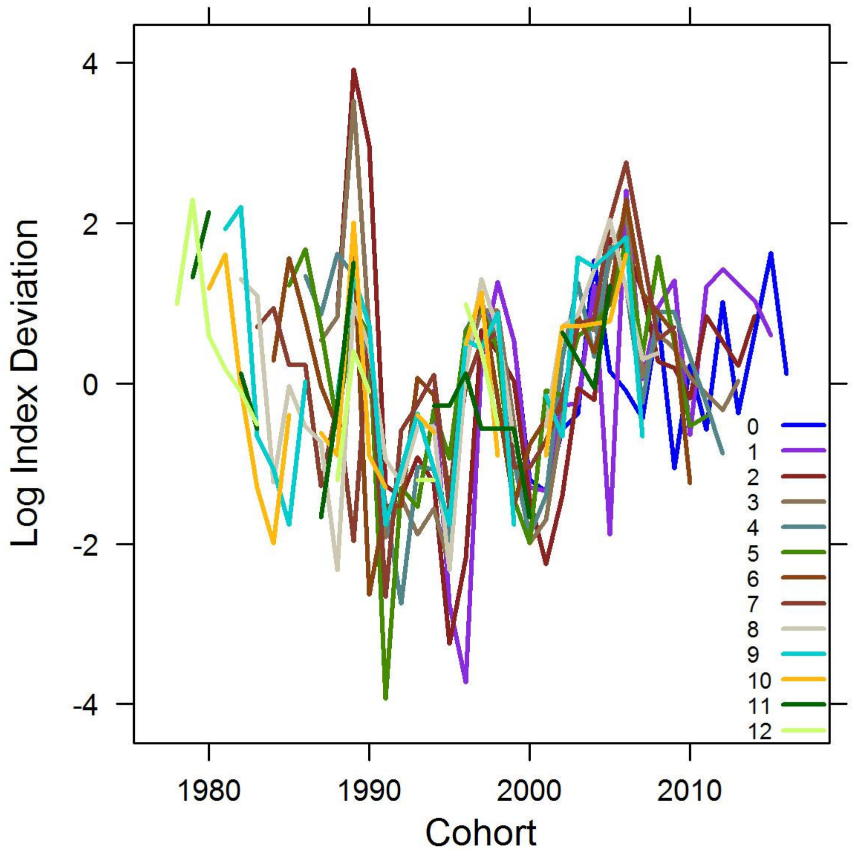


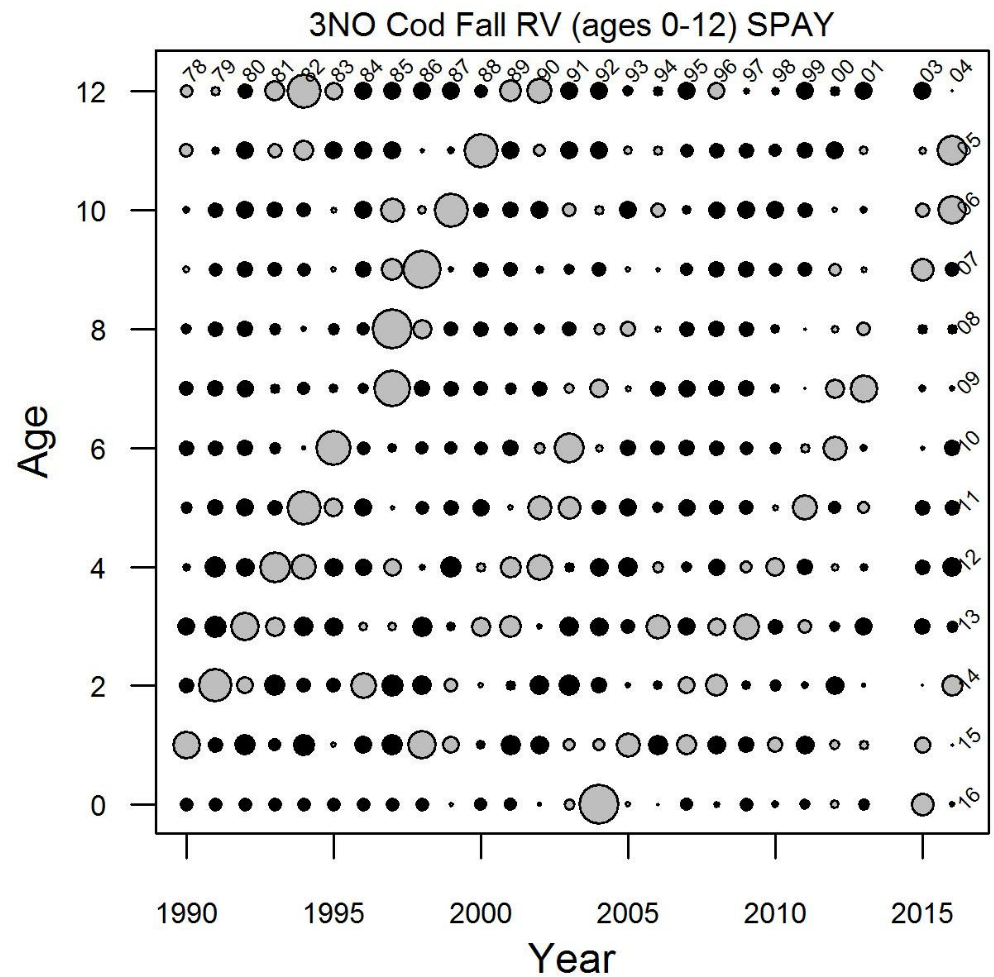
3NO cod

Messy – missing!



3NO cod SPAY


Log Index Deviation
Cohort
1980
1990
2000
2010
4
2
0
-2
-4
0
1
2
3
4
5
6
7
8
9
10
11
12


3NO Cod Fall RV (ages 0-12) SPAY
Age
Year
1990
1995
2000
2005
2010
2015
12
10
8
6
4
2
0

SURBA in R

```
> load('F3LNO.RData')
> head(vdat,n=10)
  Year Age index wt
1  1990   0  0.00  0
2  1991   0  0.00  0
3  1992   0  0.00  0
4  1993   0  0.00  0
5  1994   0  0.00  0
6  1995   0  0.00  0
7  1996   0  0.00  0
8  1997   0  0.00  0
9  1998   0  0.00  0
10 1999   0  0.39  1
```

```
vdat$YC = vdat$Year-vdat$Age
vdat$log.index = NA
ind = vdat$wt==1
vdat$log.index[ind] = log(vdat$index[ind])
```

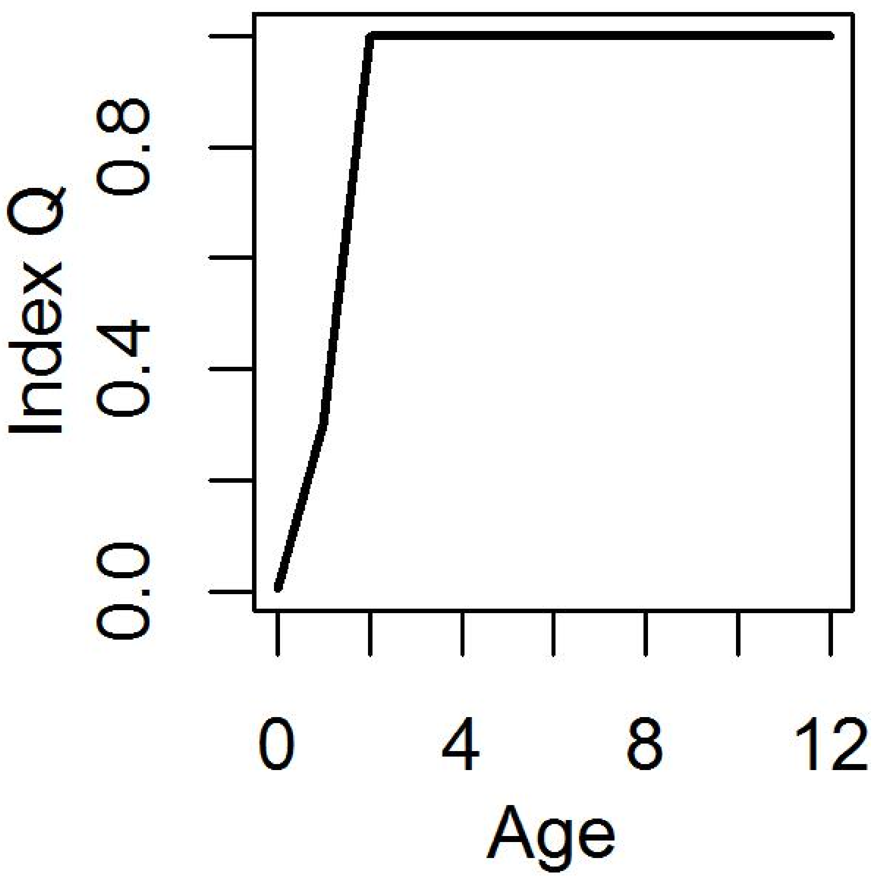
Make log index for index>0

```
> head(mdat,n=10)
  Year Age0  Age1   Age2  Age3 Age4 Age5 Age6 Age7 Age8 Age9 Age10 Age11 Age12
1  1990 0.00 18.89   6.15  3.25 3.56 1.73 0.37 0.29 0.38 0.40  0.24  0.20  0.09
2  1991 0.00 14.87 129.66  4.36 2.19 2.73 1.33 0.37 0.31 0.53  0.37  0.45  0.33
3  1992 0.00  0.41  49.65 65.00 4.70 1.02 0.61 0.18 0.03 0.03  0.07  0.00  0.06
4  1993 0.00  1.30   0.72  3.63 3.59 0.30 0.27 0.18 0.10 0.02  0.02  0.06  0.04
5  1994 0.00  0.00   0.62  0.28 0.96 1.32 0.16 0.04 0.06 0.01  0.01  0.03  0.03
6  1995 0.00  1.15   1.02  0.46 0.20 0.94 1.64 0.11 0.05 0.06  0.05  0.00  0.02
7  1996 0.00  0.08   0.74  0.29 0.06 0.01 0.02 0.02 0.01 0.00  0.00  0.00  0.00
8  1997 0.00  0.03   0.10  0.40 0.33 0.14 0.06 0.28 0.28 0.05  0.04  0.00  0.00
9  1998 0.00  1.67   0.29  0.20 0.32 0.11 0.06 0.01 0.16 0.22  0.03  0.01  0.00
10 1999 0.39  4.44   5.01  2.52 0.13 0.37 0.30 0.08 0.04 0.12  0.55  0.04  0.00
```

