Marine Institute

Lecture 11: Modern Stock Assessment Models

Noel Cadigan

The Center for the Advancement of Population Assessment Methodology (CAPAM) in collaboration with the National Institute of Water and Atmospheric Research Ltd (NIWA) hosted a technical workshop on the creation of frameworks for the next generation general stock assessment models in Wellington, New Zealand November 4-8, 2019.

CFER

Centre for Fisheries Ecosystems Research

MARINE INSTITUTE

MEMORIAL UNIVERSITY

Essential features of the next-generation integrated fisheries stock assessment package: A perspective

André E. Punt[a,b,*], Alistair Dunn[c], Bjarki Þór Elvarsson[d], John Hampton[e], Simon D. Hoyle[f], Mark N. Maunder[g,h], Richard D. Methot[i], Anders Nielsen[j]

next-generation stock assessment package needs to

a) be able to capture age and size/stage dynamics simultaneously yet computationally efficiently,
b) scale from data-rich to data-poor
c) include some multi-species capability, and
d) More appropriately deal with temporal variation (e.g., random effects and state-space models).

- In relation to data, there is a need to ensure that the next-generation stock assessment package better handles tagging data (age/size/stage models may help in this regard),
- in particular, to be able to use close-kin mark-recapture data.
- Efficient methods are needed to share parameter priors among stocks (satisfying the promise of the 'Robin Hood' paradigm).

F6004 Lecture 11 Outline

Modern Stock Assessment

- Integrated Assessments
- Data weighting
- State-space models
- Northern cod – NCAM
- Censored likelihoods
  - catch
  - survey zero's
  - partial surveys
- Age Compositions
  - Additive logistic normal multinomial
  - Multiplicative logistic normal multinomial
- Likelihood components
- Future
- Statistical Catch-at-Age (SCA): 3NO cod

Integrated Assessments

- Use as much data as possible.
- By integrating information we sometimes can deal with problems that are impossible to address using data sources individually
- Try to avoid 'model says xxx, but <u>other</u> data say something else'
- We will look at some details of the northern cod integrated model later this lecture

A state-space stock assessment model for northern cod, including under-reported catches and variable natural mortality rates
Noel G. Cadigan
Canadian Journal of Fisheries and Aquatic Sciences

Integrated Assessments. Why

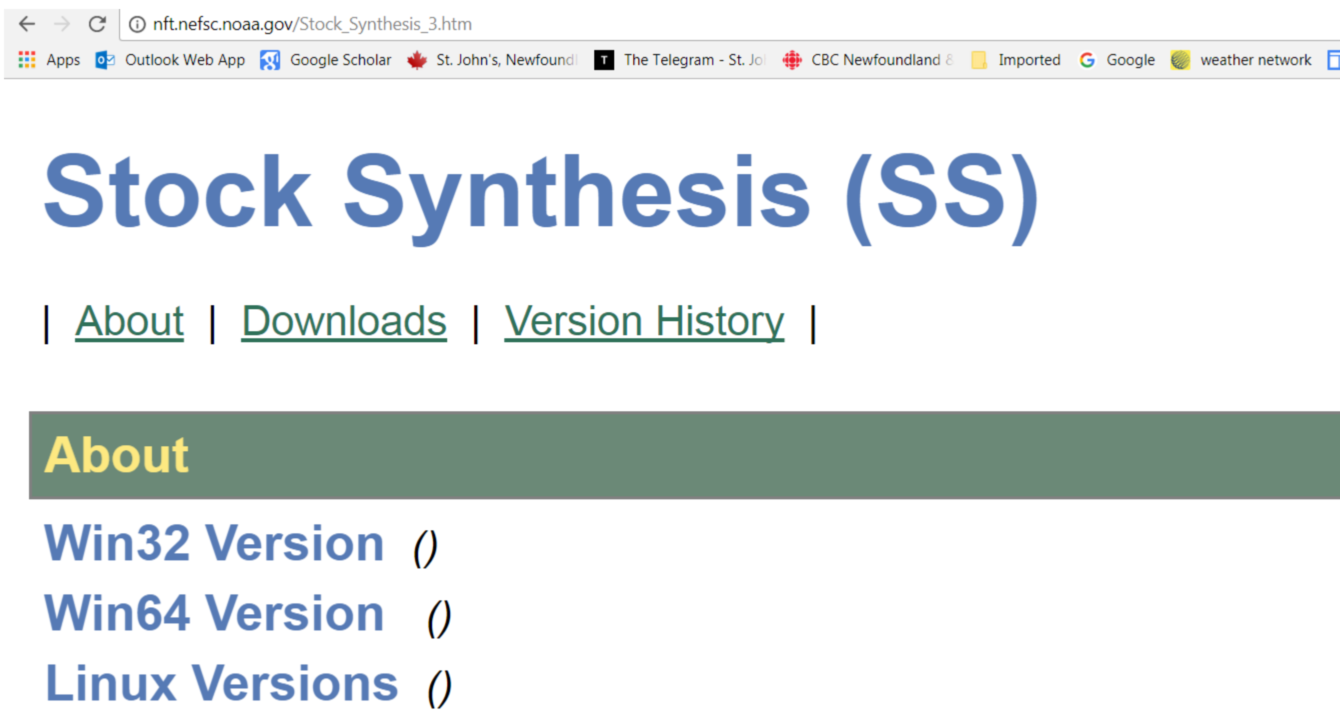
- Problems with the old approach (do statistics on statistics)

- Information may be lost when data are summarized
- two analyses may be logically inconsistent (i.e. based on substantially different assumptions)
- Difficulty in determining the appropriate likelihood function
- Difficulty in fully accounting for uncertainty
- Reduced diagnostic ability

A review of integrated analysis in fisheries stock assessment.
Maunder, M.N. and Punt, A.E., 2013. *Fisheries Research*, *142*, pp.61-74.

Integrated Assessments

- Packages may not be a good approach for a fully integrated analysis.
- Taylor-made software may work better.
- SS3 has attempted to be a software platform for integrated assessment (one ring to rule them all….).



SS

nft.nefsc.noaa.gov/Stock_Synthesis_3.htm

Apps | Outlook Web App | Google Scholar | St. John's, Newfound | The Telegram - St. Jo | CBC Newfoundland & | Imported | Google

Stock Synthesis provides a statistical framework for calibration of a population dynamics model using a diversity of fishery and survey data. It is designed to accommodate both age and size structure in the population and with multiple stock sub-areas. Selectivity can be cast as age specific only, size-specific in the observations only, or size-specific with the ability to capture the major effect of size-specific survivorship. The overall model contains subcomponents which simulate the population dynamics of the stock and fisheries, derive the expected values for the various observed data, and quantify the magnitude of difference between observed and expected data. Some SS features include ageing error, growth estimation, spawner-recruitment relationship, movement between areas. SS is most flexible in its ability to utilize a wide diversity of age, size, and aggregate data from fisheries and surveys. The ADMB C++ software in which SS is written searches for the set of parameter values that maximize the goodness-of-fit, then calculates the variance of these parameters using inverse Hessian and MCMC methods. A management layer is also included in the model allowing uncertainty in estimated parameters to be propagated to the management quantities, thus facilitating a description of the risk of various possible management scenarios, including forecasts of possible annual catch limits. The structure of Stock Synthesis allows for building of simple to complex models depending upon the data available.

CASAL

https://www.niwa.co.nz/fisheries/tools-resources/casal

CASAL

CASAL is an advanced software package developed by NIWA for fish stock assessment.

It is used for quantitative assessments of the status of most of New Zealand's fish stocks,including our deepwater (e.g. orange roughy), middle depth (e.g. hoki), inshore (e.g. snapper), and shellfish fisheries. Internationally, it has been used to assess Patagonian and Antarctic toothfish, and broadbill swordfish fisheries.

CASAL2 is an advanced software package developed by NIWA for modelling the population dynamics of marine species.

CASAL2 is NIWA's new integrated assessment tool for modelling population dynamics of marine species, including fishery stock assessments. CASAL2 expands functionality and maintainability over its predecessor, CASAL. CASAL2 can be used for quantitative assessments of marine populations, including fish, invertebrates, marine mammals and seabirds.

Introduction to CASAL

CASAL (**C++ Algorithmic Stock Assessment Laboratory**) is an advanced software package developed by NIWA for fish stock assessment. The software implements a generalised age- or length-structured fish stock assessment model that allows a great deal of choice in specifying the population dynamics, parameter estimation, and model outputs.

CASAL is designed for flexibility. It can implement either an age- or size-structured model, optionally also structuring the population by sex, maturity, and/or growth-path. It can be used for a single stock for a single fishery, or for multiple stocks, areas, and/or fishing methods. The user can choose the sequence of events in a model year. The data used can be from many different sources of information, for example catch-at-age or catch-at-size data from commercial fishing, survey and other biomass indices, survey catch-at-age or catch-at-size data, and tag-release and tag-recapture data. Estimation can be by either maximum likelihood or Bayes.

As well as generating point estimates of the parameters of interest, CASAL can calculate likelihood or posterior profiles and can generate Bayesian posterior distributions using Monte Carlo Markov Chain methods. CASAL can project stock status into the future using deterministic or stochastic recruitment and can generate a number of yield measures commonly used in New Zealand stock assessment, including MCY, CAY, F_{max}, $F_{0.1}$, deterministic MSY, and CSP.

Taylored Made: Disadvantages

- Computational demands:
  - Development time,
  - simulations,
  - sensitivity analyses,
  - inference,
  - projection uncertainty

- Convergence
- Parameter confounding
- Model misspecification
- Model selection
- Data weighting
- Level of data aggregation
- Transparency
- **MSE**

Data Weighting

- A complicated issue, esp. integrated models
- And models that use statistics as data observations
- Weight by sample size? What is it?
- Weight by sample sd's? what about process error?
- CAPAM workshop on this issue. SS3 centric

CAPAM

In 2012, the **Center for the Advancement of Population Assessment Methodology** (CAPAM) was established to: (1) improve quantitative methods generally used in stock assessment modeling efforts, whereby research is focused on parameterization and simulation involved in determining good practices for developing robust fishery models; and (2) afford the educational and training opportunities necessary to produce competent researchers and ultimately, the next generation of stock assessment scientists.

www.capamresearch.org/home/purpose

Apps Outlook Web App Google Scholar St. John's, Newfound

CAPAM

Center for the Ac

Home | Programs | Staff | Workshops | Forum | Publications

Purpose

Framework

Funding

Purpose

In 2007, the Reauthorization of the Magnuson-Stevens Fishery Conservation and Management Act strengthened the role of science in fishery management decision-making and required that formal Fishery Management Plans adopt catch limits and related accountability measures for exploited marine populations. An important stipulation in this legislation required that Federal marine management more efficiently monitor fisheries in the future to stem overfishing and develop long-term sustainable fishing practices. To achieve these goals, the National Oceanic and Atmospheric Administration (NOAA Fisheries) identified critical needs for: (1) higher quality and more frequent stock assessments on a comprehensive range of species that can be used to objectively provide sound advice to fishery managers concerning appropriate harvest strategies; and (2) increased attention to preparing potential researchers for stock assessment work, including educational commitments and focused research on quantitative methods generally employed in population ecology and specifically, stock assessment modeling.

CAPAM Updates:

Stock Synthesis Course, Chile NEW

CAPAM awarded AIFRB Outstanding Group Achievement Award NEW

Spatio-temporal modeling Mini-Workshop 2nd Announcement NEW

Spatio-temporal modelling Mini-Workshop Registration Open NEW

Stock Assessment Course, Italy

Recruitment Workshop Draft

Data Weighting

- Important to get the likelihoods right (e.g. not multinomial for compositions)
- State-space models seem like a good framework to address data weighting

Fisheries Research
Volume 192, Pages 1-140 (August 2017)

Data conflict and weighting, likelihood functions, and process error Data weighting Edited by Mark N. Maunder, Paul R. Crone, Andre E. Punt, Juan L. Valero and Brice X. Semmens

Data conflict and weighting, likelihood functions and process error
Pages 1-4
Mark N. Maunder, Paul R. Crone, Andre E. Punt, Juan L. Valero, Brice X. Semmens
PDF (219 K)

Revisiting data weighting in fisheries stock assessment models Original Research Article
Pages 5-15
R.I.C.Chris Francis
Abstract | PDF (732 K)

Dealing with data conflicts in statistical inference of population assessment models that integrate information from multiple diverse data sets Original Research Article
Pages 16-27
Mark N. Maunder, Kevin R. Piner
Abstract | Close research highlights | PDF (1337 K) | Supplementary content

Statistical State-Space Models

- the population dynamics model (latent state), including how process error is modeled
- the data and the associated observation model (the space), including the likelihood function, which represents the sampling process
- State-space models have become the favored approach in modeling time varying ecological phenomena (Pedersen et al., 2011).
- Maximum likelihood and Bayesian methods have become the standard in model fitting
- **F6005**

Northern Cod

- 2013 DFO Science Advisory Report:
- Total fishery removals are very uncertain;
- Catch-based analytical models not used.
- stock status inferred from trends in DFO RV survey indices (directly or modelled – SURBA+), and
- other indices and harvest rates from tagging studies.

- *<u>Can't</u> provide catch advice with this approach.*

Another Ncod problem: M

- In stock assessment models M is the natural mortality rate.
- M may have increased substantially around the time of the moratorium in 1992, and
- this persisted for some years after, although the level and duration is uncertain (Lilly, 2008).
- Change and uncertainty about M is a major issue for other NW Atlantic cod stocks

Another Ncod problem: Q

- Q is a parameter that links stock size (N) to survey results (I).
- Survey index of stock size; i.e. $I = Q \times N$.
- Q should be constant over time if survey trends are used directly to infer stock trends.
- There are good reasons to think Q changed substantially since the moratorium for the DFO RV survey.

Solution?

- How to deal with the CxMxQ problems????
- Solution: an integrated state-space model.
- Such models utilize all possible relevant information on population dynamics.
- By integrating information we can deal with problems that are impossible to address using data sources individually.

Northern Cod Assessment Model (NCAM)

- An <u>integrated</u> <u>state-space</u> population dynamics model
- <u>State-space</u>: Includes model process errors and observation equations for all data sources
- <u>Integrated</u>: Use as much information as possible
- Process errors necessary for sensible stochastic projections

Cadigan, N. 2016. Updates to a Northern Cod (*Gadus morhua)* State-Space Integrated Assessment Model. DFO Can. Sci. Advis. Sec. Res. Doc. 2016/022. v + 58 p.