SURBA in R

```
## create some useful objects

uage = unique(vdat$Age)
uyear = min(vdat$Year):max(vdat$Year)
A = length(uage)
Y = length(uyear)
## vectorized full model dimensions to create an index observation map
pdat = data.frame(
  Age = as.vector(matrix(uage,nrow=Y,ncol=A,byrow=T)),
  Year = as.vector(matrix(uyear,nrow=Y,ncol=A,byrow=F))
)
pdat$imap = 1:nrow(pdat)

vdat.nz = subset(vdat,wt==1); #use data with wts=1
temp1 = paste(pdat$Year,"_",pdat$Age,sep='')
temp2 = paste(vdat.nz$Year,"_",vdat.nz$Age,sep='')
vdat.nz$imap = pdat$imap[temp1 %in% temp2]
```

SURBA Q's



- Q age-pattern is a subjective choice
- A stock assessment meeting will have to decide if these make sense
- The 3LNO survey trawl is a campelen shrimp-type and whole stock available to survey so one expects Q pattern like this
- Maybe we will see something in residuals

SURBA in R

```
Npop = function(Nfirst,No,Z){
  Nmatrix = matrix(NA,nrow=Y,ncol=A)
  next.age = 2:A
  Nmatrix[,1] = No
  Nmatrix[1,next.age] = Nfirst
  for(y in 2:Y){
    Nmatrix[y,next.age] = Nmatrix[y-1,next.age-1]*exp(-Z[y-1,next.age-1]);
  }
  return(Nmatrix)
}
```

	Age				
Year	1	2	3	4	5
1	N_{11}	N_{21}	N_{31}	N_{41}	N_{51}
2	N_{12}	N_{22}	N_{32}	N_{42}	N_{52}
3	N_{13}	N_{23}	N_{33}	N_{43}	N_{53}
4	N_{14}	N_{24}	N_{34}	N_{44}	N_{54}
5	N_{15}	N_{25}	N_{35}	N_{45}	N_{55}
6	N_{16}	N_{26}	N_{36}	N_{46}	N_{56}
7	N_{17}	N_{27}	N_{37}	N_{47}	N_{57}
8	N_{18}	N_{28}	N_{38}	N_{48}	N_{58}

SURBA in R

```
Npop = function(Nfirst,No,Z){
 Nmatrix = matrix(NA,nrow=Y,ncol=A)
 next.age = 2:A
 Nmatrix[,1] = No
 Nmatrix[1,next.age] = Nfirst
 for(y in 2:Y){
  Nmatrix[y,next.age] = Nmatrix[y-1,next.age-1]*exp(-Z[y-1,next.age-1]);
 }
 return(Nmatrix)
}
```

	Age				
Year	1	2	3	4	5
1	N_{11}	N_{21}	N_{31}	N_{41}	N_{51}
2	N_{12}	N_{22}	N_{32}	N_{42}	N_{52}
3	N_{13}	N_{23}	N_{33}	N_{43}	N_{53}
4	N_{14}	N_{24}	N_{34}	N_{44}	N_{54}
5	N_{15}	N_{25}	N_{35}	N_{45}	N_{55}
6	N_{16}	N_{26}	N_{36}	N_{46}	N_{56}
7	N_{17}	N_{27}	N_{37}	N_{47}	N_{57}
8	N_{18}	N_{28}	N_{38}	N_{48}	N_{58}

SURBA in R

```
pred_index = function(logNo,logNfirst,sparm,logf){
  No = exp(logNo)
  Nfirst=exp(logNfirst)
  f = exp(logf)
  age_diff = uage-6
  s = exp(sparm[1]*age_diff + sparm[2]*age_diff*age_diff)
  Z = f %o% s
  N = Npop(Nfirst,No,Z)
  logR = log(Qm*N*exp(-0.75*Z))
  vN = as.vector(N)
  vlogR = as.vector(logR)
  return(cbind(vlogR,vN))
}
```

nls() will use this

```
q_age = c(0.005,0.3,1,1,1,1,1,1,1,1,1,1,1)
Qm = matrix(q_age,nrow=Y,ncol=A,byrow=T)
```

SURBA nls() starting values

```
N.start = mdat[,2:14]/Qm
Nfirst = apply(N.start,2,mean)[2:A]
No = rep(mean(N.start[,1]),Y)
sparm=c(0,0)
age_diff = uage-6
s_age = exp(sparm[1]*age_diff + sparm[2]*age_diff*age_diff)
f_year = rep(0.4,Y)
Z = f_year %o% s_age
Ninit = Npop(Nfirst,No,Z)
```

nls() didn't converge with these

```
vec_pred = function(x){
 logNo = x[1:27]; logNfirst = x[28:39];
 sparm = x[40:41];  logf = x[42:length(x)];
 return(pred_index(logNo,logNfirst,sparm,logf)[vdat.nz$imap,1])
}

fit = function(x){return(sum((vdat.nz$log.index- vec_pred (x))**2))}

x.start=optim(x,fit,control=list(trace=1))$par
```

Use deriv-free optim to get better starting values

SURBA nls() fit

```
start.parms = list(
  logNo = x.start[1:27],
  logNfirst = x.start[28:39],
  sparm = x.start[40:41],
  logf = x.start[42:length(x)]
)
```

```
> surba.fit <- nls(log.index ~
pred_index(logNo,logNfirst,sparm,logf)[imap,1],
+   algorithm="port",data=vdat.nz,start = start.parms,
+   control=list(maxiter=10000))
Error in nls(log.index ~ pred_index(logNo, logNfirst, sparm, logf)[imap,
:
  Convergence failure: singular convergence (7)
```

Some of the logf -> -∞

Let's lower bounds on logf's to log(0.2)
So that Z = s x f >= 0.2 at age 6 (i.e. s_6=1)

SURBA nls() starting values


```
lower = list(
  logNo = rep(-Inf,length(start.parms$logNo)),
  logNfirst = rep(-Inf,length(start.parms$logNfirst)),
  logs = rep(-Inf,length(start.parms$logs)),
  logf = rep(log(0.2),length(start.parms$logf))
)

upper = list(
  logNo = rep(Inf,length(start.parms$logNo)),
  logNfirst = rep(Inf,length(start.parms$logNfirst)),
  logs = rep(Inf,length(start.parms$logs)),
  logf = rep(Inf,length(start.parms$logf))
)
lower=unlist(lower)
upper=unlist(upper)
```

I like to set up bounds like this (first as a list) so that they match with starting values

nls() wants the bounds as vectors

```
start.parms$logf[start.parms$logf<log(0.2)]=log(0.3)
```

SURBA nls() bounded fit

```
> surba.fit <- nls(log.index ~
pred_index(logNo,logNfirst,logs,logf)[imap,1],
+   algorithm="port",data=vdat.nz,start = start.parms,
+   control=list(maxiter=10000),lower=lower,upper=upper)
> surba.fit
Nonlinear regression model
  model: log.index ~ pred_index(logNo, logNfirst, sparm, logf)[imap, 1]
   data: vdat.nz
    logNo1     logNo2     logNo3     logNo4     logNo5     logNo6     logNo7     logNo8     logNo9
   6.64630    6.37062    8.73874    7.26008    3.94689    2.65314    3.91006    3.01629    3.04270
   logNo10    logNo11    logNo12    logNo13    logNo14    logNo15    logNo16    logNo17    logNo18
   2.69835    2.75854    4.10542    4.41039    4.82569    3.80940    4.89969    4.82241    3.80238
   logNo19    logNo20    logNo21    logNo22    logNo23    logNo24    logNo25    logNo26    logNo27
   5.19941    4.00489    3.95639    4.15244    4.46553    5.29138    6.45667    8.49249    5.06012
 logNfirst1 logNfirst2 logNfirst3 logNfirst4 logNfirst5 logNfirst6 logNfirst7 logNfirst8 logNfirst9
   5.91075    2.75746    1.62672    1.28314    0.21419   -1.28368   -1.08231   -0.71089   -0.77610
logNfirst10 logNfirst11 logNfirst12    sparm1     sparm2      logf1      logf2      logf3      logf4
  -1.16160   -1.16529   -2.29019   -0.11077    0.01174   -1.60944   -1.60944    0.37703    0.49702
     logf5      logf6      logf7      logf8      logf9     logf10     logf11     logf12     logf13
  -0.64694   -0.53110   -0.12657   -1.60944   -1.60944   -1.60944   -1.60944   -0.39045   -0.54792
    logf14     logf15     logf16     logf17     logf18     logf19     logf20     logf21     logf22
  -0.29191   -1.60944   -0.49946   -1.59371   -1.60944   -0.37002   -1.57725   -0.61201   -0.91820
    logf23     logf24     logf25     logf26     logf27
  -1.60944   -0.81975   -1.60944    0.50684   -0.69617
 residual sum-of-squares: 128.3

Algorithm "port", convergence message: relative convergence (4)
```