Data used in NCAM in 2016

- Reported catches considered lower bounds (1983-2015)
- Catch age composition information (2-14; 1983-2015)
- *New! Landings by month (1983-2015)*
- Offshore DFO survey indices (1983-2015)
- Inshore Sentinel gillnet survey indices (1995-2014)
- Inshore acoustic biomass surveys and age composition data (1995-2009). *Can do offshore surveys too*.
- Tagging data, 116K releases, 11K recaptures
  - Ages estimated
  - 280 experiments during 1983-2015.
  - Adjustments for tagging mortality, tag loss, **reporting rates**

and there is still more data out there!

Other Biological Inputs

- Proportion mature at age (mats; modelled)
- Beginning of year and mid-year stock weights at age (modelled)
- Baseline (mean) M:
  - 0.5 age 2; 0.3 age 3; 0.2 ages 4+
- Projection inputs for mats, weights, baseline M

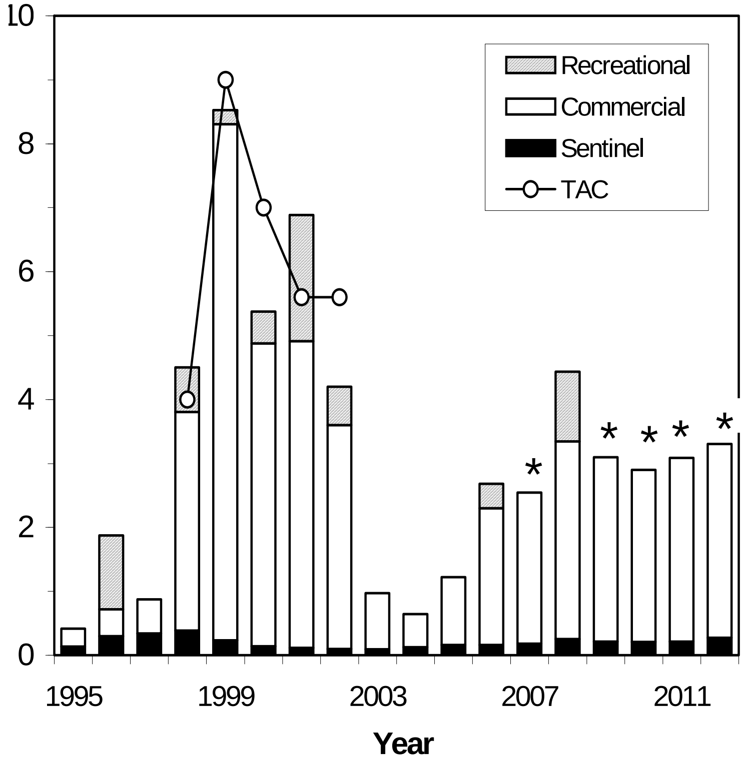
Under the NCAM hood:

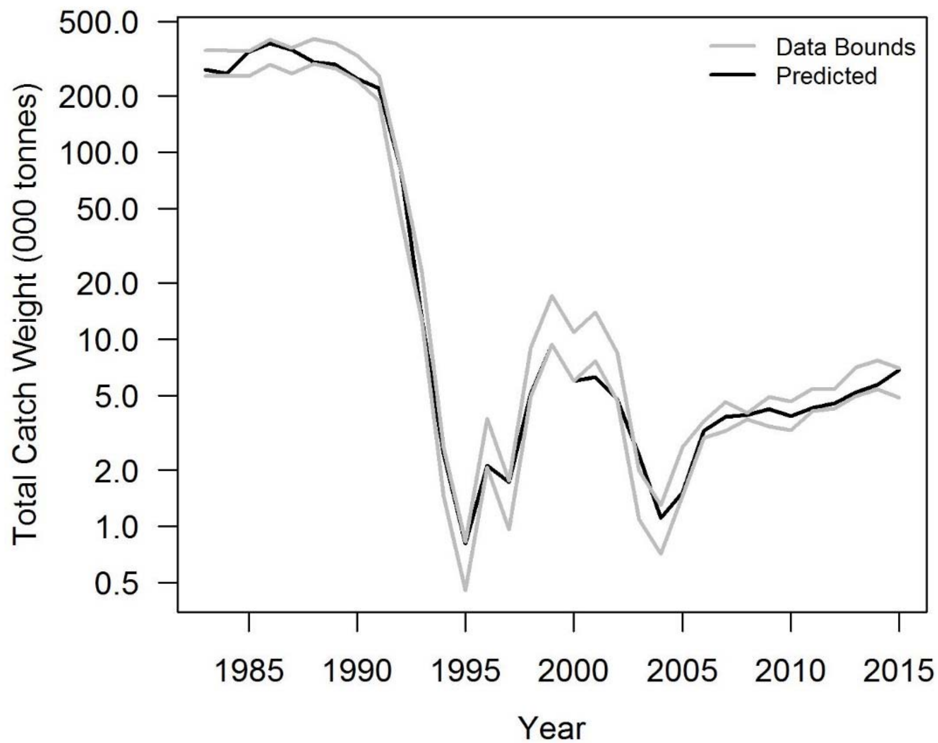
- Censored likelihood (L) for total catches
- Compositional L for catch age comps
- Lognormal L for surveys, but with a censored component for many zeros in DFO RV survey
- Lognormal L for inshore acoustic biomass
- Negative Binomial L for acoustic age comps
- Poisson L for tag catches, but with random exp. F's; or Negative Binomial L with stock F's
- Likelihood approximation for reporting rates
- Censored L for offshore acoustic biomass

NCAM: Catches

Ncod reported catch treated as a lower bound, and user sets upper bound

bounds used via a censored likelihood (Hammond and Trenkel (2005)





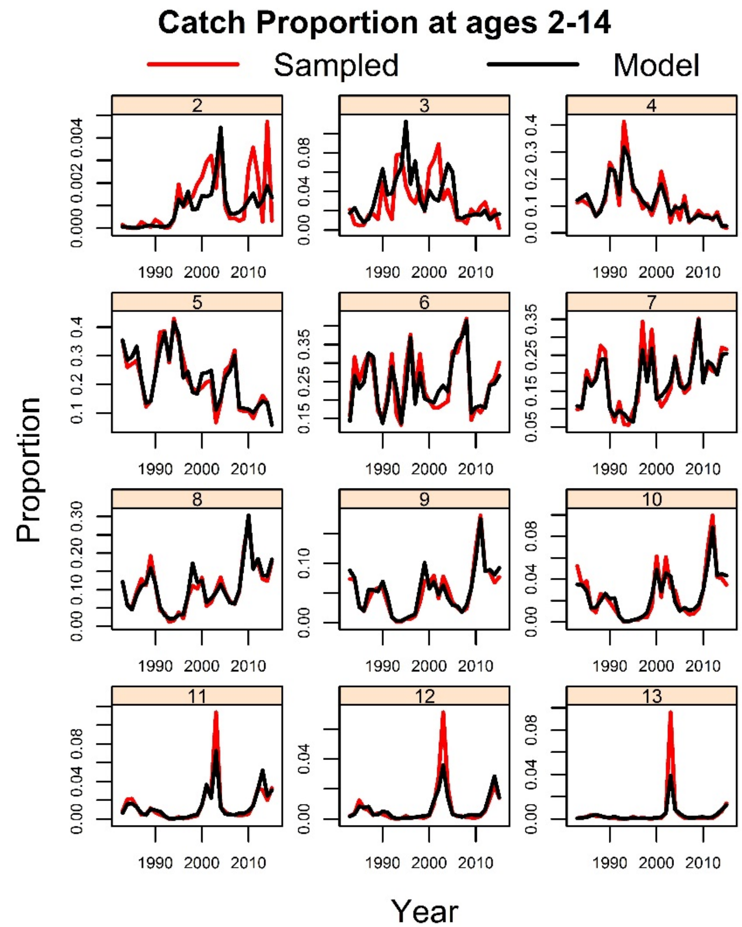
Detecting and correcting underreported catches in fish stock assessment: trial of a new method

Nicolas Bousquet, Noel Cadigan, Thierry Duchesne, and Louis-Paul Rivest

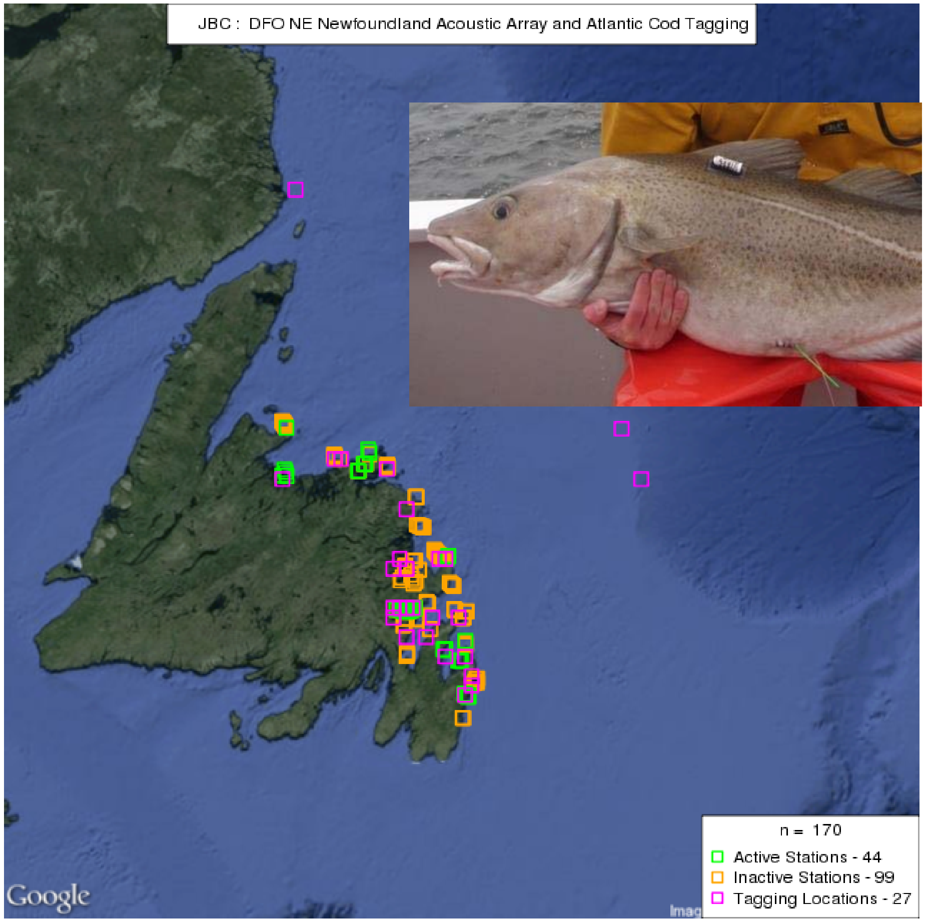
Catches

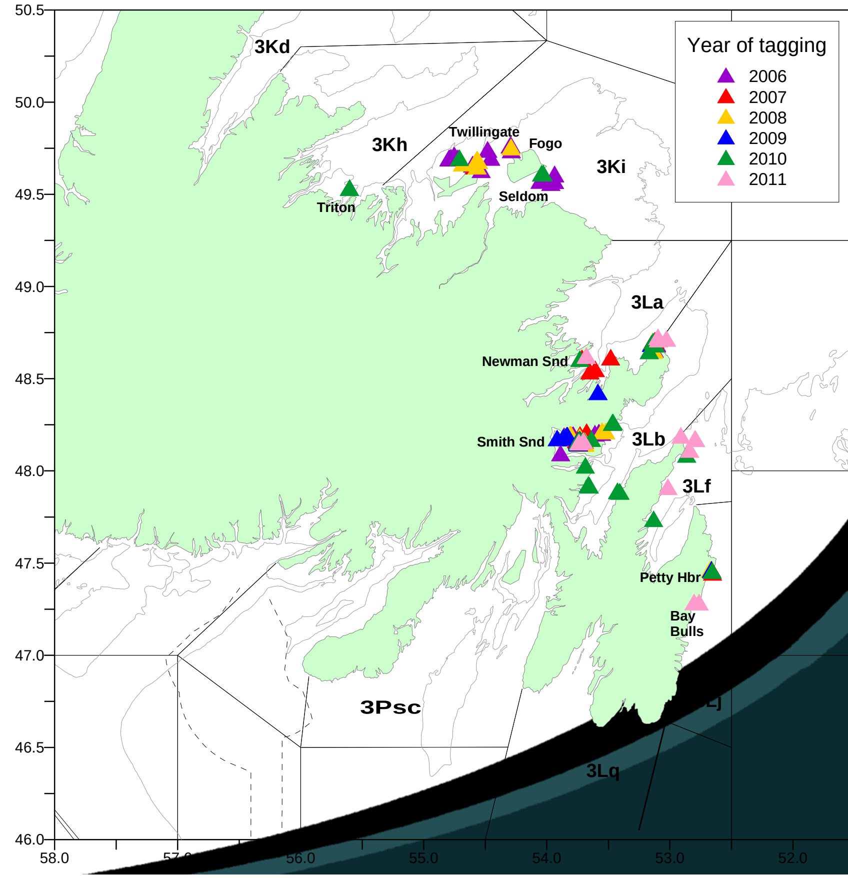
- statistical properties of the catch age composition information are very complex.

- I used an approach recommended for compositional data.