Hit lower bounds

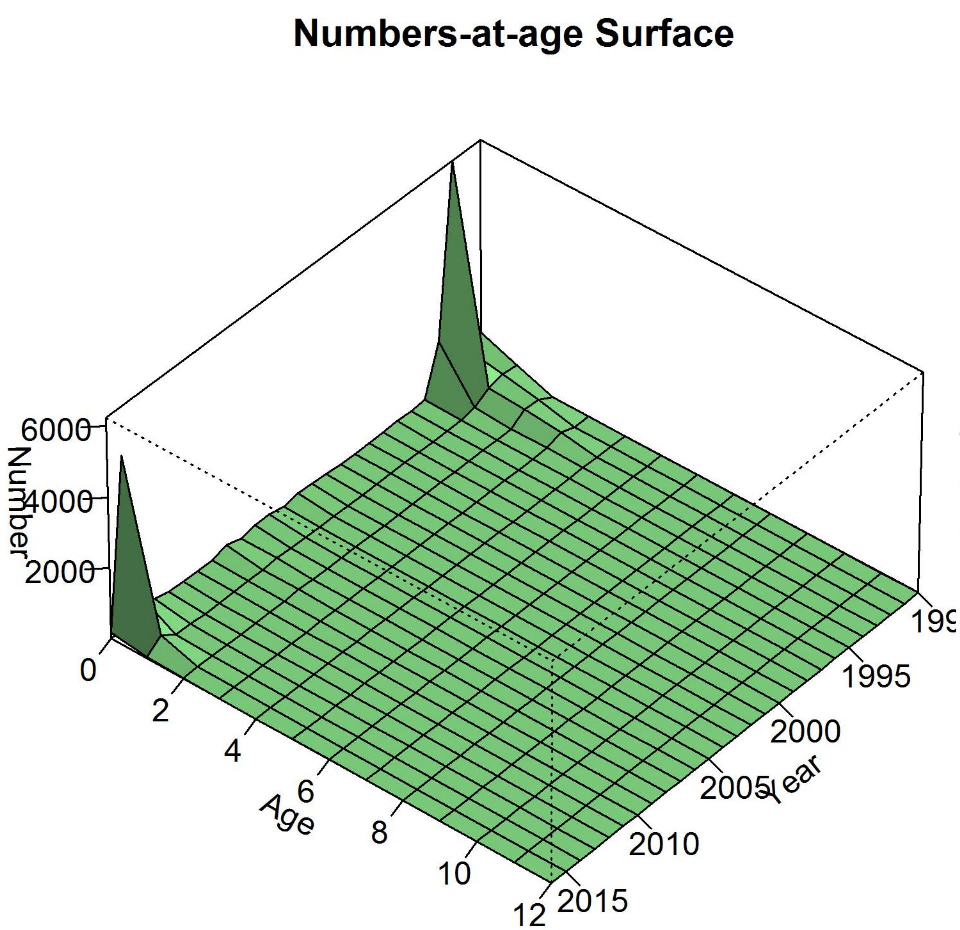
SURBA nls() results via stargazer

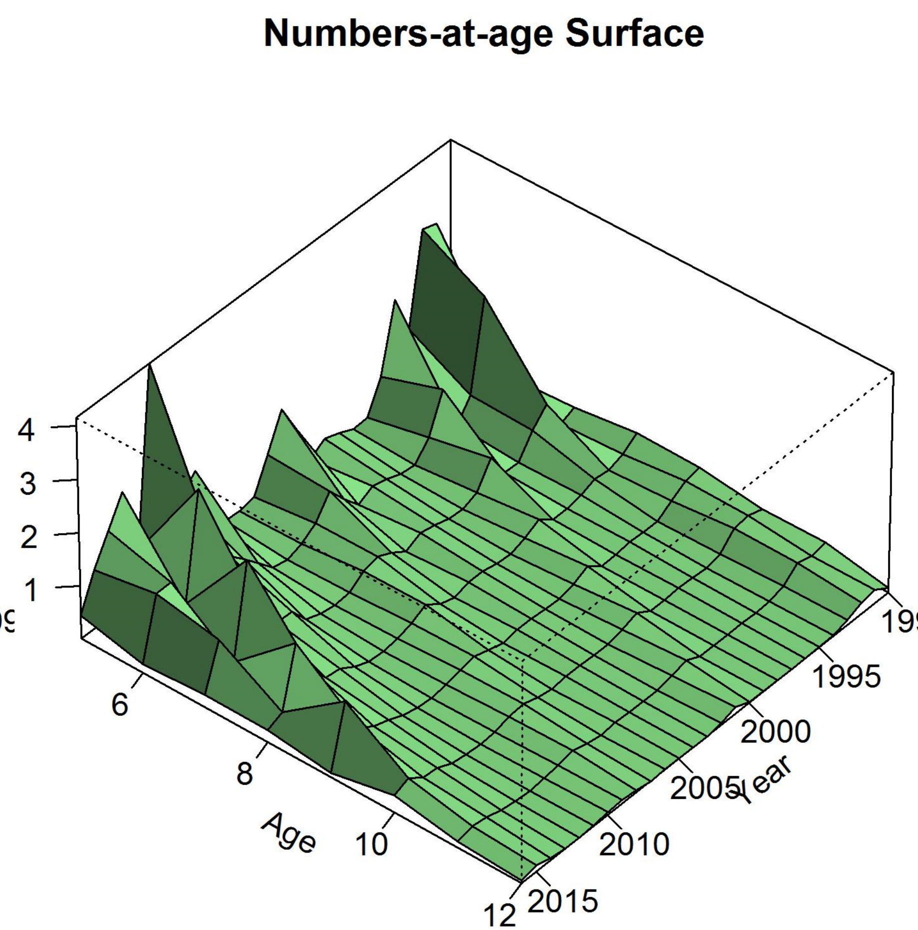
Fall Survey 3LNO cod SURBA

	Estimate	Std. Error	t value	Pr(> | t|)
logNo1	6.646	37.319	0.178	0.859
logNo2	6.371	14.226	0.448	0.655
logNo3	8.739	5.622	1.554	0.122
logNo4	7.260	2.471	2.938	0.004
logNo5	3.947	1.445	2.731	0.007
logNo6	2.653	1.239	2.141	0.033
logNo7	3.910	1.059	3.692	0.0003
logNo8	3.016	1.015	2.971	0.003
logNo9	3.043	0.988	3.080	0.002
logNo10	2.698	0.979	2.757	0.006
logNo11	2.759	0.870	3.170	0.002
logNo12	4.105	0.897	4.578	0.00001
logNo13	4.410	0.870	5.072	0.00000
logNo14	4.826	0.874	5.522	0.00000
logNo15	3.809	0.837	4.551	0.00001
logNo16	4.900	0.862	5.687	0.00000
logNo17	4.822	0.827	5.828	0.00000
logNo18	3.802	0.836	4.549	0.00001
logNo19	5.199	0.873	5.955	0
logNo20	4.005	0.850	4.712	0.00000
logNo21	3.956	0.892	4.435	0.00001
logNo22	4.152	0.885	4.693	0.00000
logNo23	4.466	0.913	4.890	0.00000
logNo24	5.291	0.994	5.323	0.00000
logNo25	6.457	1.193	5.413	0.00000
logNo26	8.492	11.408	0.744	0.457
logNo27	5.060	4.581	1.104	0.271
logNfirst1	5.911	29.202	0.202	0.840
logNfirst2	2.757	23.348	0.118	0.906
logNfirst3	1.627	19.147	0.085	0.932
logNfirst4	1.283	16.080	0.080	0.936
logNfirst5	0.214	13.830	0.015	0.988
logNfirst6	-1.284	12.182	-0.105	0.916
logNfirst7	-1.082	10.986	-0.099	0.922
logNfirst8	-0.711	10.145	-0.070	0.944
logNfirst9	-0.776	9.590	-0.081	0.936
logNfirst10	-1.162	9.292	-0.125	0.901
logNfirst11	-1.165	9.210	-0.127	0.899
logNfirst12	-2.290	9.646	-0.237	0.813
sparm1	-0.111	0.015	-7.478	0
sparm2	0.012	0.005	2.247	0.026
logf1	-1.609	81.640	-0.020	0.984
logf2	-1.609	31.030	-0.052	0.959
logf3	0.377	1.628	0.232	0.817
logf4	0.497	0.632	0.787	0.432
logf5	-0.647	1.179	-0.549	0.584
logf6	-0.531	0.898	-0.592	0.555
logf7	-0.127	0.550	-0.230	0.818
logf8	-1.609	2.417	-0.666	0.506
logf9	-1.609	2.336	-0.689	0.492
logf10	-1.609	2.241	-0.718	0.473
logf11	-1.609	1.982	-0.812	0.418
logf12	-0.390	0.573	-0.681	0.497
logf13	-0.548	0.674	-0.813	0.417
logf14	-0.292	0.526	-0.555	0.580
logf15	-1.609	1.968	-0.818	0.414
logf16	-0.499	0.649	-0.770	0.442
logf17	-1.594	1.905	-0.836	0.404
logf18	-1.609	1.929	-0.835	0.405
logf19	-0.370	0.566	-0.654	0.514
logf20	-1.577	1.895	-0.832	0.406
logf21	-0.612	0.727	-0.842	0.401
logf22	-0.918	0.983	-0.934	0.351
logf23	-1.609	1.960	-0.821	0.412
logf24	-0.820	0.896	-0.915	0.361
logf25	-1.609	16.040	-0.100	0.920
logf26	0.507	3.019	0.168	0.867
logf27	-0.696	4.108	-0.169	0.866

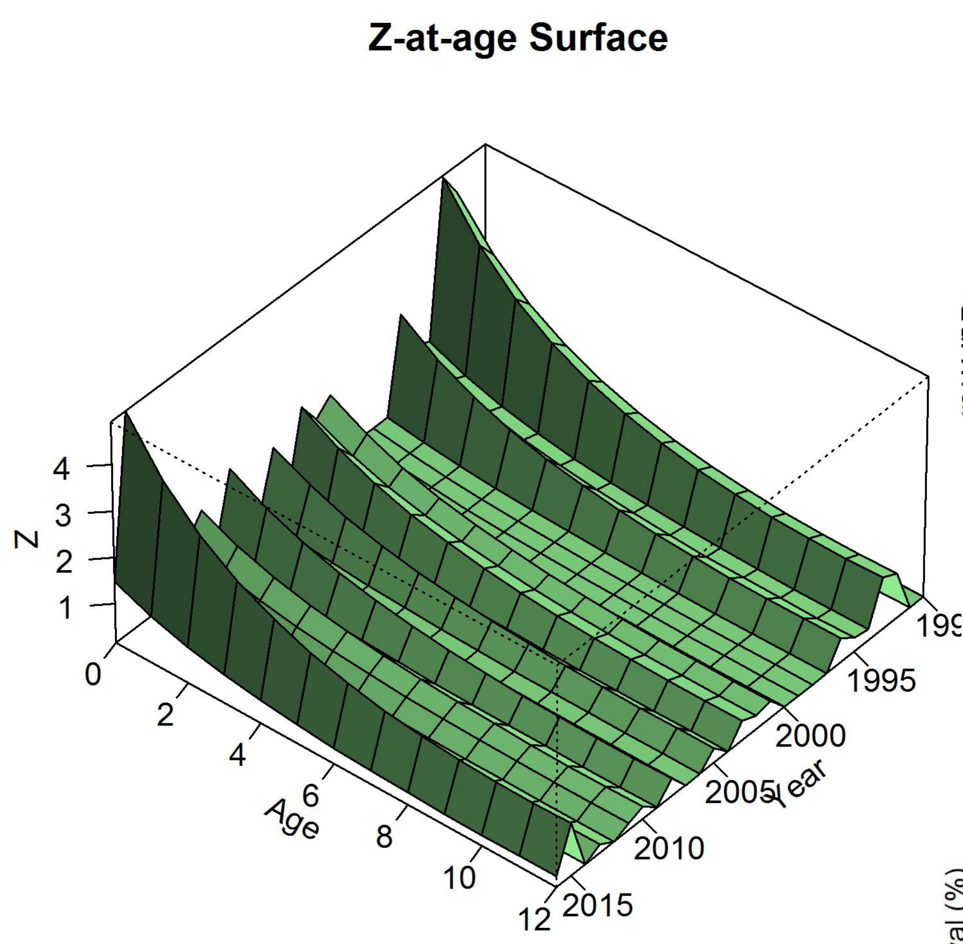
```
sum.surba = summary(surba.fit)

stargazer(sum.surba[11], type =
"html",  out="out.doc",align =
TRUE,
 title="Fall Survey 3LNO cod
SURBA", single.row=TRUE)
```