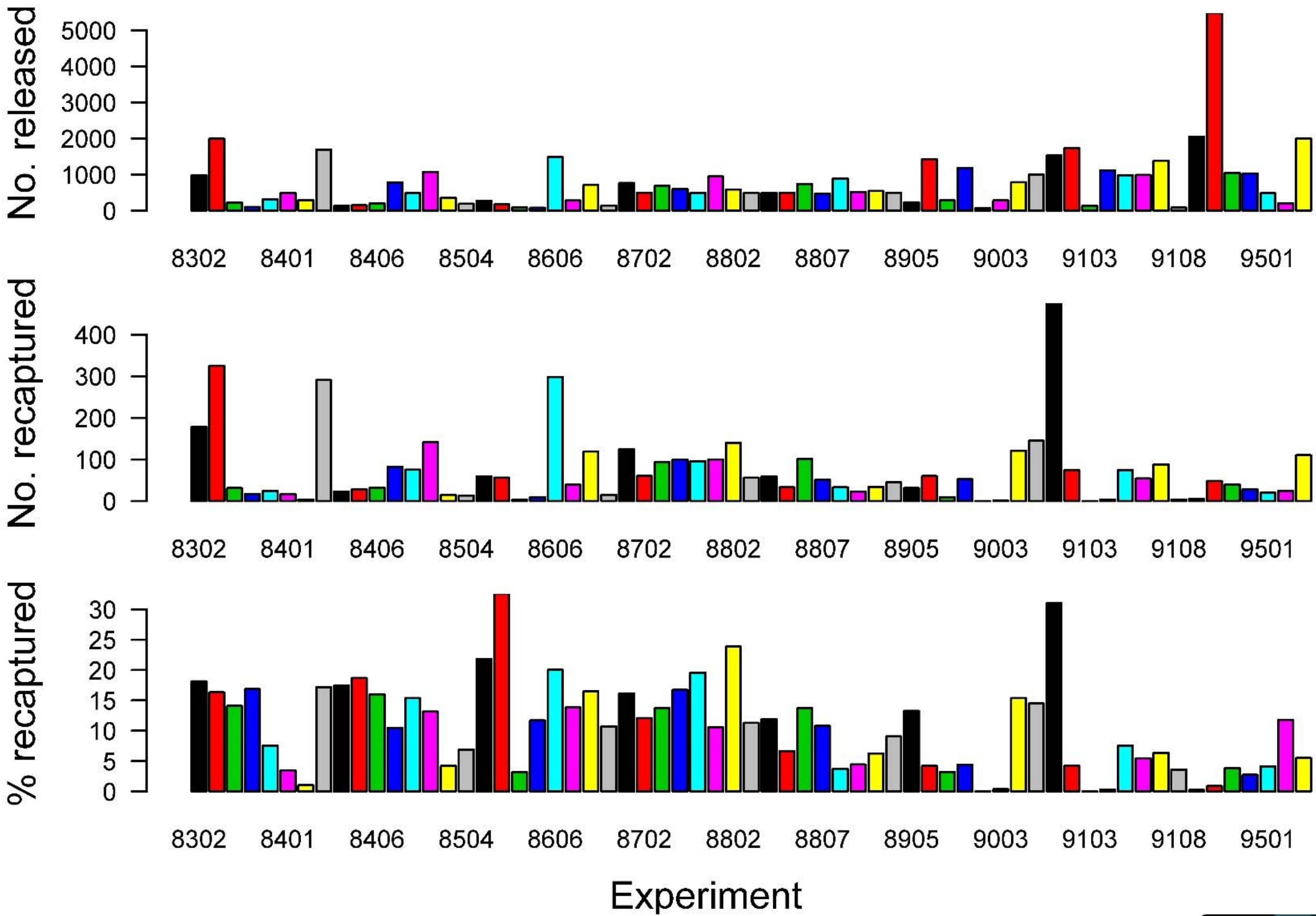


Data: Tagging data

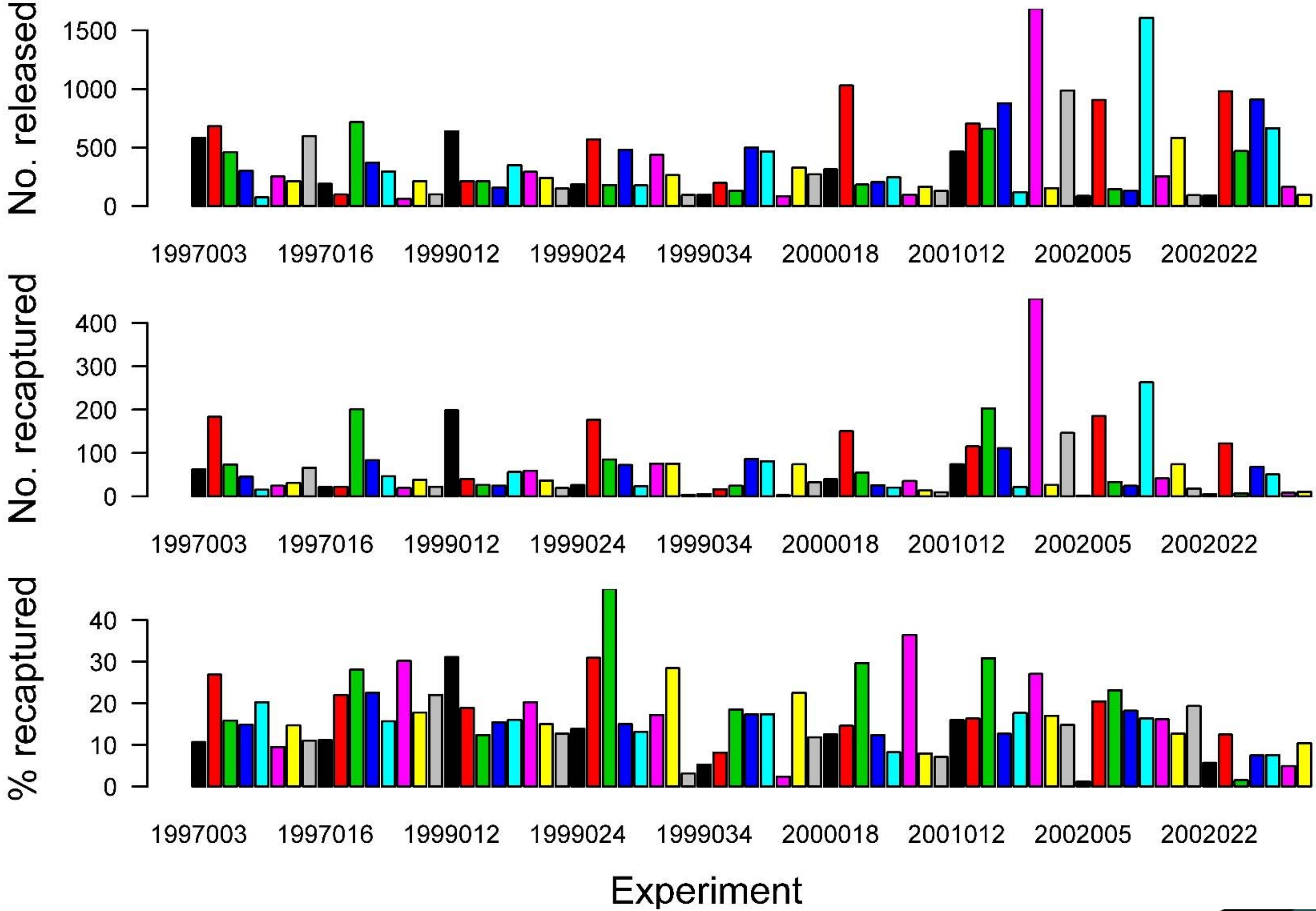




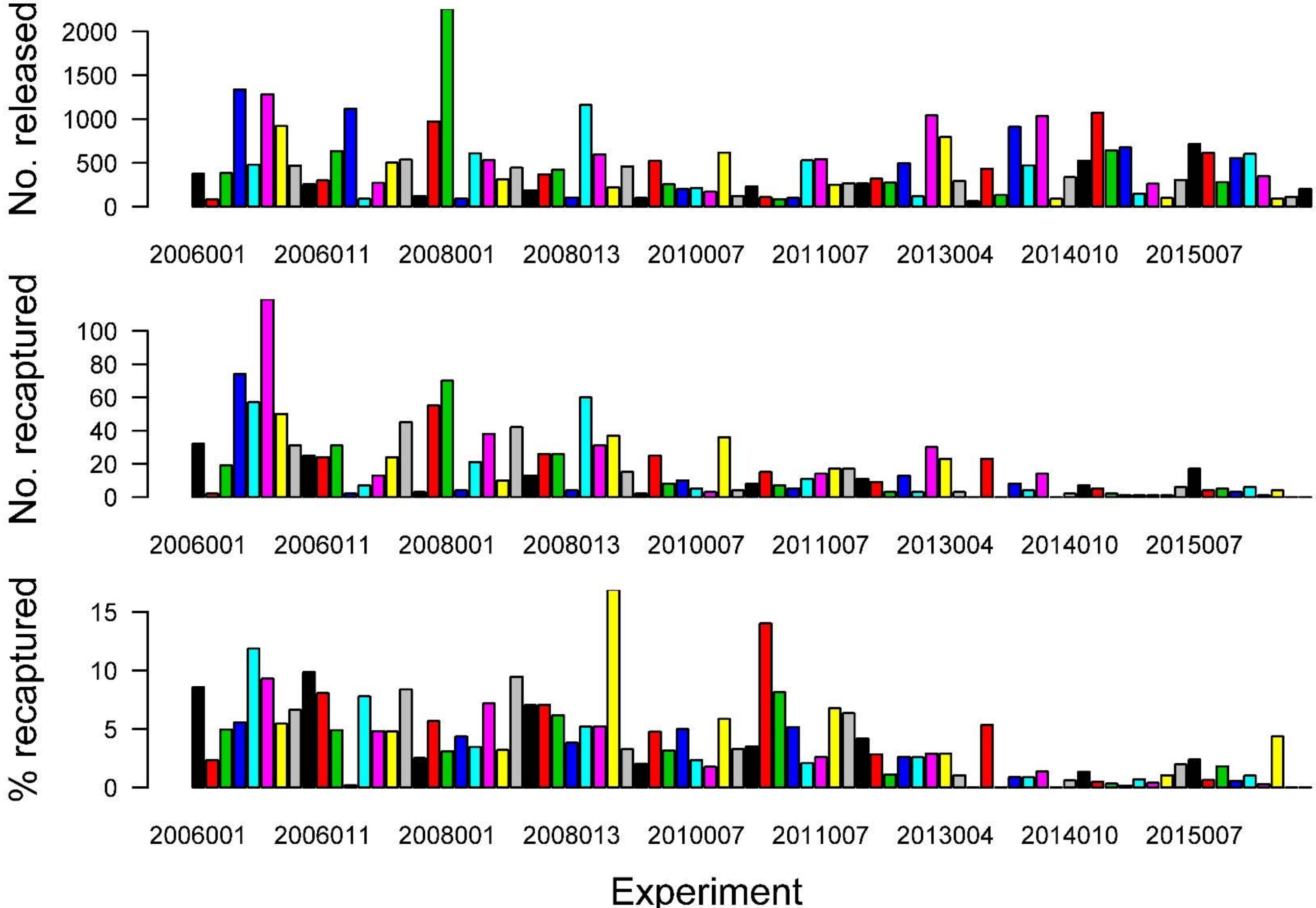
Tagging data selected for model



Tagging data selected for model



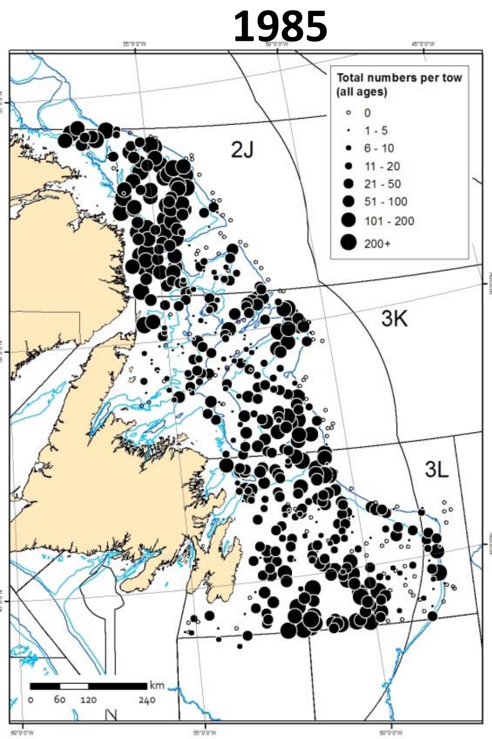
Tagging data selected for model

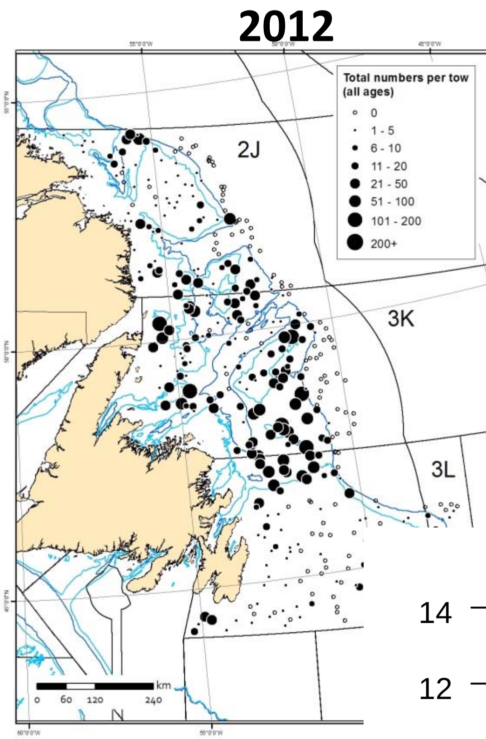


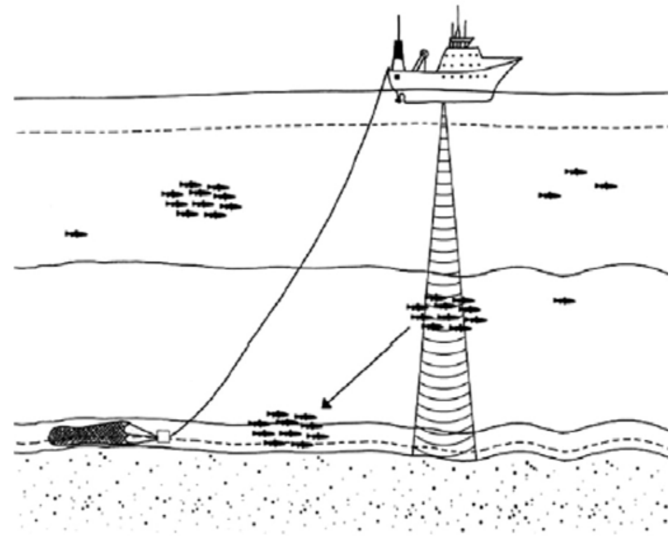
Monthly Landings

- NCAM uses the monthly distribution of landings over years to infer how much fishing mortality (*F*) to apply to tagging experiments in their release year
- The idea is to apply a fraction of *F*, depending on how much of the cumulative total annual catch has occurred before tagging
- Assume tagged fish released at beginning of a month and are immediately available to fishery