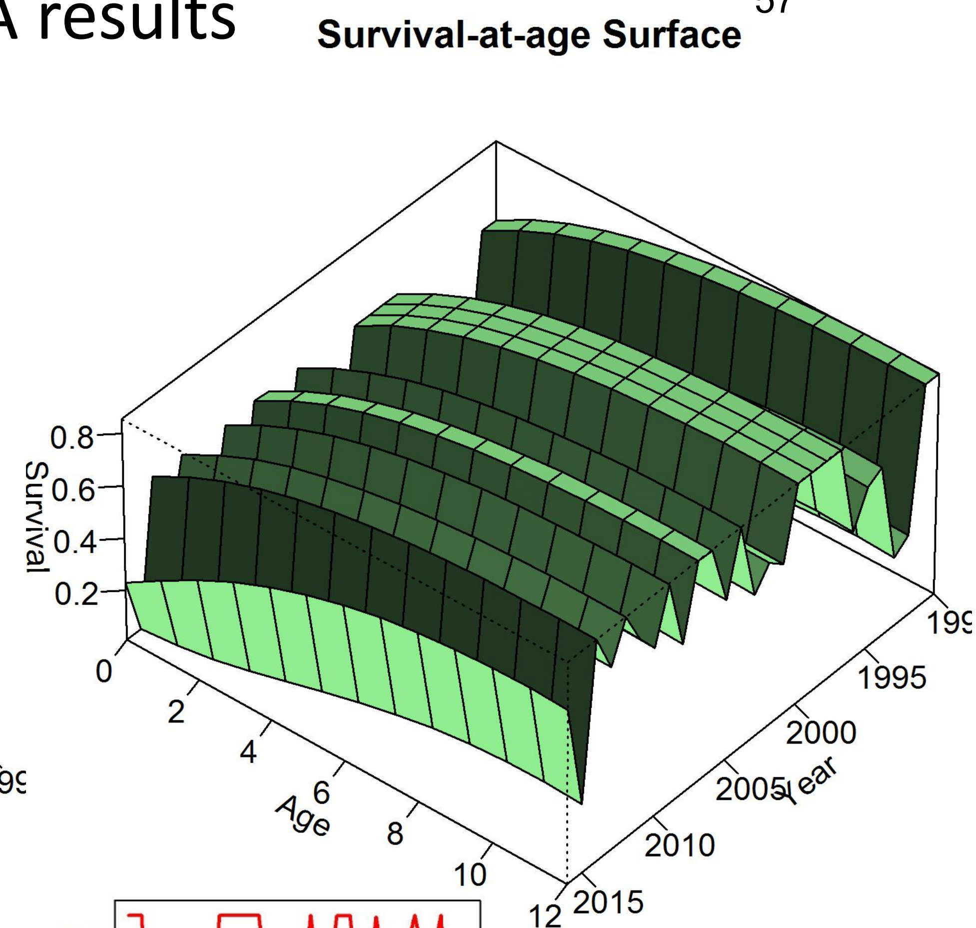
SURBA results

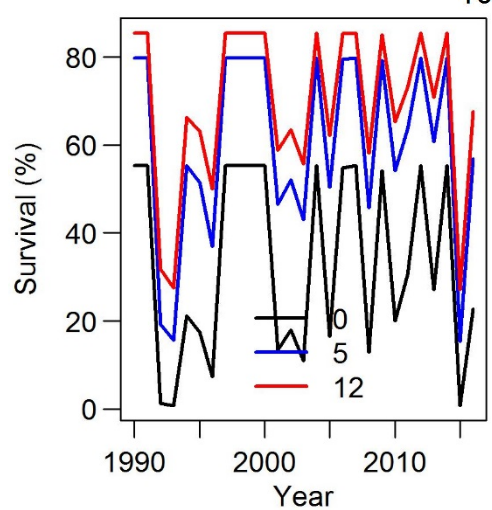




SURBA results

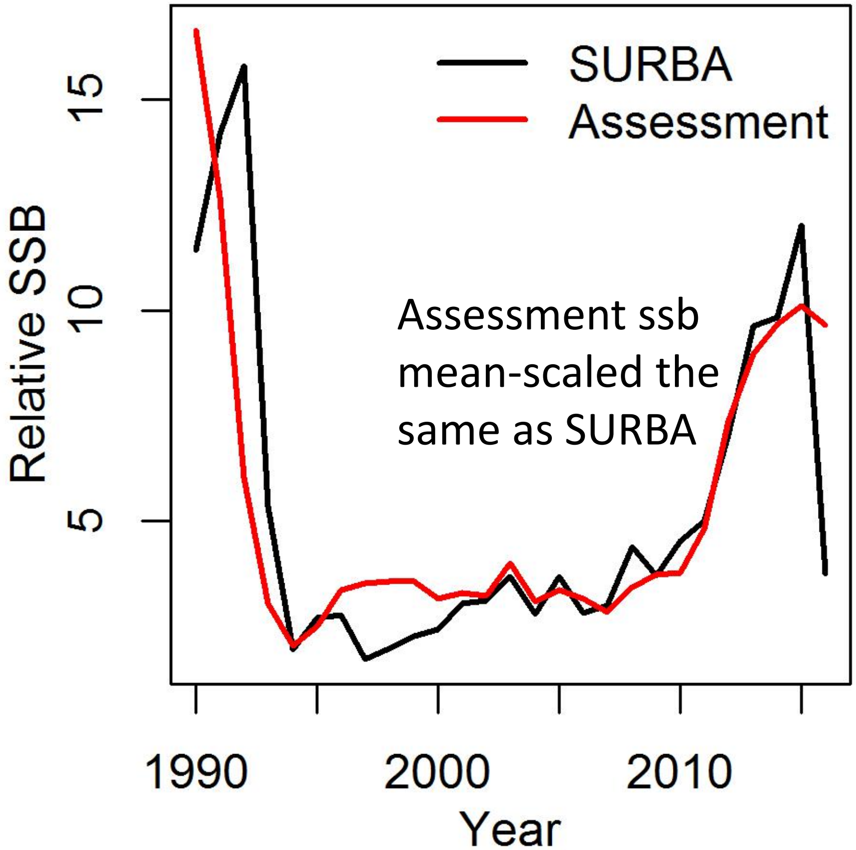


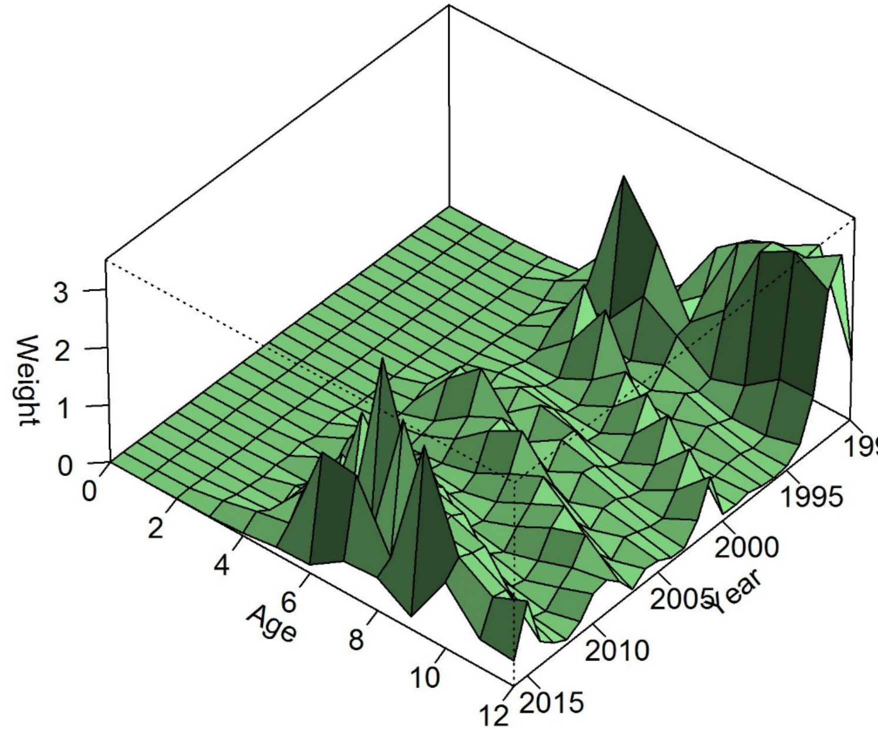


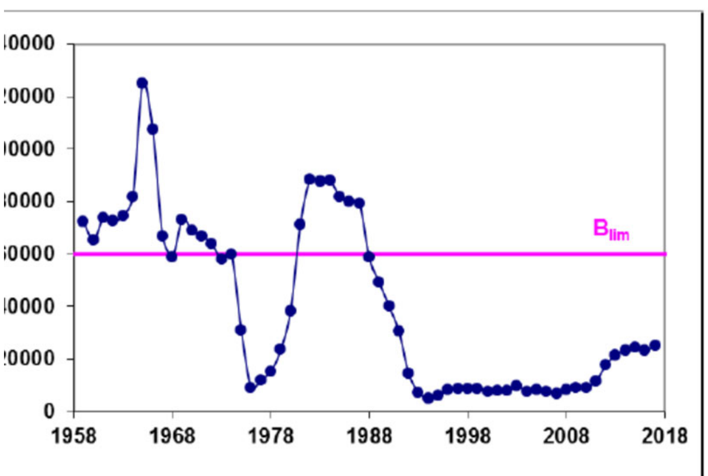


The high levels of Z do not seem plausible – probably confounded with No

Compared to Assessment







98

NOT TO BE CITED WITHOUT PRIOR
REFERENCE TO THE AUTHOR(S)