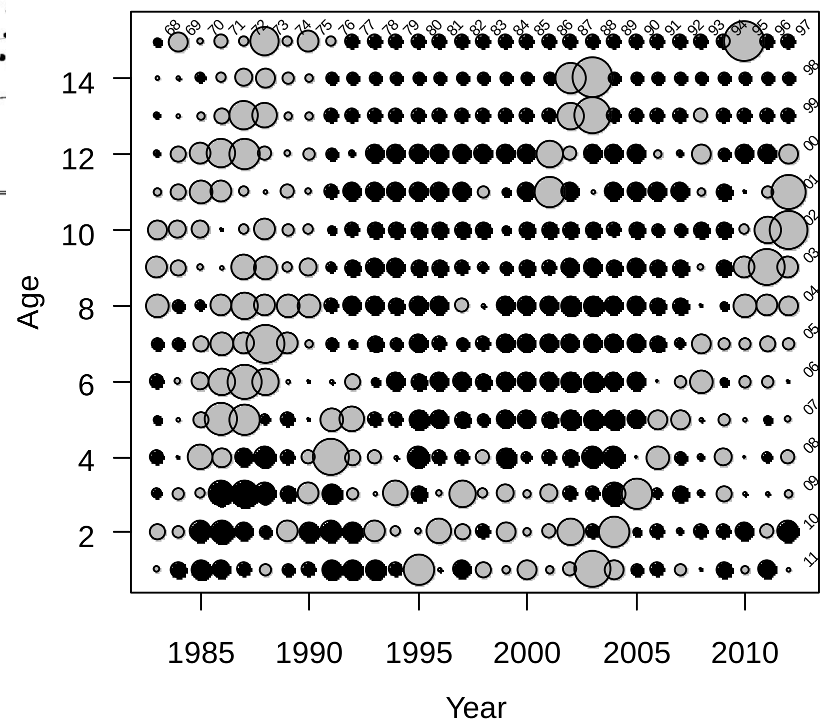
Data: Survey indices



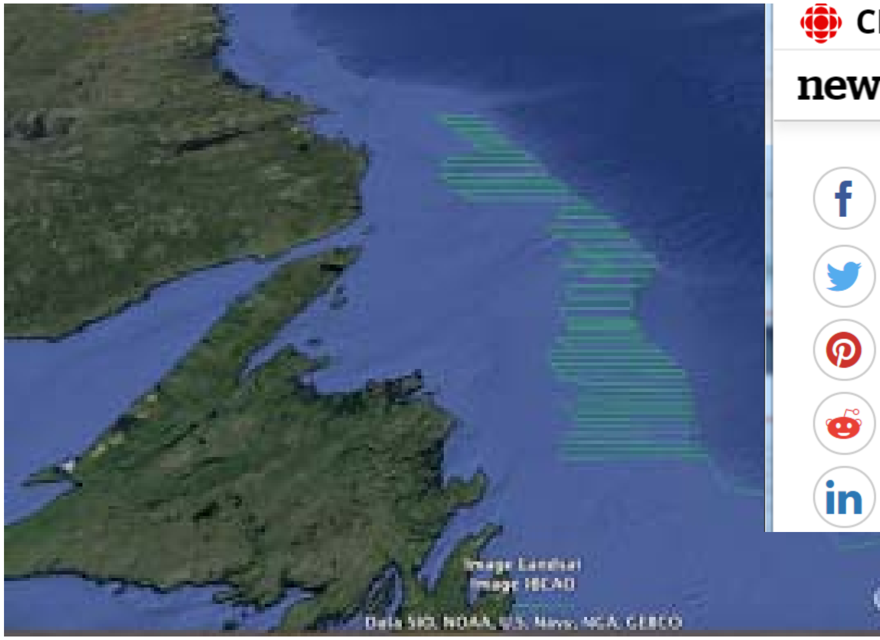




- Length at age
- Weight at age
- Relative Condition Index
- Relative Liver Index
- Age at maturity



Data Source 6: Acoustic Surveys



CBC | MENU

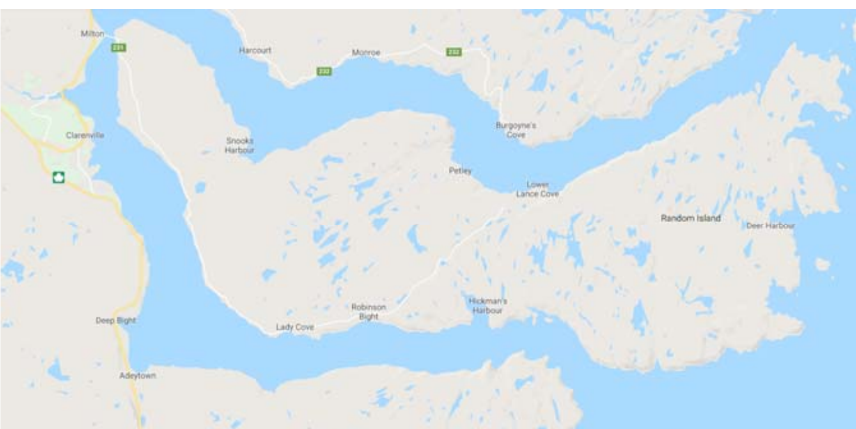
news Smith Sound cod disappearing

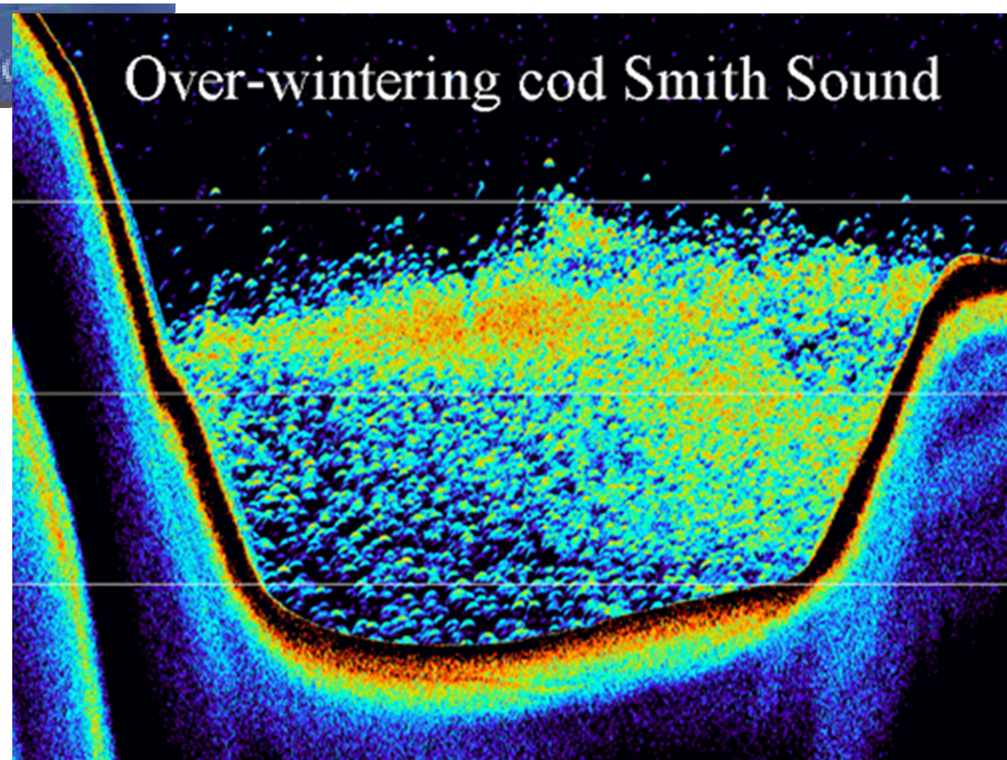
Smith Sound cod disappearing

CBC News · Posted: January 24th 2003, 9:57:00 AM | Last Updated: January 24, 2003

The preliminary results from a survey of the Smith Sound cod stock show the population has dropped for the second year in a row.

Two years ago, the stock in Trinity Bay, near Clarenville, was estimated at 25,000 tonnes. Now a survey this month indicates it's 16,000 tonnes.





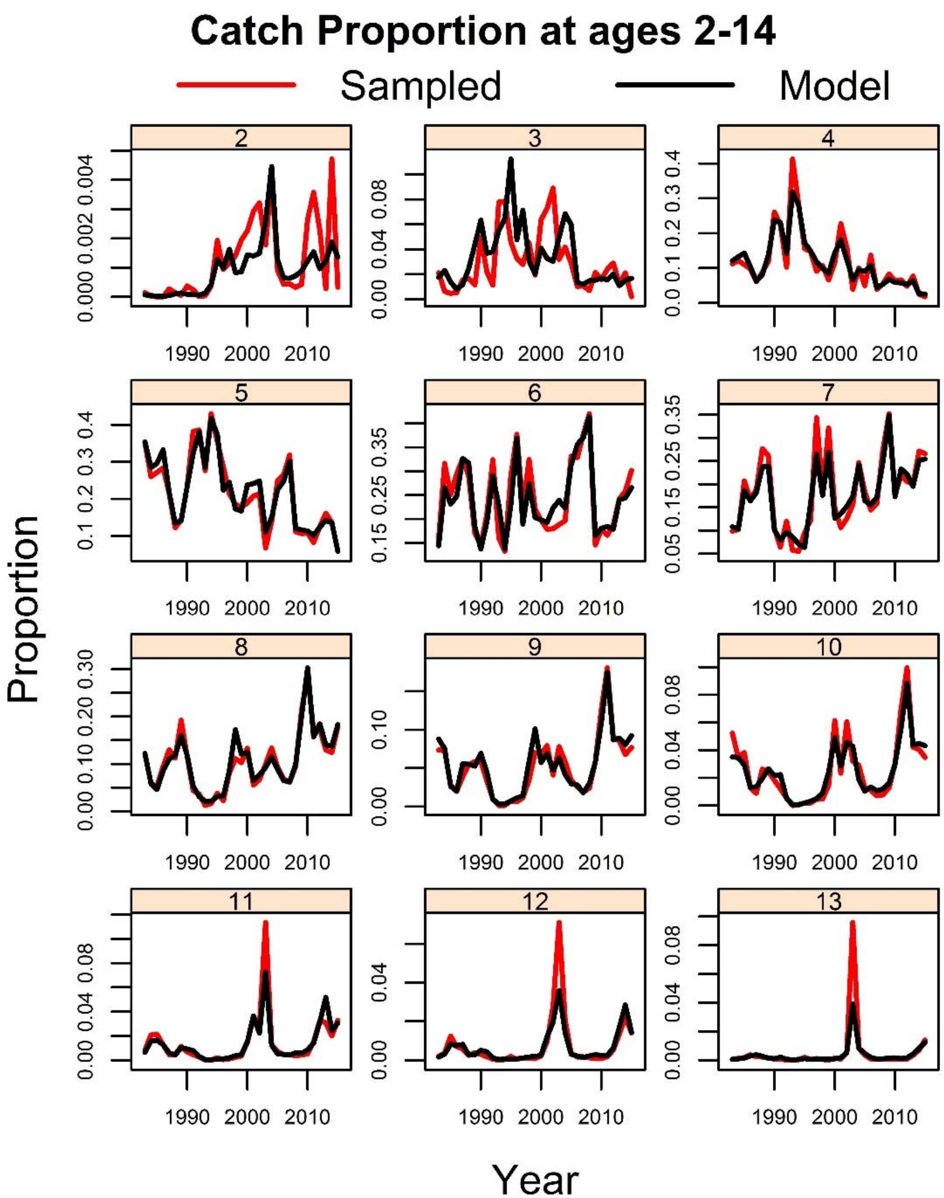
DFO RV survey q

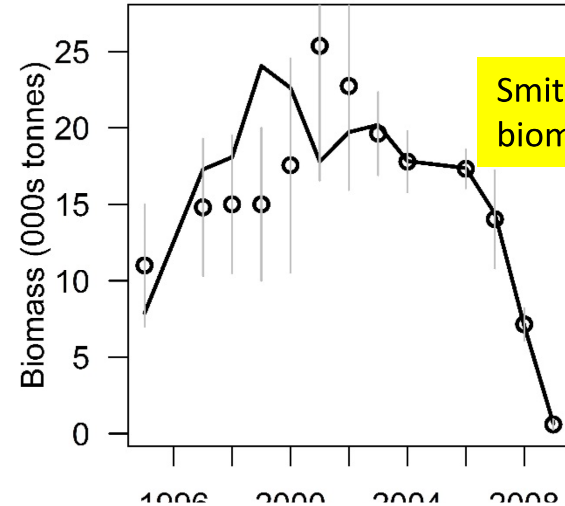
- DFO q's probably changed during 1995-2009 because of a shift in the over-wintering stock distribution:
- from mostly in the offshore to mostly in Smith Sound in the inshore.

$$E(I_{DFO,a,y}) = \begin{cases} \mathrm{q}_a \mathrm{N}_{a,y} \exp(-\mathrm{t}\mathrm{Z}_{a,y}), & y \neq 1995, \dots, 2009, \\ \mathrm{q}_a (\mathrm{N}_{a,y} - \mathrm{N}_{sa,y}) \exp(-\mathrm{t}\mathrm{Z}_{a,y}), & y = 1995, \dots, 2009. \end{cases}$$

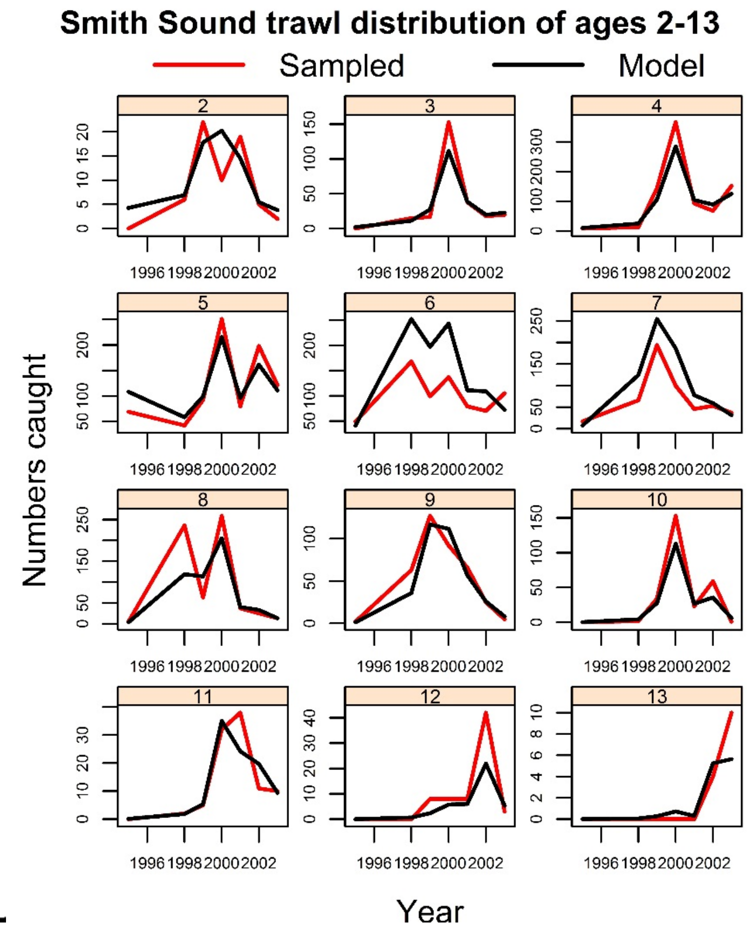
- $N_{sa,y}$ is the abundance of offshore fish that migrated and over-wintered in Smith Sound during 1995 to 2009

The Fit

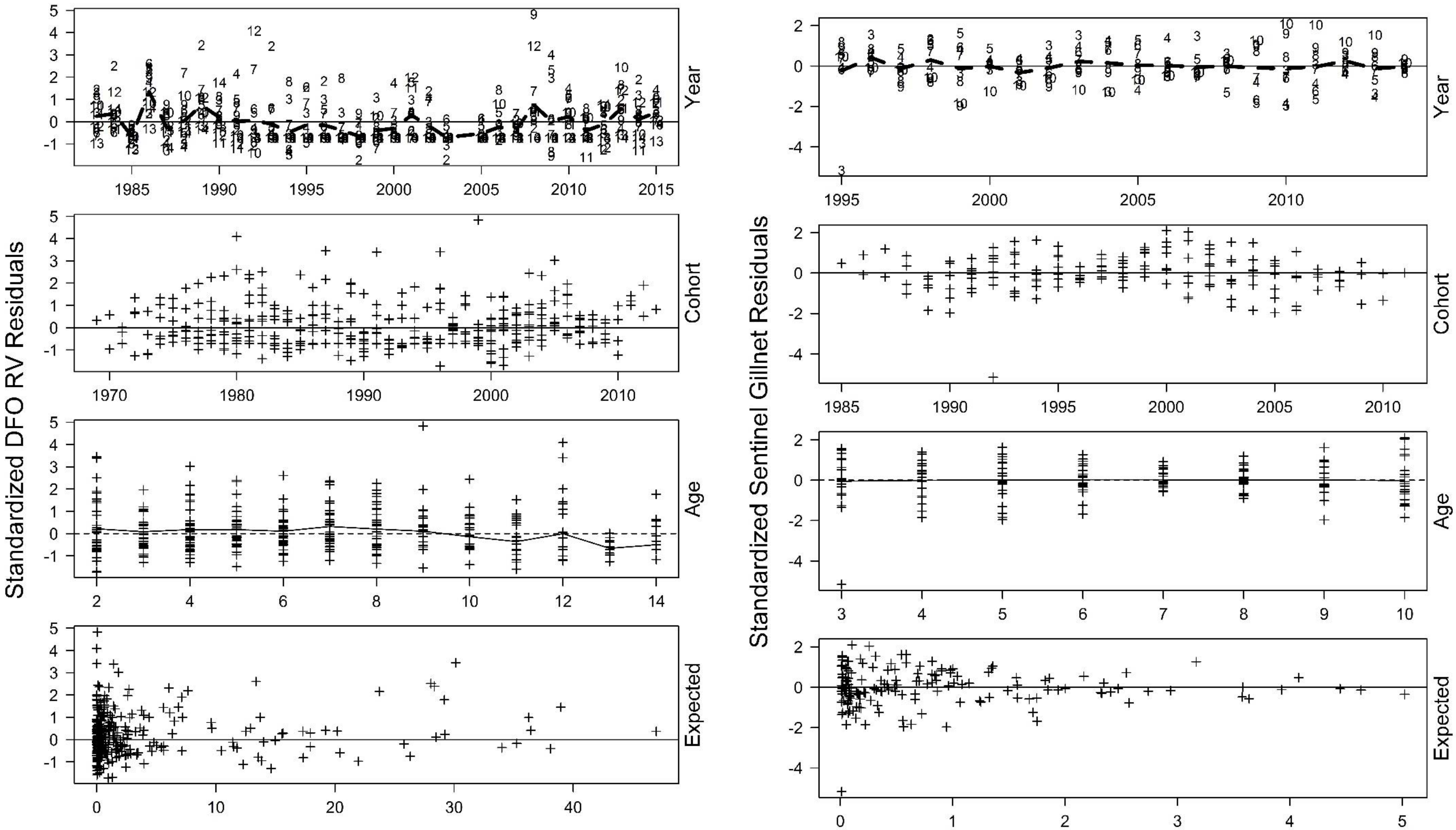




Smith Sound acoustic biomass estimates



The Fit



Standardized DFO RV Residuals
Year
Cohort
Age
Expected
Standardized Sentinel Gillnet Residuals
Year
Cohort
Age
Expected

Future NCAM research

- Include more tagging information / acoustic surveys +
- Re-think inshore Sentinel survey indices ?
- Include catches prior to 1983 +

DFO has done this

- Fishery timing data
- Include fleet composition information ?
- Include inshore juvenile surveys / age 1 RV index ?
- Spatio-temporal model +

Censored likelihood: what is it?

- EG. Assume $Y = \beta + \varepsilon$, where ε is error
- The likelihood is $L(\beta|y) = Pr(Y=y|\beta)$
- If we only know that y was in some interval y_l,y_u then L is still the same, $L(\beta|y_l,y_u) = Pr(y_l \leq Y \leq y_u|\beta)$
- This is an interval censored likelihood
- If we only know y_l (lower bound) or y_u (upper bound) then we can still form a L
- i.e. if only y_l, $L(\beta|y_l) = Pr(Y \geq y_l|\beta)$

Simple Censored Example

- $Y_1,\ldots,Y_n \sim N(10,\sigma^2)$, independent
- Observed data: we only know that $Y_i \in (L_i,U_i)$ – interval censored

```
n=1000
y = rnorm(n,10,10)
#case 1 - uncensored data
nll = function(parms){
  mu=parms[1]
  s = exp(parms[2])
  dn = dnorm(y,mu,s,log=TRUE)
  return(-sum(dn))
}
```

```
> fit = nlminb(c(10,5),nll)
> fit$par[1]
[1] 10.00658
> exp(fit$par[2])
[1] 10.45253
>
> mean(y)
[1] 10.00667
> sqrt((n-1)*var(y)/n)
[1] 10.4526
```

Simple Censored Example

```
#case 2 - censored data
yl = y - 4
yu = y + 4

nllc = function(parms){
  mu=parms[1]
  s = exp(parms[2])
  return(-
sum(log(pnorm(yu,mu,s)
  -pnorm(yl,mu,s))))
}
```

```
> fitc = nlminb(c(10,10),nllc)
> fitc$par[1]
[1] 10.00676
> exp(fitc$par[2])
[1] 10.194
```

```
> yl = y - 8
> yu = y + 8
>
> fitc = nlminb(c(10,10),nllc)
> fitc$par[1]
[1] 10.0081
> exp(fitc$par[2])
[1] 9.376591
```

```
#case 3 – censored but with N(10,1) !!
> n=1000
> y = rnorm(n,10,1)
> yl = y - 4
> yu = y + 4
> fitc = nlminb(c(10,10),nllc)
> fitc$par[1]
[1] 9.999999
> exp(fitc$par[2])
[1] 0.002478752
```

Censored log-likelihood for total catches

- Reported catch treated as a lower bound, and user sets an upper bound
- C_y is the total model predicted catch
- U_y is the upper bound
- $L_y = C_{oy}$ is the observed catch
- assume lognormal measurement error

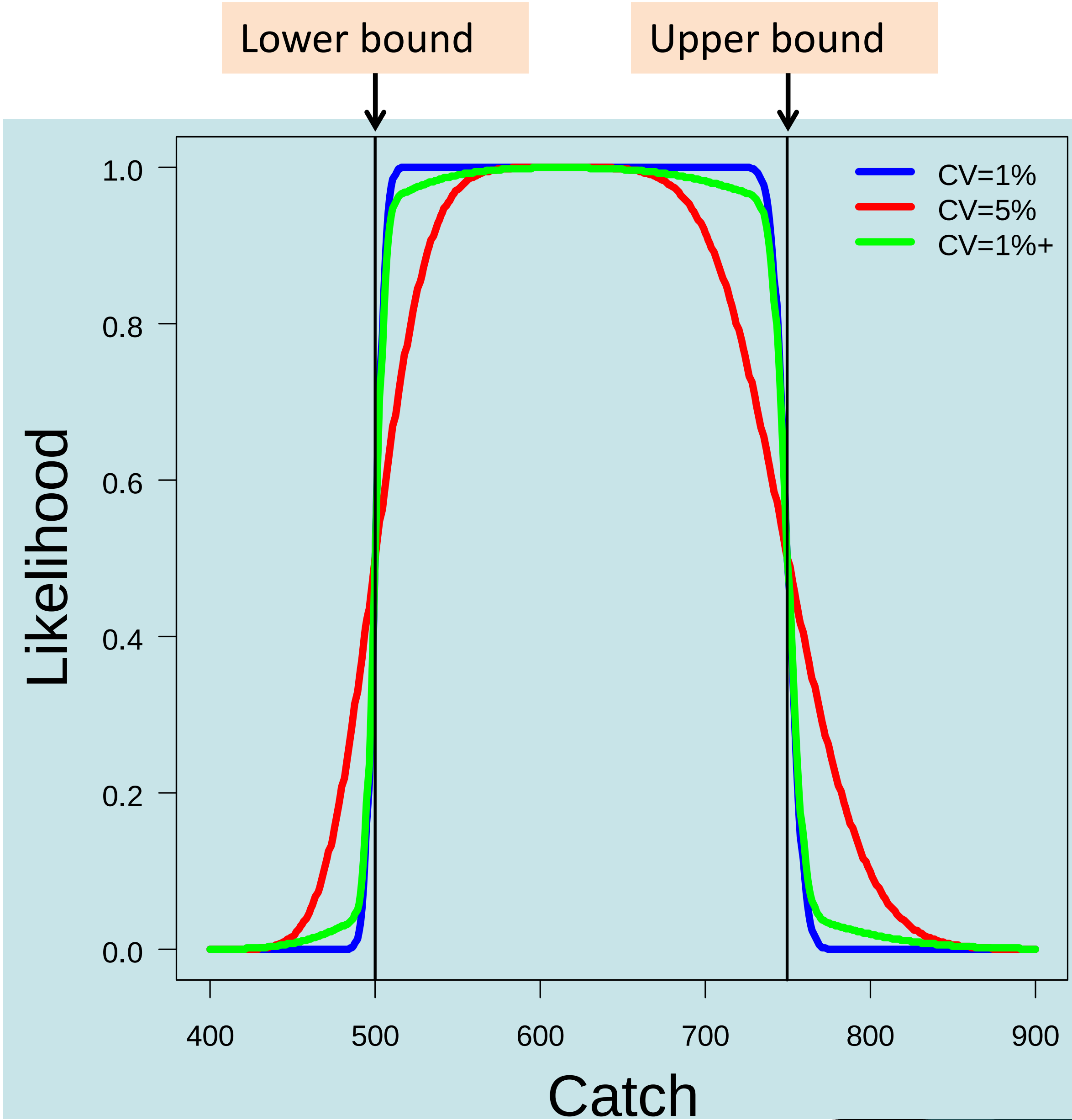
from Baranov catch equation

$$l(\theta|\{L,U\}|) = \sum_{y=1}^{Y} log\left[\Phi\left\{\frac{\log(U_y/C_y)}{\sigma_C}\right\} - \Phi\left\{\frac{\log(L_y/C_y)}{\sigma_C}\right\}\right]$$

$[\cdot]= \Pr(L_y \le C_y \le U_y | \theta)$

Assumed measurement error

Censored likelihood for catch, example



Lognormal likelihood for surveys

- Offshore DFO autumn RV survey indices for ages 2-14 and years 1983-2014.

- Inshore Sentinel gillnet indices for ages 3-10 and years 1995-2014.

- Model predicted survey catch: $E\left(I_{s,a,y}\right) = \mathrm{q}_{s,a}\mathrm{N}_{a,y}$

$\varphi_{\mathrm{N}}(.)$ - N(0,1) pdf

$$l(\{I_{s,a,y}\}|\theta) = \sum_a \sum_y log\left(\underbrace{\varphi_N\left[\frac{\log(I_{a,y}) - \log(\mathrm{q}_{s,a}\mathrm{N}_{a,y})}{\sigma_I}\right]/\sigma_I}_{\text{N(mean,std) pdf}}\right).$$

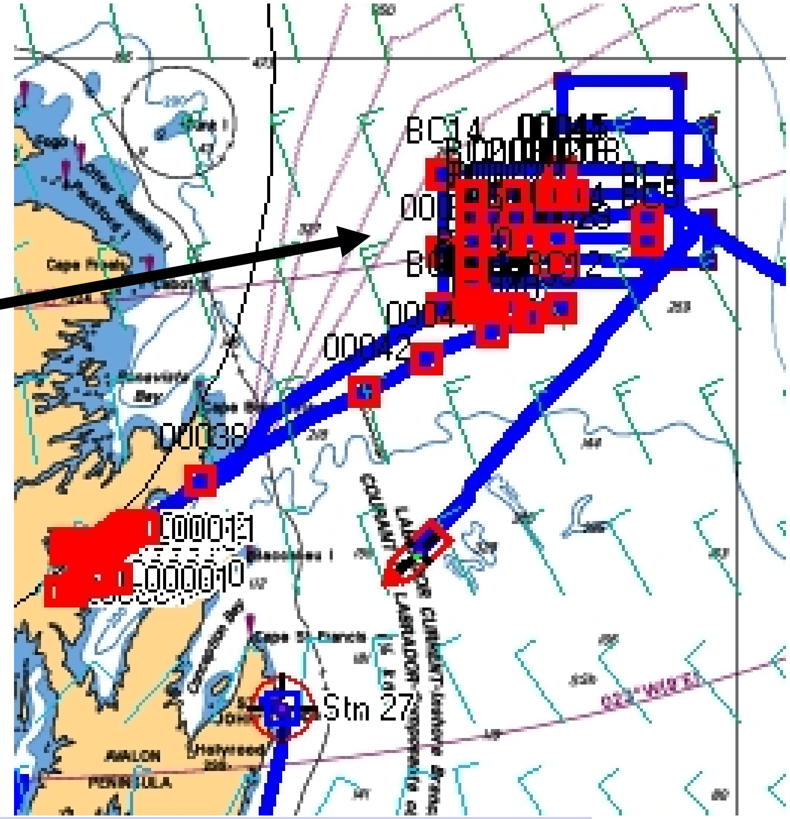
- Zero's included via a censored likelihood

Likelihood for surveys

- I assume there is a DFO RV survey detection limit of 0.005 mean numbers per tow.
- The likelihoods for zero indices are replaced by censored likelihoods, all with upper bound 0.005.
- Effect: estimation will not care how much less $E\left(I_{s,a,y}\right)$ is from the detection limit.
- This is very different than replacing a zero with small value.