Northwest Atlantic Fisheries Organization

NAFO SCR Doc. No. 17-042

SCIENTIFIC COUNCIL MEETING – JUNE 2017

An Assessment of the Cod Stock in NAFO Divisions 3NO

by

R.M. Rideout, D.W. Ings, J. Brattey

Compared to Assessment

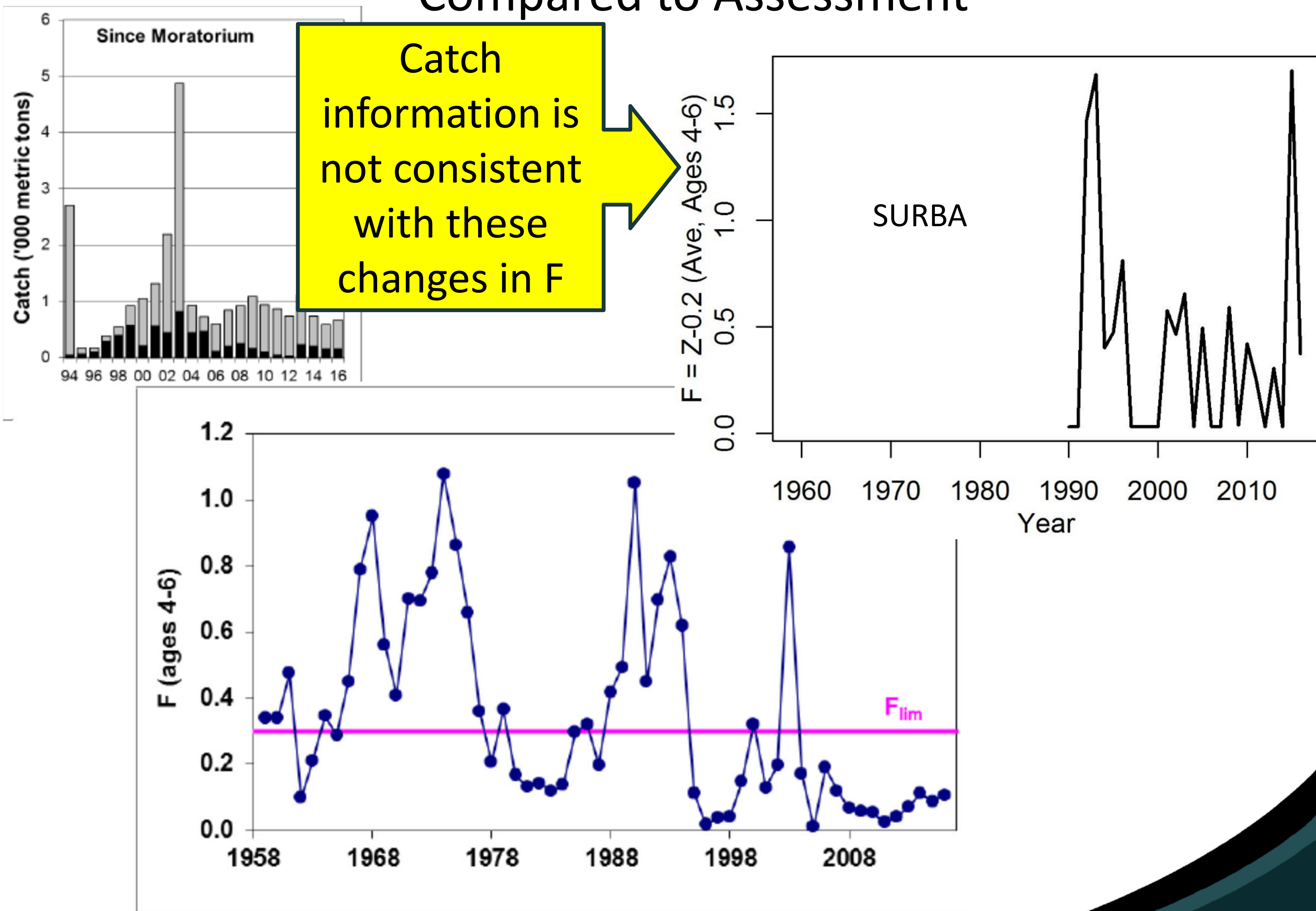


Fig. 23. Fishing Mortality for cod in Div. 3NO as estimated from ADAPT.

SURBA relative catch

- Can infer fishing mortality rates (*F*) using *F = Z – 0.2*.
- or use whatever value of *M* is appropriate.
- Can infer relative catch using Baranov catch equation.

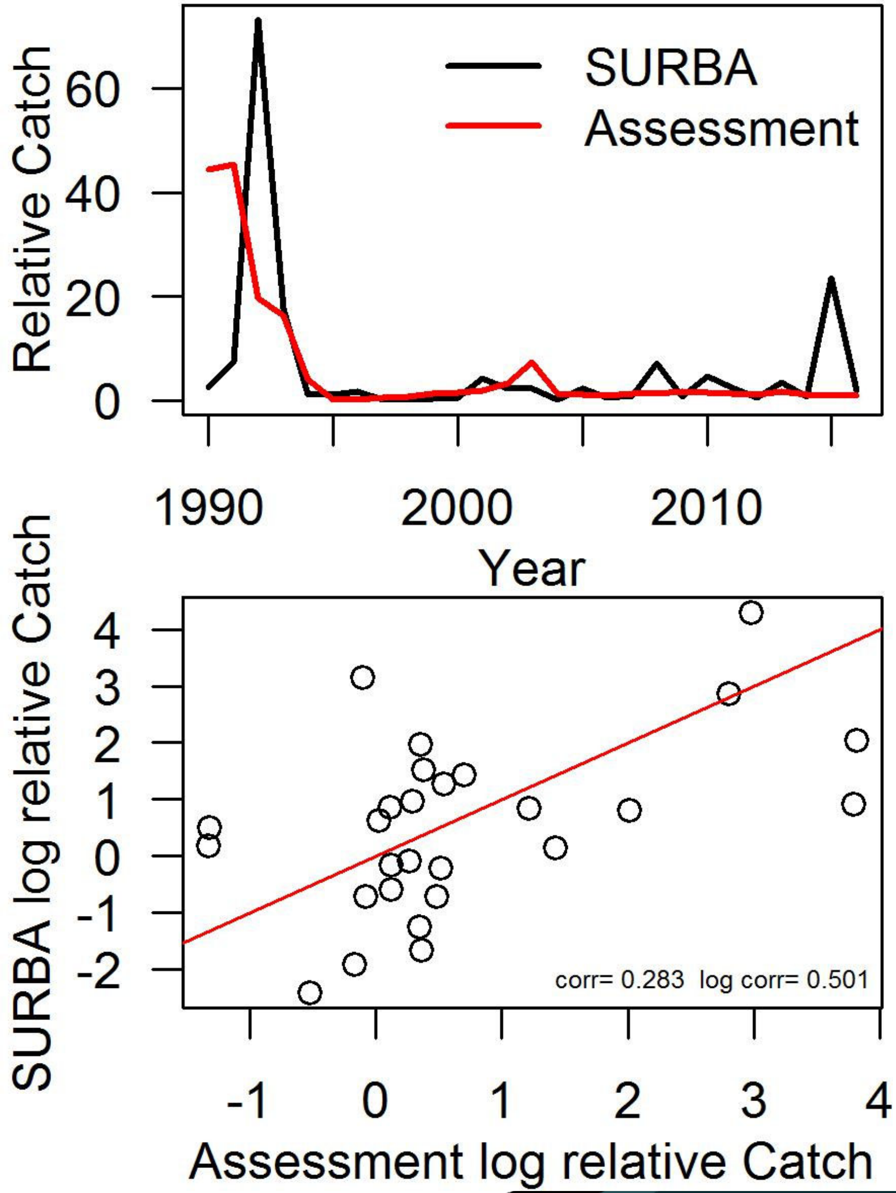
$$CW_{ay} = W_{ay} N_{ay} \left(1 - e^{-Z_{ay}}\right) \frac{F_{ay}}{M_a + F_{ay}}$$

- Useful for comparing with reported trends in landings.

3No Cod SURBA relative catch


```
pcw = wtm*N*(1-exp(-Z))*(abs(Z-0.2))/Z
pland = apply(pcw,1,sum)

assmnt.catch = c(28846,29454,12752,
10646,2702,172,174,383,547,919,1050,
1310,2194,4870,934,724,600,848,923,
1083,946,867,734,1113,734,586,666)

rel.catch = assmnt.catch*mean(pland)/
mean(assmnt.catch)
```

Not good agreement between SURBA and reported landings



SURBA fitting

In practise ad hoc penalties on parameter variation (i.e. shrinkage) were used to smooth between year changes in f_y's

$$l(\theta) = -\frac{1}{2}\sum_{a,y} \log(2\pi) + \log(\sigma^2) + \sigma^{-2} e_{a,y}^2 + \lambda_{f_y}\Delta_{f_y}^2 + \lambda_{s_a}\Delta_{s_a}^2,$$

θ – vector of all parameters

$$e_{a,y} = I_{a,y} / q_a N_{ay} \exp(-pZ_{ay}),$$

$$N_{a,y} = N_{a,y}(N_{1y}, f_y, s_a) = \text{a function of } N_{1y}\text{'s}, f_y\text{'s, and } s_a\text{'s},$$

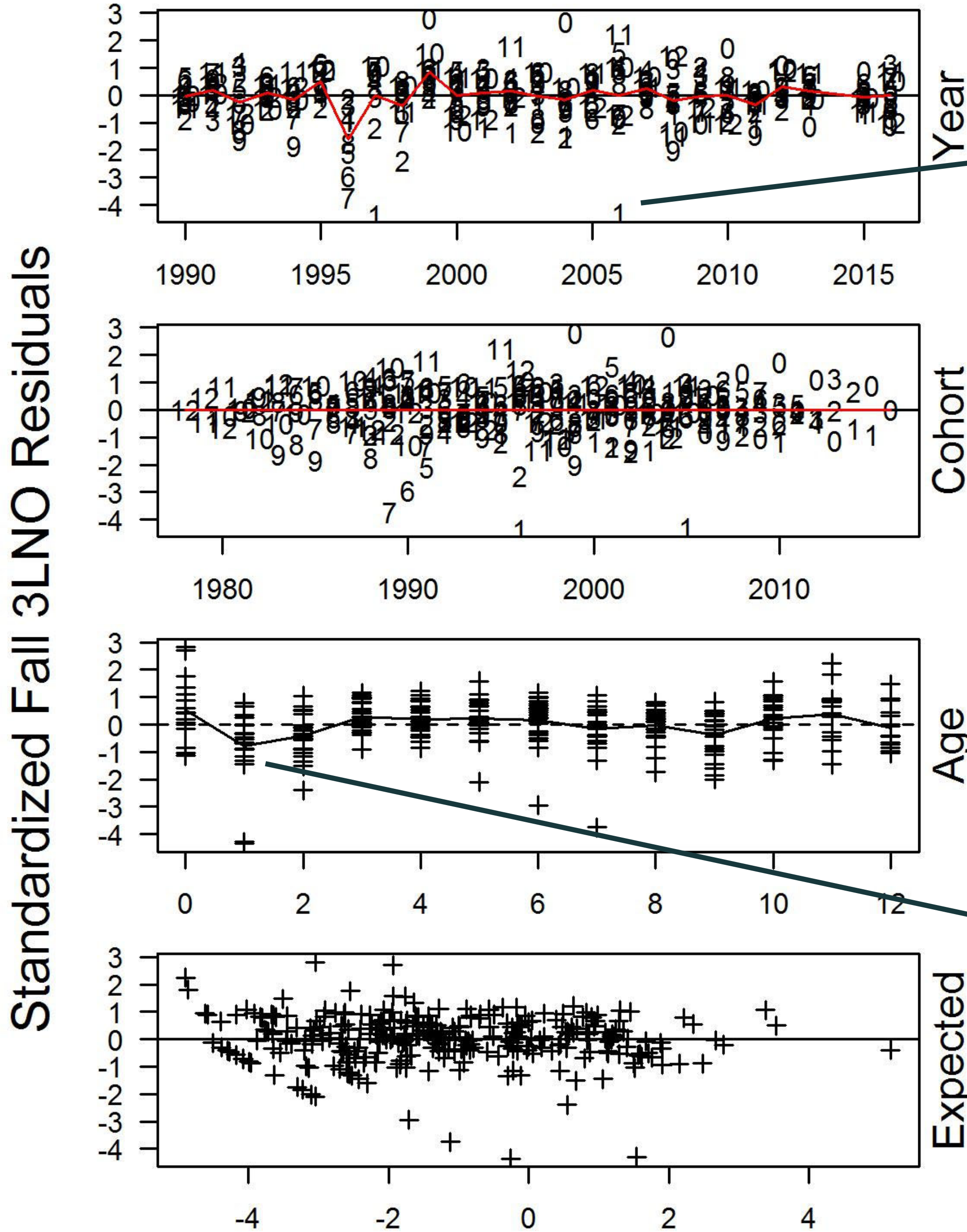
$$\Delta_{f_y} = \begin{cases} \log(f_y) - \log(f_{y-1}), & y > 1, \\ 0, & y = 1, \end{cases}$$