Censored L for offshore acoustic biomass

How to use a 500 Kt offshore acoustic biomass estimate of some spawning aggregations in 2014




```
// Imaginary Acoustic survey nll;

 Type ZU = (log(500000.0) - log_ssb(2014))/0.15;
 nll -= log(one - pnorm(ZU));
```

Age Compositions

- Recall how we estimated catch-at-age from catch-at-length and an age-length key
- Reality is fisheries landings are estimated via some monitoring scheme that we do not understand well
- And science samples catch for length and age using a complicated sampling scheme that we also don't understand well
- Age sampling involves some ad hoc decisions because '*the best-laid plans of mice and men often go awry*'

Age Compositions

- SS3 uses the multinomial distribution with some ad hoc and fairly obscure methods to adjust the multinomial "effective" sample sizes to better "weight" data
- Thorson proposed the Dirichlet-multinomial distribution for age and length composition data, and I think that is now an option in SS3
- Francis argued that the multinomial or Dirichlet-multinomial distributions are not 'proper' for compositions

Francis, R.C., 2014. Replacing the multinomial in stock assessment models: A first step. Fisheries Research, 151, pp.70-84.

Thorson, J.T., Johnson, K.F., Methot, R.D. and Taylor, I.G., 2016. Model-based estimates of effective sample size in stock assessment models using the Dirichlet-multinomial distribution. *Fisheries Research*.

Albertsen, C.M., Nielsen, A. and Thygesen, U.H., 2016. Choosing the observational likelihood in state-space stock assessment models. *Canadian Journal of Fisheries and Aquatic Sciences*, *74*(5), pp.779-789.

Age Compositions

- In NCAM, bounds on total annual landings (derived from industry reps) are treated as a data source via a censored nll
- And annual catch proportion at age are treated as another data source
- But we don't know the sample sizes to determine these age comps

Compositional likelihood for catch age comps

- Practically impossible *for me* to evaluate from first principles the statistical properties of the Ncod catch age composition information – very complex sampling

- That has been poorly documented over time

- Francis (2014) reviewed this area and also studied some of the common approaches, using many data sets and simulation analyses

- He concluded that the **logistic normal multinomial (LNM)** distribution showed great promise for modelling compositional data in stock assessment.

Francis, R.C., 2014. Replacing the multinomial in stock assessment models: A first step. *Fisheries Research*, *151*, pp.70-84.

Additive LNM distribution

- Compositional data $P_1,\ldots,P_B$ for states $1,\ldots,B$ where $P_1+\ldots+P_B = 1$
- let $X_b = log\{P_b/(1-\sum_{i=1}^{B-1} P_i)\}$, $b = 1,\ldots,B-1$.
- The LNM distribution is based on assuming $X_1,\ldots,X_{B-1}$ has a multivariate normal distribution with some correlation structure.

$$P_b = \begin{cases} \dfrac{\exp(X_b)}{1+\sum_{i=1}^{B-1}\exp(X_i)}, & b = 1,\ldots,B-1. \\ \dfrac{1}{1+\sum_{i=1}^{B-1}\exp(X_i)}, & b = B. \end{cases}$$

Multiplicative LNM distribution

- Aitchison (2003) referred to this as the additive logistic transformation.
- He argued that for ordered compositional data, such as for ages and lengths, the multiplicative logistic transformation is more appropriate

$$P_b = \begin{cases} \dfrac{\exp(X_b)}{\prod_{i=1}^{b}\{1+\exp(X_b)\}}, & b = 1, \dots, B-1. \\ \dfrac{1}{\prod_{i=1}^{B-1}\{1+\exp(X_b)\}}, & b = B. \end{cases}$$

Multiplicative LNM distribution

- Let $\pi_b = P_b/(P_b + \ldots + P_B)$
- $\pi_b = Prob(State = b \,|\, State \geq b)$

$$X_b = log\left(\frac{\pi_b}{1-\pi_b}\right) = log\left(\frac{P_b}{P_{b+1}+,\ldots,+P_B}\right), b = 1, \ldots, B-1$$

- This is the continuation-ratio logit as defined by Agresti (1992)
- has been used for modelling length and age distributions by Kvist et al. (2000), Rindorf and Lewy (2001), Berg and Kristensen (2012), ….

Age composition log-likelihood

- In NCAM, errors in catch age compositions are assumed to be independent

$$\boldsymbol{X_{ay} = log\left(\frac{C_{ay}}{C_{a+1y}+,\dots,+C_{Ay}}\right)}$$

$$l(\{X_{oay}\}|\theta,\Psi) = \sum_{a=1}^{A-1}\sum_{y=1}^{Y} log\left\{\varphi\left(\frac{X_{oay}-X_{ay}}{\sigma_{Pa}}\right)/\sigma_{Pa}\right\}$$

$$\sigma_{Pa} = \begin{cases} 3\sigma_P, & a \le 2, \\ \sigma_{P,} & 3 \le a \le 7 \\ 2\sigma_P, & a \ge 8, \end{cases}$$

=> Ad hoc adjustments to account for different residual variations

- σ_P is estimated

Zero catches are an issue

Integrated Assessments

A review of integrated analysis in fisheries stock assessment. Maunder, M.N. and Punt, A.E., 2013. *Fisheries Research*, *142*, pp.61-74.

- Contemporary use of integrated analysis involves using <u>all</u> available data, in as raw a form as appropriate

- through <u>likelihood functions</u> that include multiple data sources

Likelihoods

- If data x are assumed to have some probability density function $f_X(x|\theta)$, that is known except for an unknown set of parameters θ
- The likelihood of θ is $L_X(\theta|x) = f_X(x|\theta)$
- Integrated: If two independent data sources X and Y are available, $L_{X,Y}(\theta|x,y) = f_X(x|\theta)f_Y(y|\theta)$
- $l_{X,Y}(\theta|x,y) = log\{L_{X,Y}(\theta|x,y)\}$
  $= log\{f_X(x|\theta)\} + log\{f_Y(y|\theta)\}$

Future: Integrated/Spatial. Why?

- There are important spatial dimensions to stock productivity and life-history processes that are currently not adequately accounted for in many stock assessments
- Increasingly fisheries managers and industry are asking for spatial advice (e.g. impacts of seasonal spawning closures, areal quotas).
- Often there are important spatial dimensions to stock assessment data (e.g. inshore/offshore surveys) that are better accommodated in a spatial model.

Future: Integrated/Spatial

- Hence, state-of-the-art integrative assessment models will need to be spatial in the future.
- This may be simple, such as including broad scale spatial dimensions in survey catchability (e.g. Ncod, Ghal)
- Less simple meta-substock models
- More complex spatial gridded approach with movements etc.
- 2018 CAPAM WS: Spatio-temporal modelling Mini-Workshop

Future: Meta-Stock Models

- the "Robin Hood" approach (Punt et al 2011)
- Modelling several stocks of the same or similar species simultaneously, with some shared parameters for different stocks
- Similar to using meta-data priors
- Mixed fishery models
- Ecosystem models

Albertsen, C.M., Nielsen, A. and Thygesen, U.H., 2017. Connecting single-stock assessment models through correlated survival. *ICES J Mar Sci*.

Cadigan, N.G. and Campana, S.E., 2017. Hierarchical model-based estimation of population growth curves for redfish (*Sebastes mentella* and *Sebastes fasciatus*) off the Eastern coast of Canada. *ICES J Mar Sci*, *74*, pp.687-697.

Punt, A.E., Smith, D.C. and Smith, A.D., 2011. Among-stock comparisons for improving stock assessments of data-poor stocks: the "Robin Hood" approach. *ICES J Mar Sci*, *68*, pp.972-981.

Thorson, J.T., Stewart, I.J., Taylor, I.G. and Punt, A.E., 2013. Using a recruitment-linked multispecies stock assessment model to estimate common trends in recruitment for US West Coast groundfishes. Marine Ecology Progress Series, 483: 245-256.

Statistical Catch-at-Age (SCA): 3NO cod

- An **illustrative** model of how this works
- $E\left(R_{say}\right) = q_{sa}N_{ay}\exp(-f_{sy}Z_{ay})$ – the usual survey observation model
- Use a forward cohort model, like SURBA (not backwards like ADAPT)
- Lets assume a separable F model, $F_{ay} = s_a f_y$ and $Z_{ay} = F_{ay} + M$ where M is known
- Model $log(f_y)$ as a smooth spline function of y (just a linear model of some basis function values)
- Model s_a as a simple parametric (exponential) function of a, $log(s_a) = \beta_1(a\text{-}6) + \beta_2|a\text{-}6|$.

Statistical Catch-at-Age (SCA): 3NO cod

- Treat catch as observations measure with error
- Catch observation model:

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

- The additive δ catch model errors is a simple but maybe not a good approach – lets see
- Model predicted catch is

$$E\left(C_{ay}\right) \cong N_{ay}\left(1 - e^{-Z_{ay}}\right)\frac{F_{ay}}{M_a + F_{ay}}$$

- A mesh between SURBA and ADAPT

3NO cod simple SCA

Substantial R code to read and plot data inputs

```
> load('data\\3LNO.RData')
> ls()
[1] "catch"    "catch.vec" "indices"   "mat"      "mat.vec"   "wt"       "wt.vec"
```

```
# get log indices and wt=0 for index=0
indices$YC = indices$Year-indices$Age
#set fraction of year survey takes palce
indices$fs=NA
indices$fs[indices$survey=="Juvenile"] = 8.5/12
indices$fs[indices$survey=="Spring"] = 5.5/12
indices$fs[indices$survey=="Fall"] = 10.5/12

indices$log.index = NA
indices$wt = 0
ind = indices$index==0
indices$wt[!ind]=1
indices$log.index[!ind] = log(indices$index[!ind])
```

Same as in R ADAPT

3NO cod simple SCA

```
pop.dat = list(
  catch=catch,
  mat=mat,
  wt=wt,
  Year=sort(unique(wt.vec$Year)),
  Age=sort(unique(wt.vec$Age))
)
pop.dat$M = 0.2
pop.dat$A = length(pop.dat$Age)
pop.dat$Y = length(pop.dat$Year)
pop.dat$n = length(wt.vec$Age)
pop.dat$imap = 1:pop.dat$n
pop.dat$age_diff = pop.dat$Age-6
pop.dat$abs_age_diff = abs(pop.dat$Age-6)
pop.dat$veccatch = as.vector(as.matrix(pop.dat$catch))
pop.dat$sqrt_veccatch = sqrt(pop.dat$veccatch)
```

Almost same in R ADAPT

3NO cod simple SCA

```
indices.nz = subset(indices,wt==1); #use data with estimation wt=1

temp1 = paste(wt.vec$Year,"_",wt.vec$Age,sep='')
temp2 = paste(indices.nz$Year,"_",indices.nz$Age,sep='')
indices.nz$imap = pop.dat$imap[match(temp2,temp1)]

#define q parmeters and map
indices.nz$qparm = paste(indices.nz$survey,"_",indices.nz$Age,sep='')
q.name = unique(indices.nz$qparm)
nq = length(q.name)
indices.nz$qmap = (1:nq)[match(indices.nz$qparm,q.name)]
```

Almost same in R ADAPT

```
#define index std parmseters and map

indices.nz$stdparm = indices.nz$survey; ##self-weight by survey
std.name = unique(indices.nz$stdparm)
nstd = length(std.name)
indices.nz$stdmap = (1:nstd)[match(indices.nz$stdparm,std.name)]
```

3NO cod simple SCA

```
#For now cheat and take starting values from assessment

qfall = c(11,11,8,7,5,4,4,3,3)/10
qspring = c(10,14,7,5,3,3,3,3,4)/10
qjuvenile = c(35,18,13,11,8,6,5,3,3)/10
q.start = c(qfall,qspring,qjuvenile)
names(q.start)=q.name
```

```
std.start=c(0.3,0.3,0.3,0.1)
names(std.start) = c(std.name,'Catch')
```

```
Nfirst = c(53067,92911,19327,16484,12049,4268,3076,3217,2287,324)/1000
No = c(63623,98989,130098,94606,135041,195489,252970,221171,121541,154111,96818,
101649,74517,42189,44126,27764,32970,54572,50070,20911,23722,33074,26374,42559,49825,
39693,10693,7819,15588,15505,6207,6865,24684,7868,801,505,969,1342,466,2809,5975,5554,
2168,990,889,1704,4661,4405,8111,13375,2704,5508,4393,1497,2174,1411,2400,5503)/1000
sparm=0
```

```
## spline basis functions for f)year
Xf = bs(pop.dat$Year[1:(pop.dat$Y-1)],knots=seq(1960,2015,by=3))
Xf = cbind(rep(1,nrow(Xf)),Xf)
pop.dat$Xf = Xf

fbeta = as.vector(c(log(0.2),rep(0,ncol(Xf)-1)))
f_year = exp(Xf%*%fbeta)
```

3NO cod simple SCA

```
Npop = function(Nfirst,No,sparm,fbeta,pop.dat){

  A=pop.dat$A
  Y=pop.dat$Y
  M=pop.dat$M

  s = exp(sparm[1]*pop.dat$age_diff +
       sparm[2]*pop.dat$abs_age_diff)
  f_year = exp(pop.dat$Xf%*%fbeta)
  F = f_year %*% s
  Z = F + M;

  N = matrix(NA,nrow=Y,ncol=A)
  all.age = 1:A
  next.age = 2:A
  N[1:(Y-1),1] = No
  N[Y,1] = exp(mean(log(N[(Y-3):(Y-1),1])))
  N[1,next.age] = Nfirst
```

```
  for (y in seq(2,Y)){
    for (a in next.age){
      N[y,a] <- N[y-1,a-1]*exp(-Z[y-1,a-1])}
  }
  C = F*(1-exp(-Z))*N[1:(Y-1),]/Z
  B = pop.dat$wt*N
  SSB = pop.dat$mat*B
  pop = list(N=N,C=C,B=B,SSB=SSB,F=F)
  return(pop)
}
```

3NO cod simple SCA

```
create.parm = function(x){
 pname = names(x)
 ret = list(
     Nfirst = exp(x[substr(pname,1,5)=="logNf"]),
     No=exp(x[substr(pname,1,5)=="logNo"]),
     sparm=x[substr(pname,1,5)=="sparm"],
     fbeta=as.vector(x[substr(pname,1,5)=="fbeta"]),
     q=exp(x[substr(pname,1,4)=="logq"]),
     std=exp(x[substr(pname,1,6)=="logstd"]))
 return(ret)
```

```
fit = function(parms,pop.dat,indices.nz){
  x = create.parm(parms)
  pop = Npop(x$Nfirst,x$No,x$sparm,x$fbeta,pop.dat)
  N=pop$N
  vecN = as.vector(N);
  ## make sure this works right, and years and ages are vec the same as wt.vec
  Elog_index = log(x$q[indices.nz$qmap]) + log(vecN[indices.nz$imap])
  std_log_index = x$std[indices.nz$stdmap]
  nll = -sum(dnorm(indices.nz$log.index,Elog_index,std_log_index,log=TRUE))
  nll = nll - sum(dnorm(pop.dat$sqrt_veccatch,as.vector(sqrt(pop$C)),x$std[4],log=TRUE))
return(nll)
}
```

nll = nll(log_indices) + nll($\text{catches}^{1/2}$)