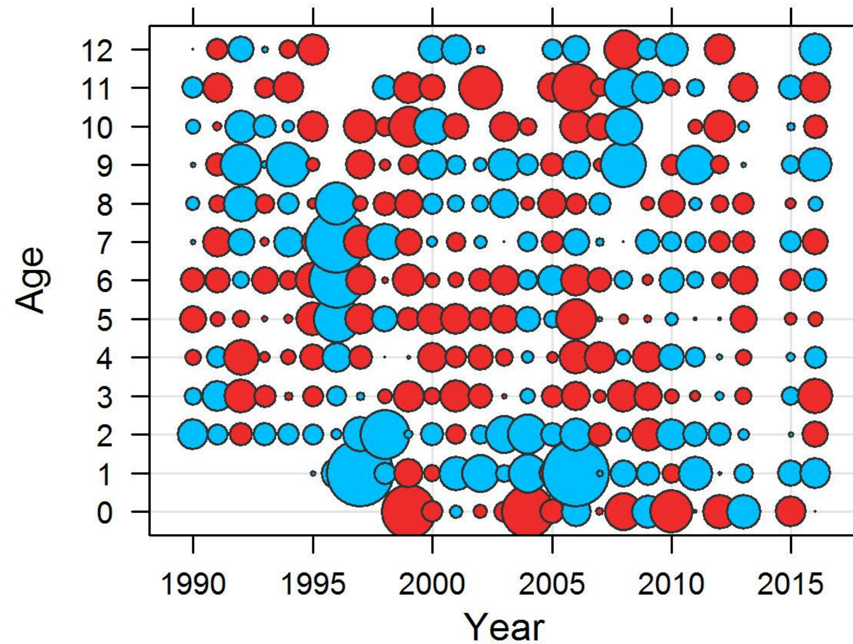
$$\Delta_{s_a} = \begin{cases} \log(s_a) - \log(s_{a-1}), & a > 1, \\ 0, & a = 1, \end{cases}$$

Not relevant for us because we used a parametric model for *s(a)*

I don't think we can get nls() to do this, but easy to setup using nlminb()

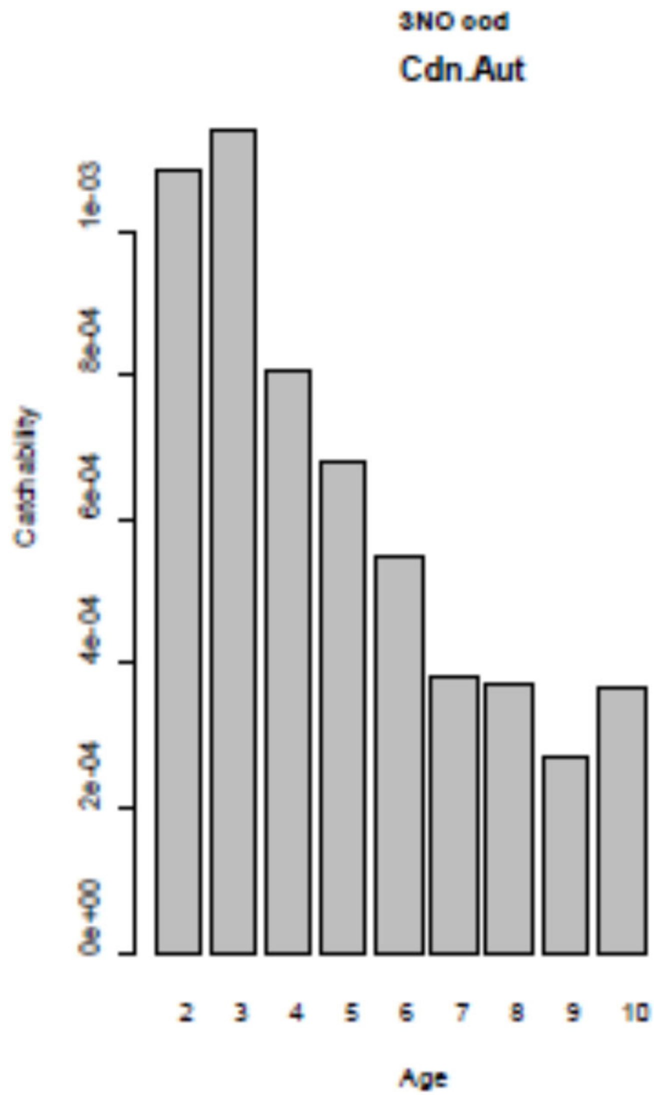
SURBA fit

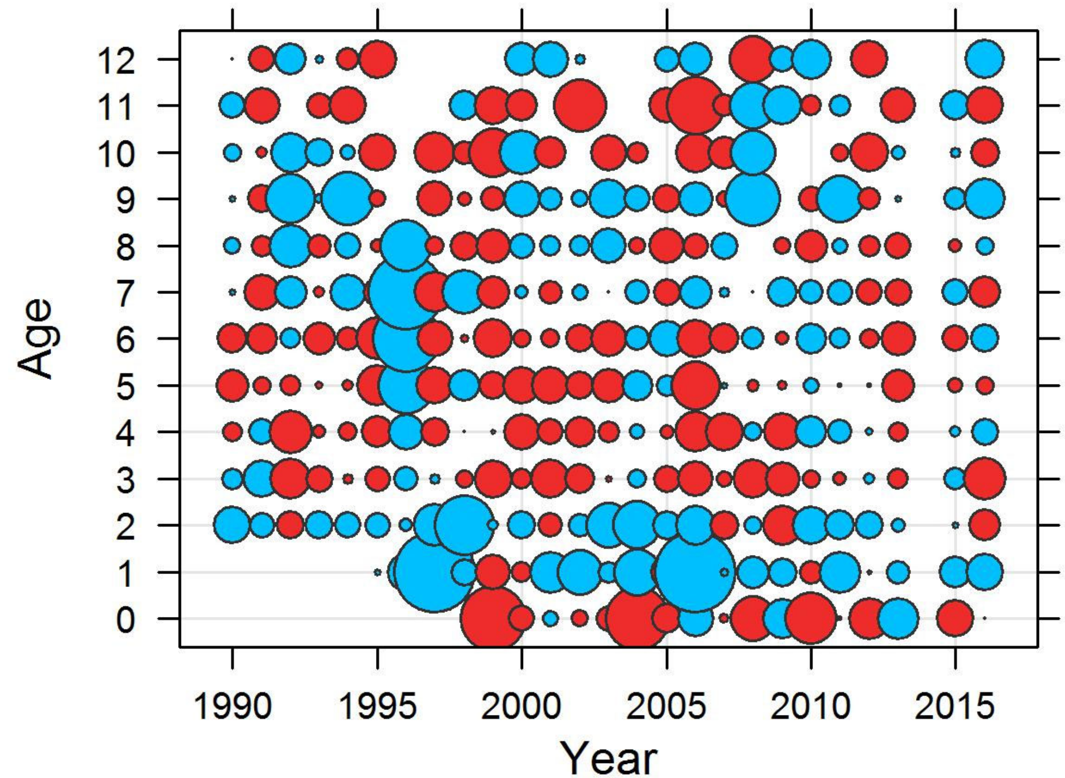




The assessment for this stock has a dome-pattern in q's?

SURBA fit





The assessment for this stock has a dome-pattern in q's?

They only use indices for ages 2-10?

SURBA

- SURBA does not use commercial catch sampling information, which is a deficiency
- But SURBA concept has much utility!

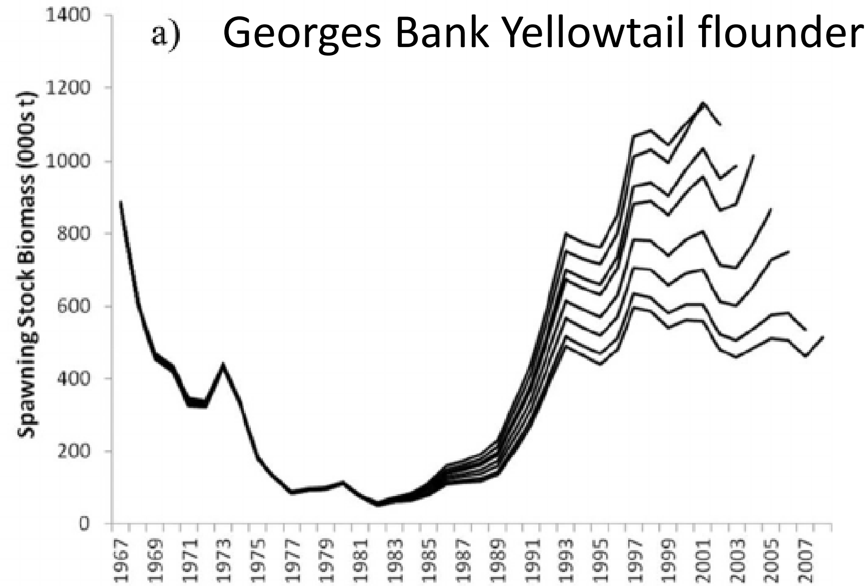
Kumar, R., Cadigan, N.G., Zheng, N., Varkey, D.A. and Morgan, M.J., 2020. A state-space spatial survey-based stock assessment (SSURBA) model to inform spatial variation in relative stock trends. *Canadian Journal of Fisheries and Aquatic Sciences*, *77*(10), pp.1638-1658.

In Prep: An age-based length-structured statistical catch-at-length model for hard-to-age fisheries stocks. Zhang, F. Cadigan, N,….