3NO cod simple SCA

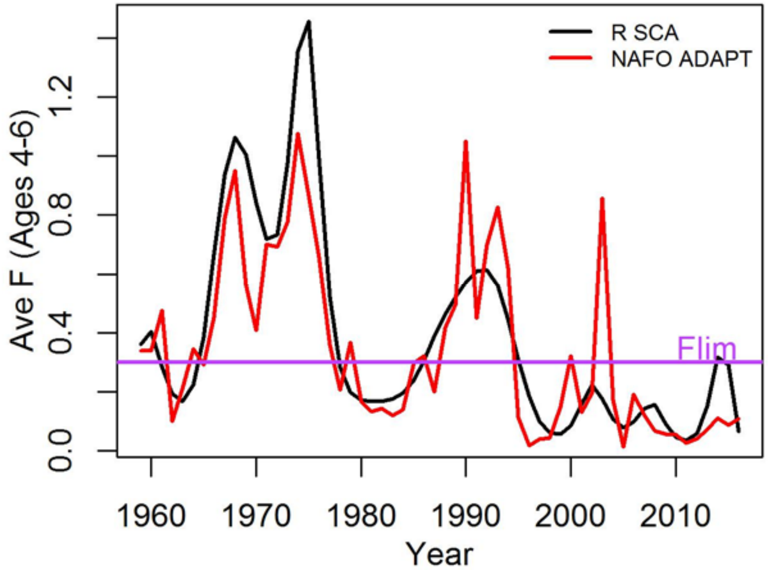
```
start.parms.list = list(
  logNfirst = log(Nfirst),
  logNo = log(No),
  sparm = sparm,
  fbeta = fbeta,
  logq = log(q.start),
  logstd = std.start
)
start.parms = unlist(start.parms.list)
```

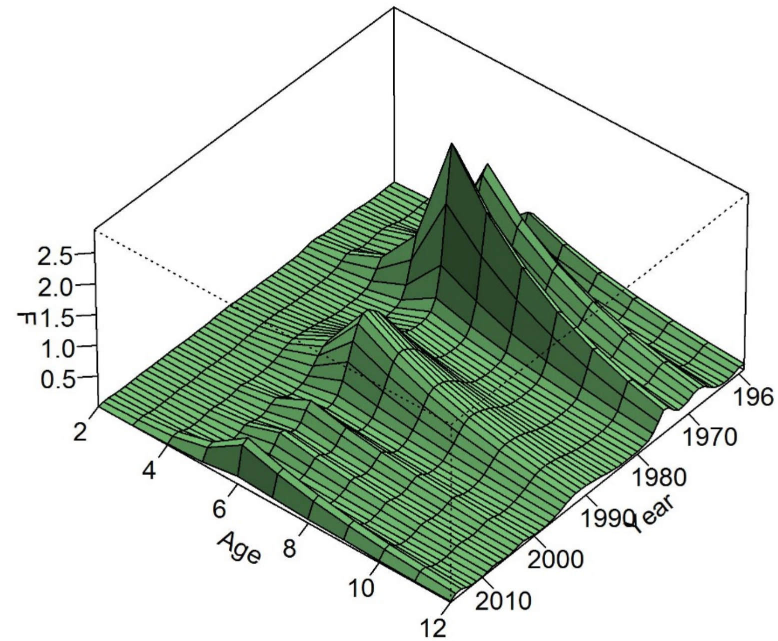
```
lower = list(
  logNfirst = rep(-10,length(start.parms.list$logNfirst)),
  logNo = rep(-10,length(start.parms.list$logNo)),
  sparm = rep(-10,length(start.parms.list$sparm)),
  fbeta = rep(-100,length(start.parms.list$fbeta)),
  logq = rep(-10,length(start.parms.list$logq)),
  logstd = log(rep(0.01,length(start.parms.list$std)))
)

upper = list(
  logNfirst = rep(100,length(start.parms.list$logNfirst)),
  logNo = rep(100,length(start.parms.list$logNo)),
  sparm = rep(100,length(start.parms.list$sparm)),
  fbeta = rep(100,length(start.parms.list$fbeta)),
  logq = rep(10,length(start.parms.list$logq)),
  logstd = log(rep(10,length(start.parms.list$std)))
)
lower=unlist(lower)
upper=unlist(upper)
```

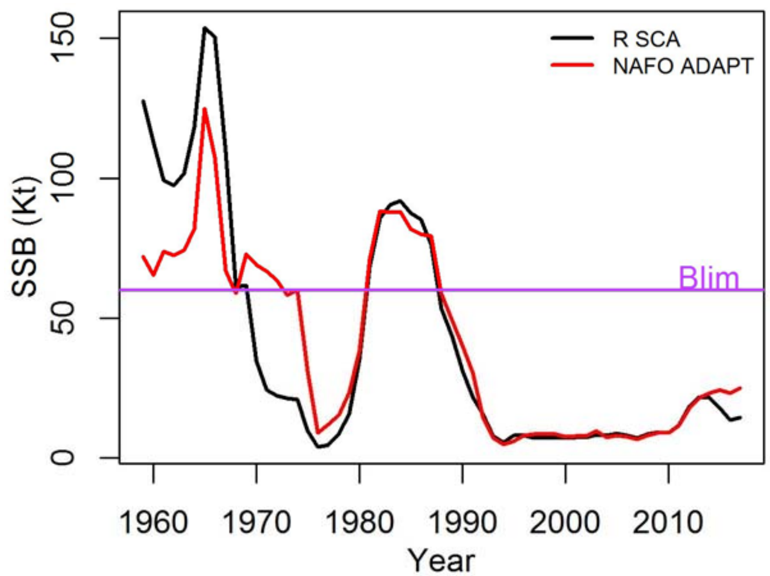
3NO cod simple SCA

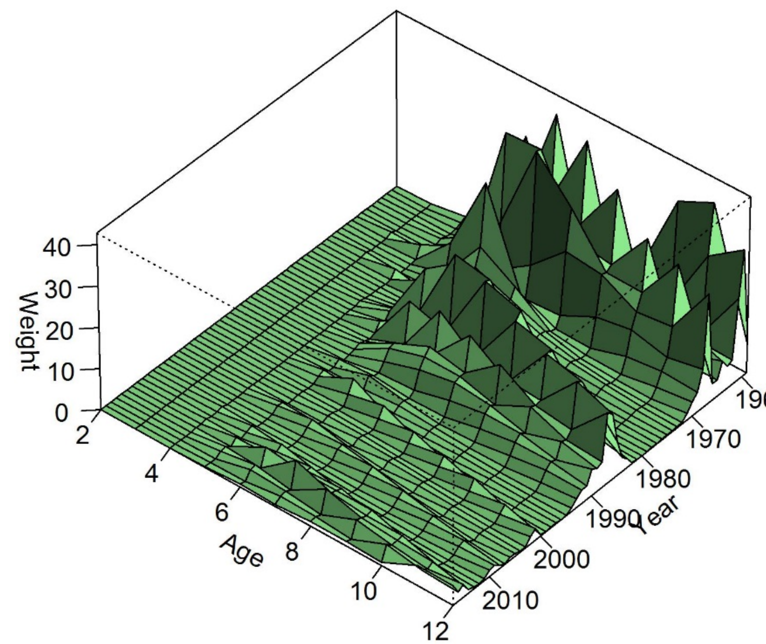
```
model.fit <- nlminb(start.parms,fit,,,pop.dat,indices.nz,
control=list(eval.max=10000,iter.max=1000),lower=lower,upper=upper)

## check gradients
grad(fit,model.fit$par,,,,pop.dat,indices.nz)

##get hessian
Hfit = hessian(fit,model.fit$par,,,pop.dat,indices.nz)
```