Retrospective Analysis

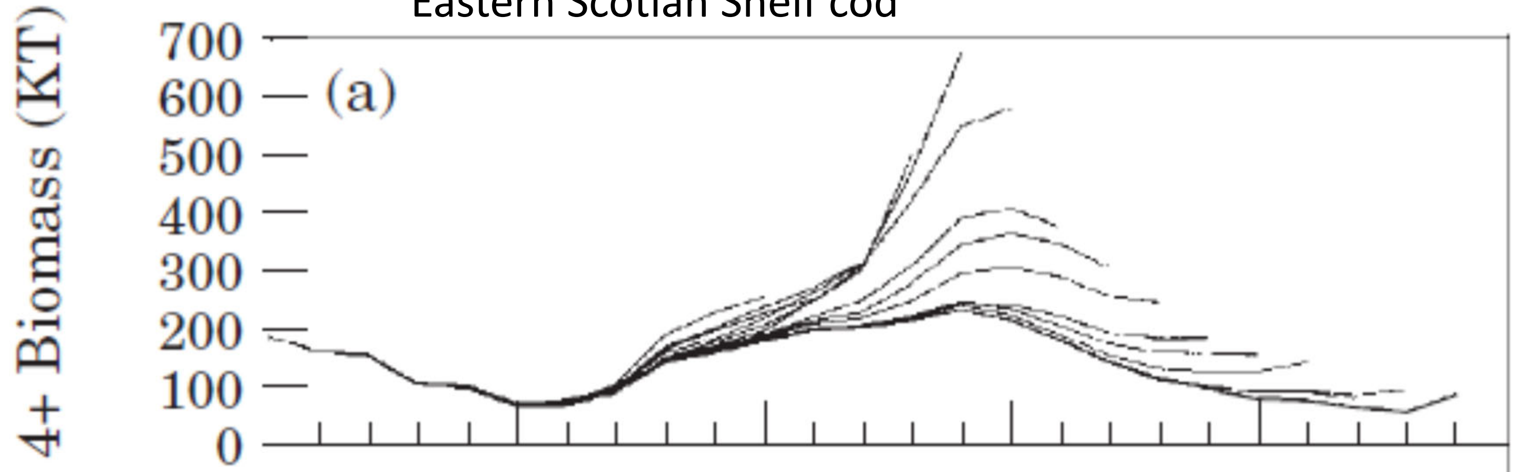
- Involves removing recent years of data (i.e.peels) and re-estimating model
- And then examining if assessment estimates change retrospectively
- A big problem if they change substantially
- Because it suggests that future assessment estimates of stock size this year will also change substantially

Bad Retro's



Mohn, R., 1999. The retrospective problem in sequential population analysis: An investigation using cod fishery and simulated data. *ICES Journal of Marine Science*, *56*(4), pp.473-488.

Eastern Scotian Shelf cod



Recent Retro Meeting

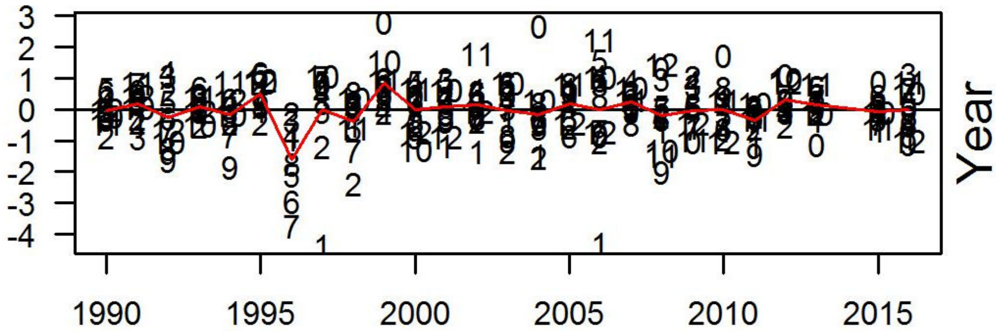
ICES
CIEM

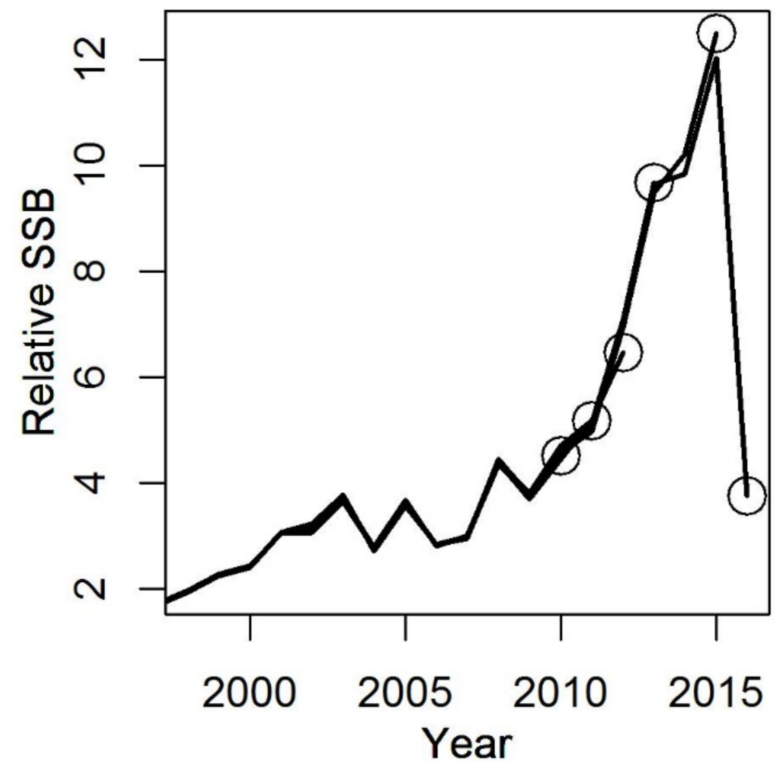
WORKSHOP ON CATCH FORECAST FROM BIASED ASSESSMENTS (WKFORBIAS; outputs from 2019 meeting)

VOLUME 2 | ISSUE 28

1. Examine the extent and the magnitude of retrospective bias identified in ICES category 1 and 2 stock assessments;
2. Determine the acceptability of such assessment; and
3. Investigate potential correction of population metrics using Mohn's rho statistics.

3NO cod SURBA retro

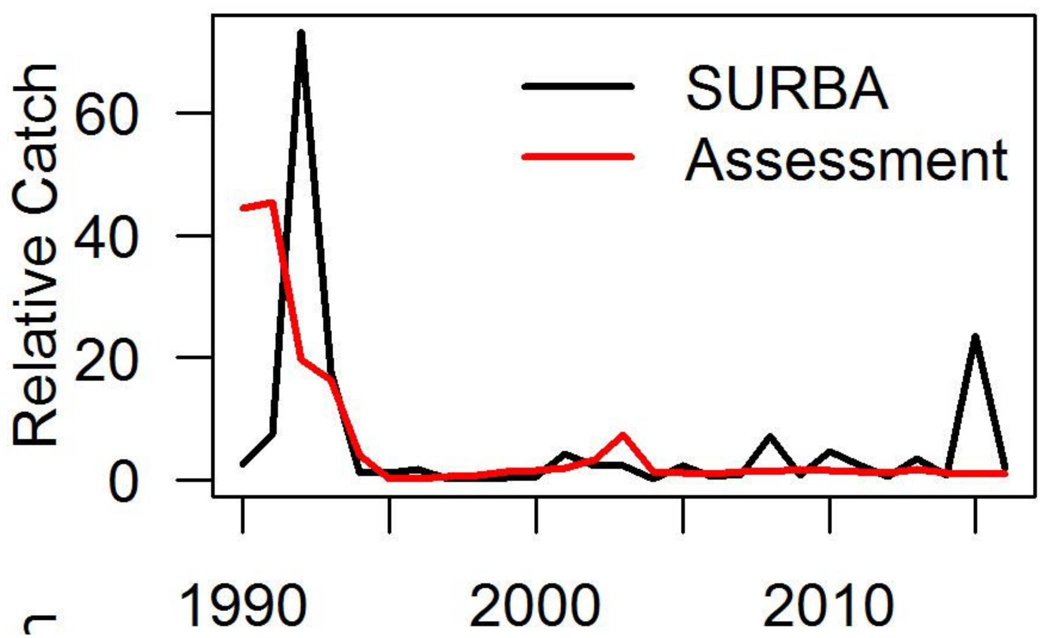




- Usually get retro patterns when there are temporal (i.e. annual) residual patterns
- Don’t have this with the 3NO cod SURBA

Retro's

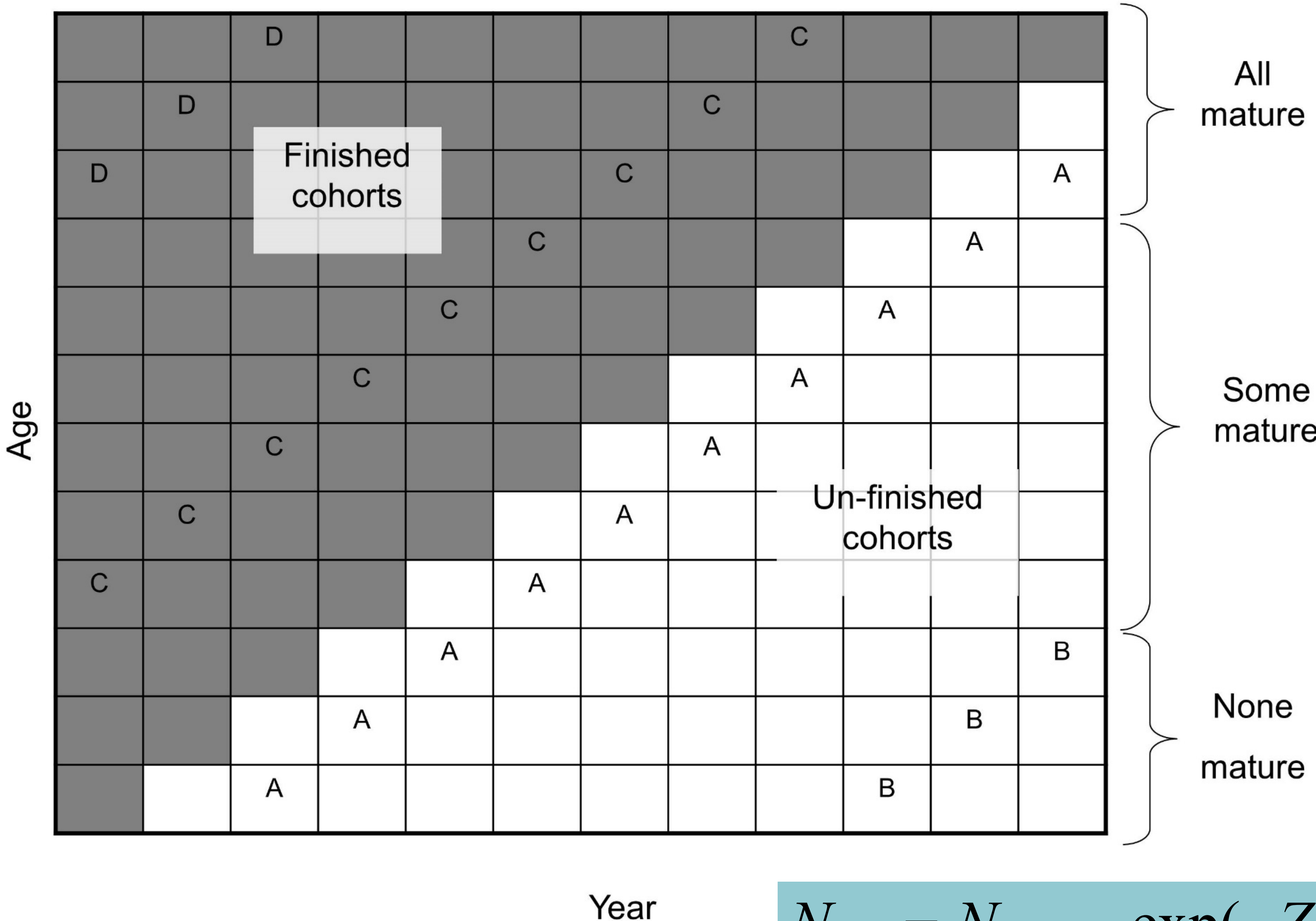
- Presence of a significant retrospective pattern means there is an assessment problem
- Lack of a retrospective pattern does not mean a good assessment



Harvest Advice

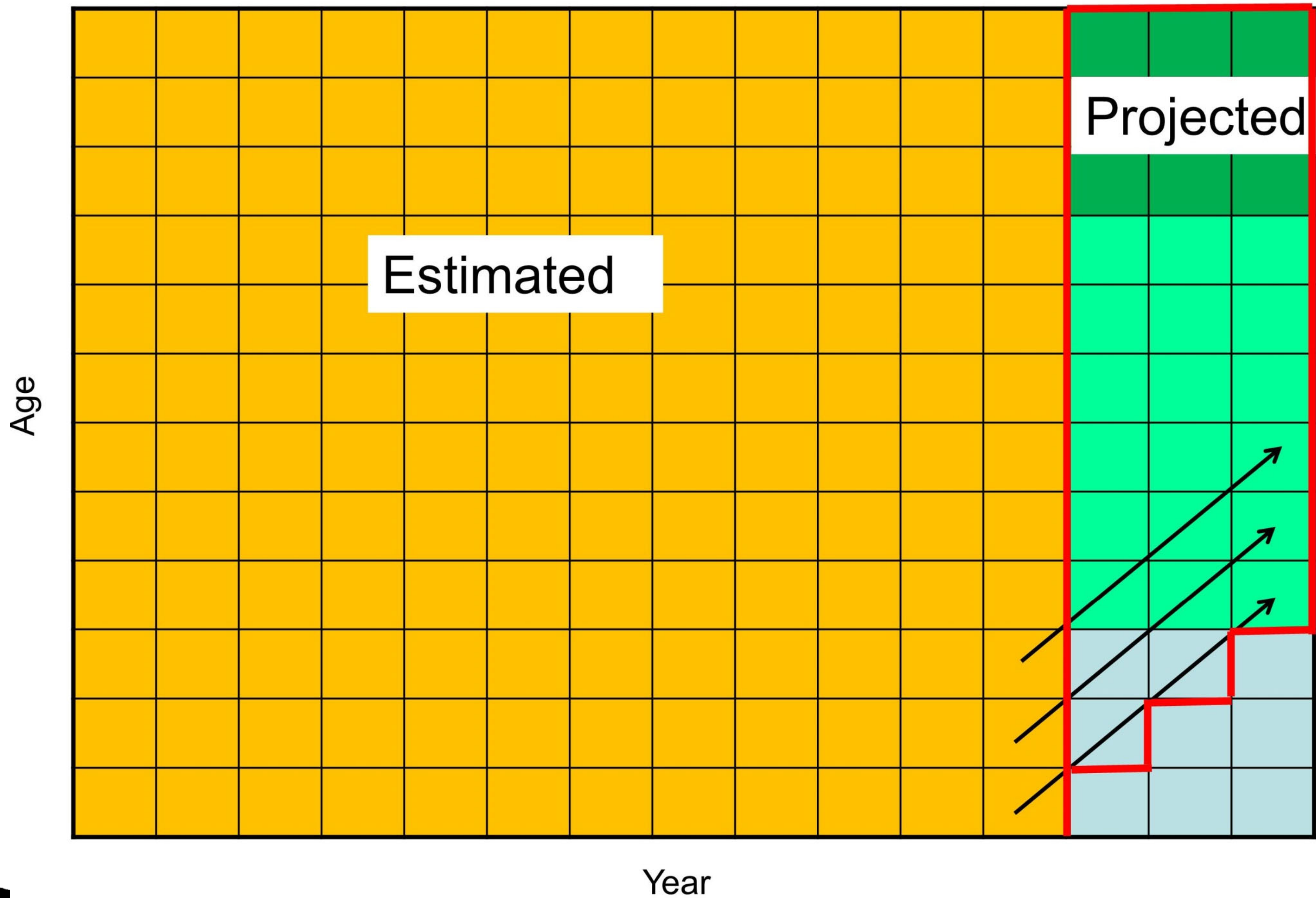
- Harvest advice is a problem
- Previously, SURBA used for 3Ps cod
- and 3-year Z-multiplier projections were provided.
- But managers don’t know what to do with these

Quick Review



$$N_{a,y} = N_{a-1,y-1} \exp(-Z_{a-1,y-1})$$

Projections for management advice



Catch multiplier projections

- We usually project selectivity based on average in the last 3 years

```
F = Z-0.2
ave.F = apply(F[(Y-2):Y,],2,mean)
sel_a = ave.F/max(ave.F)
```

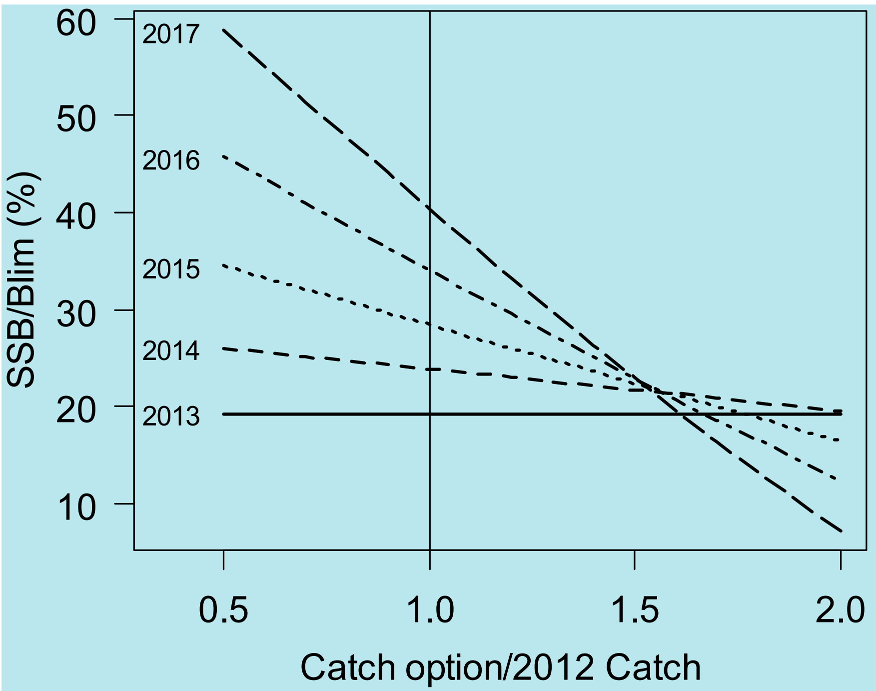
- If we provide a value for N_o in the projection year, we can project SSB at various levels of F using $F_a = F x s_a$ and $Z_a = F_a + M$.
- This is called an F projection
- For a catch projection we need to find the F such that the predicted catch equal what we want
- Then we do the projection and see if SSB does what we want (i.e. increase x% etc)

Catch multiplier projections

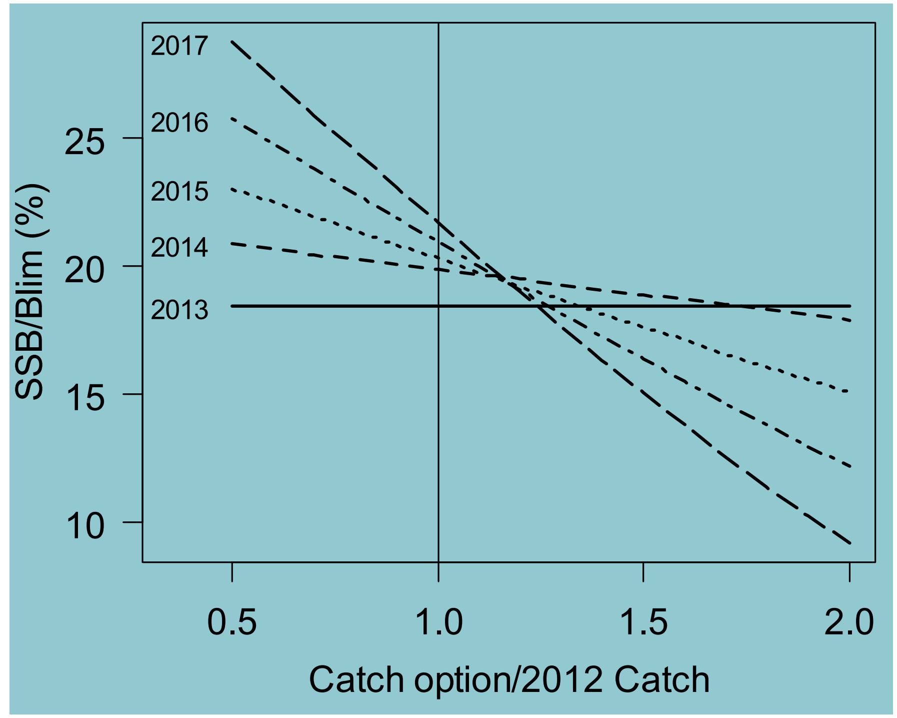
- The 3NO cod SURBA seems too implausible
- I won't bother doing projections
- But we will soon!

An Ncod example I did a few years back

M=0.2



M=0.4



This is direct harvest advice for fisheries managers!