3NO cod SCA results

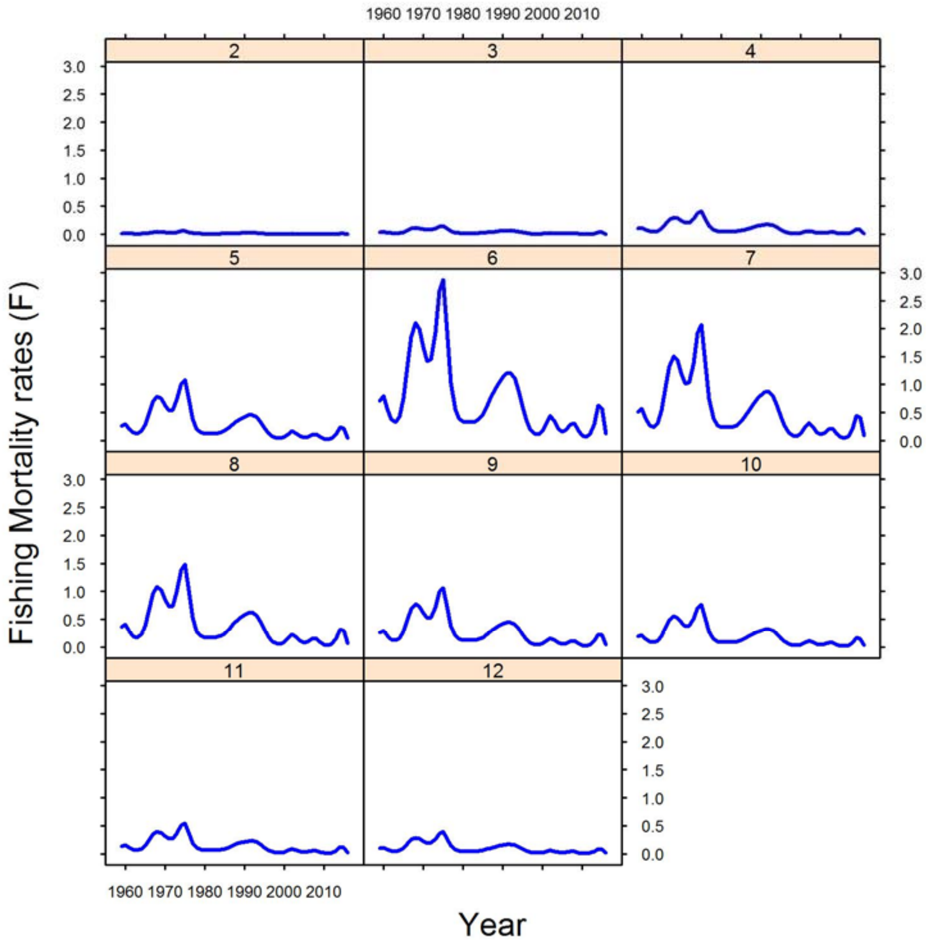


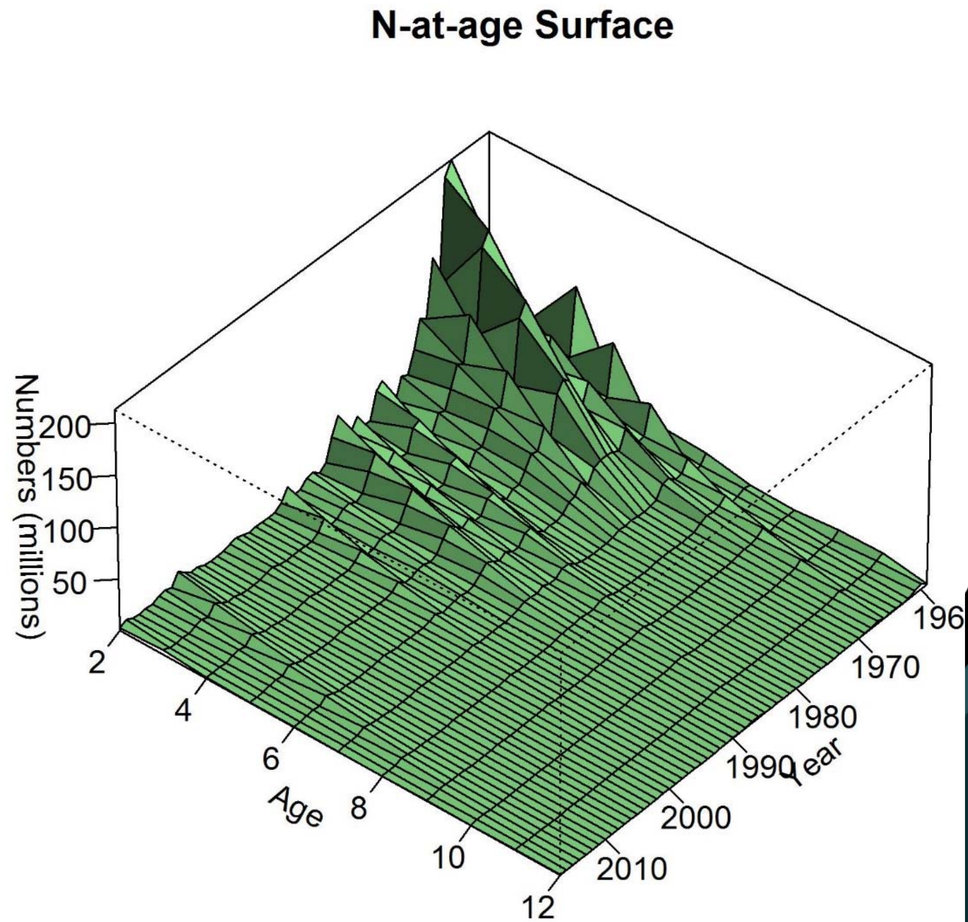




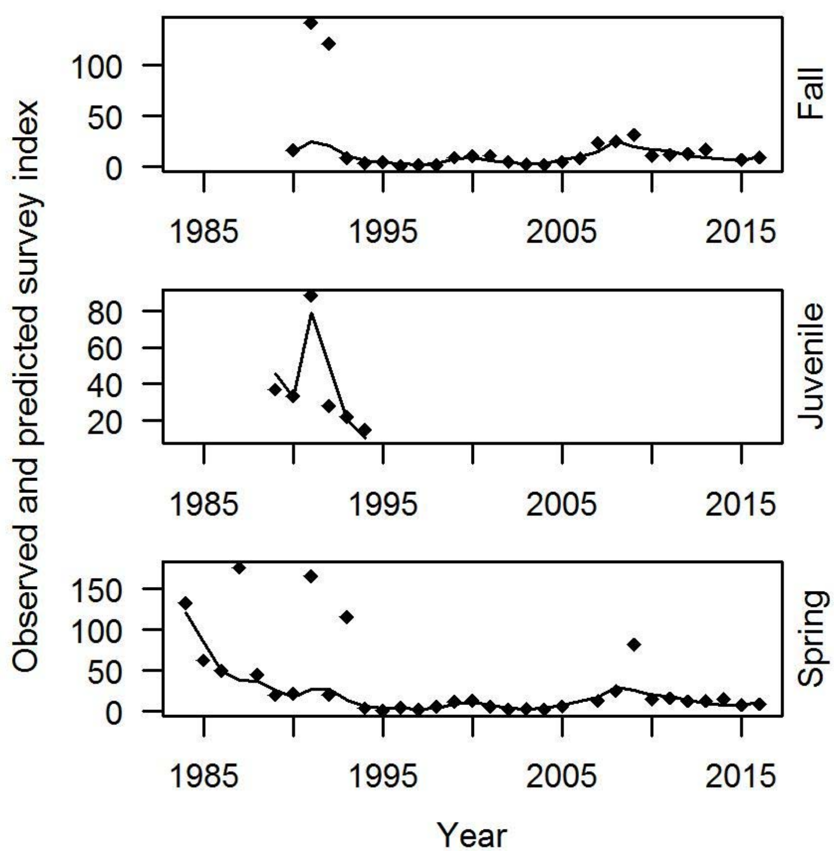


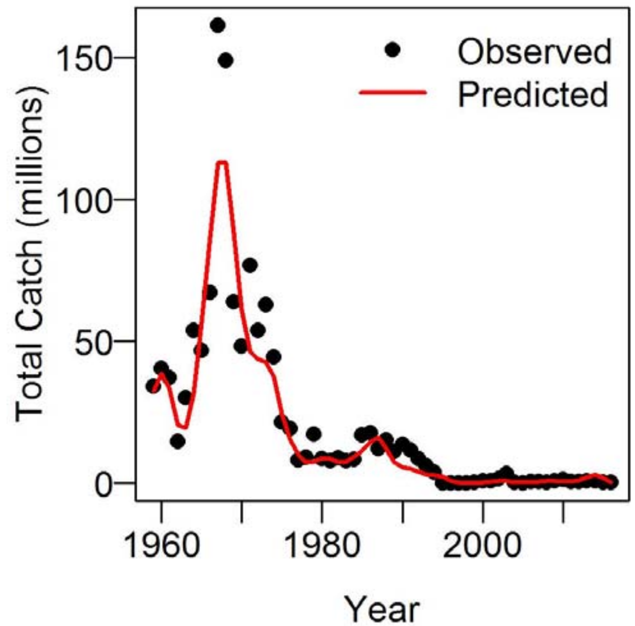
3NO cod SCA results

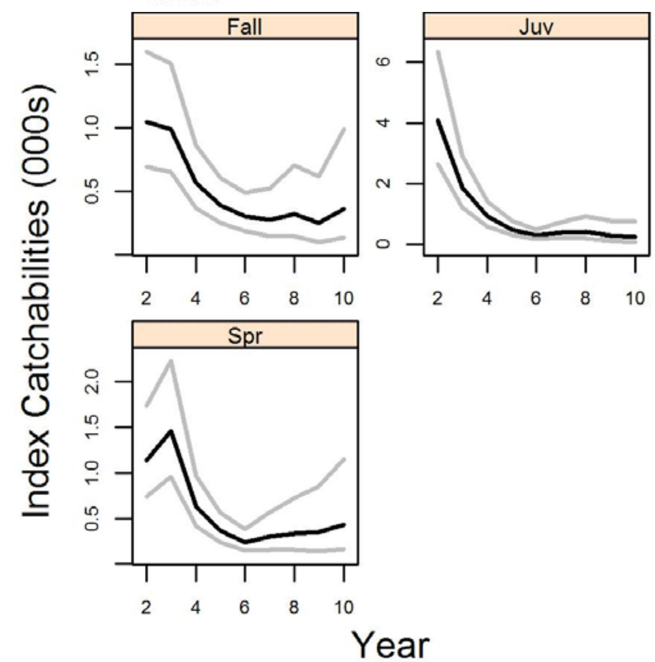




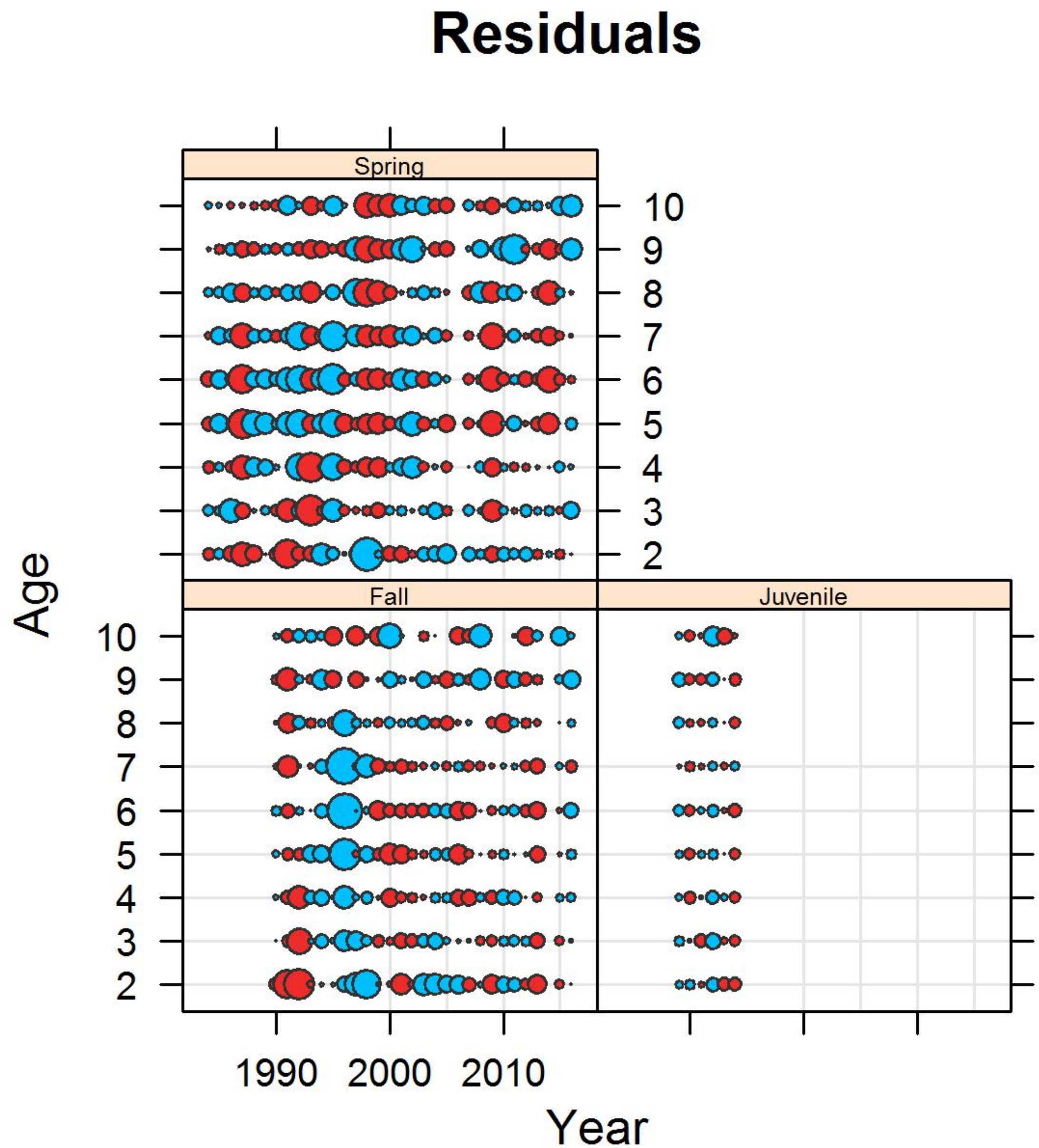
3NO cod SCA fits



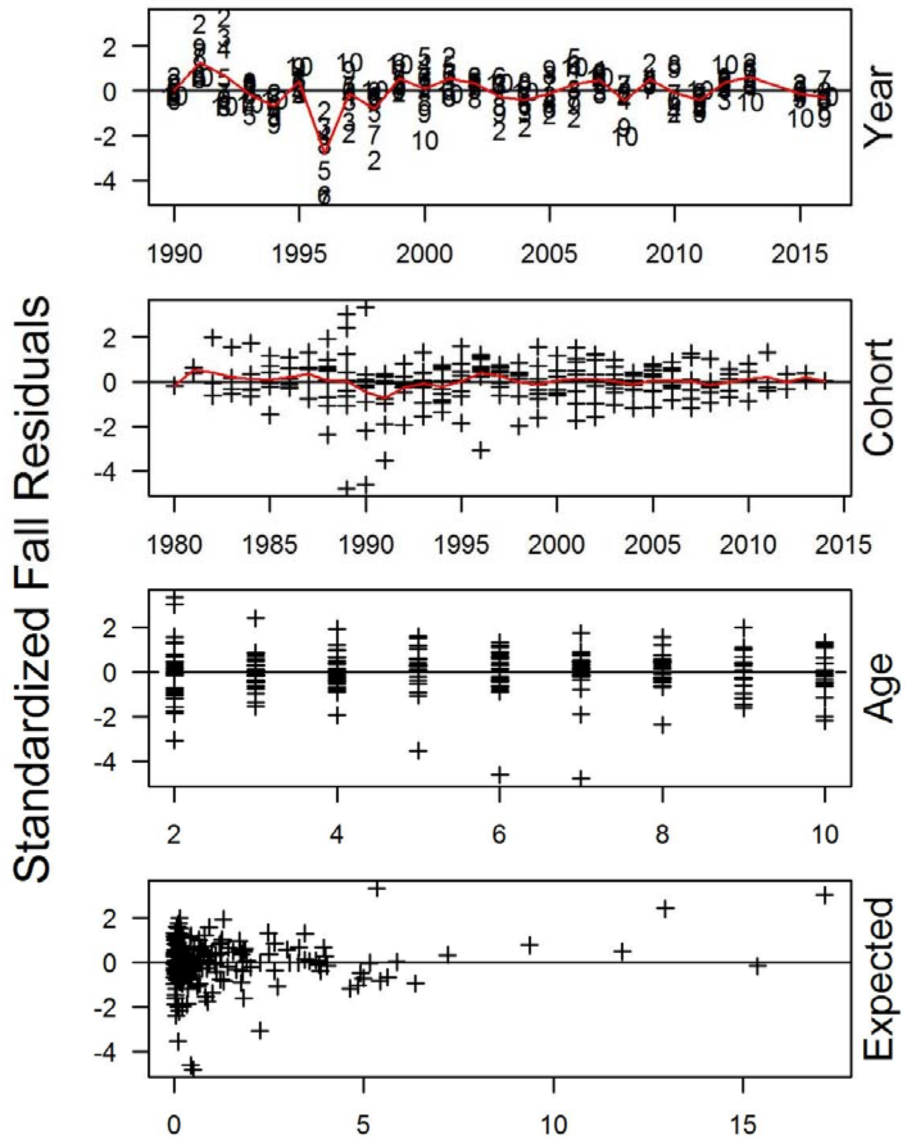


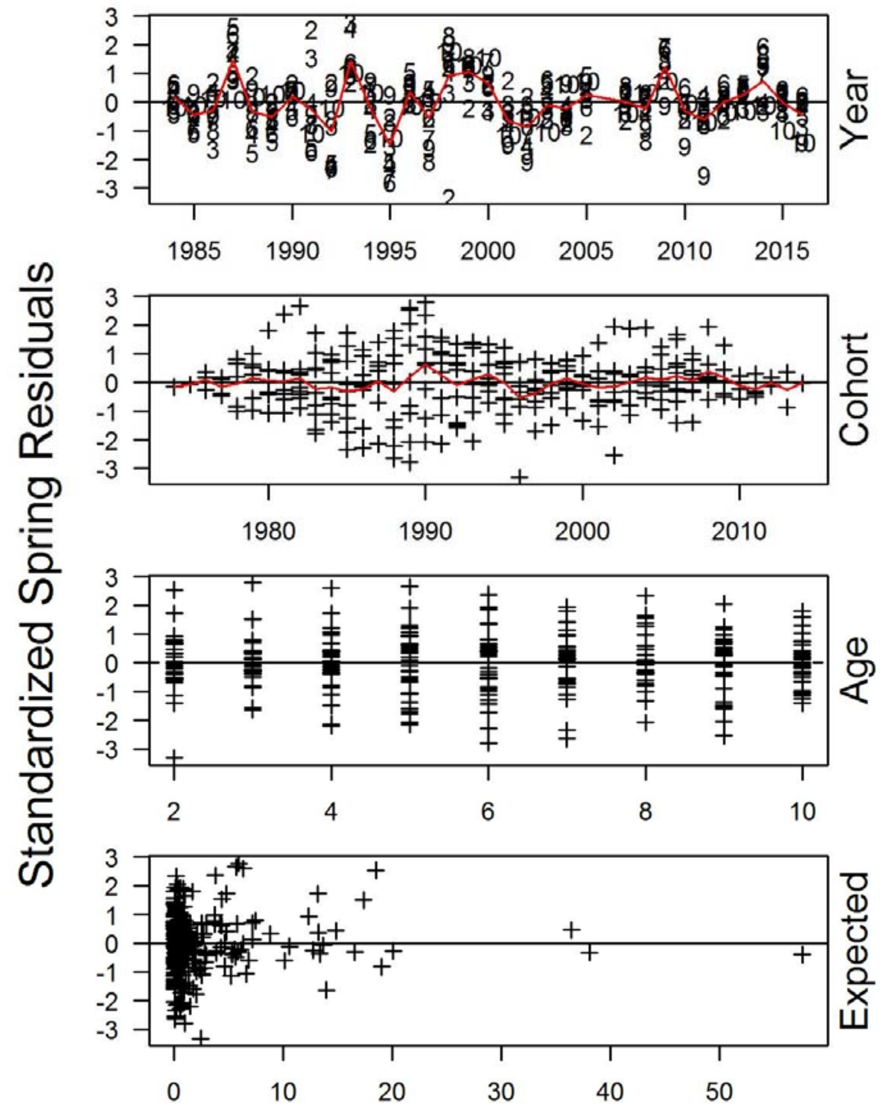


3NO cod SCA fits

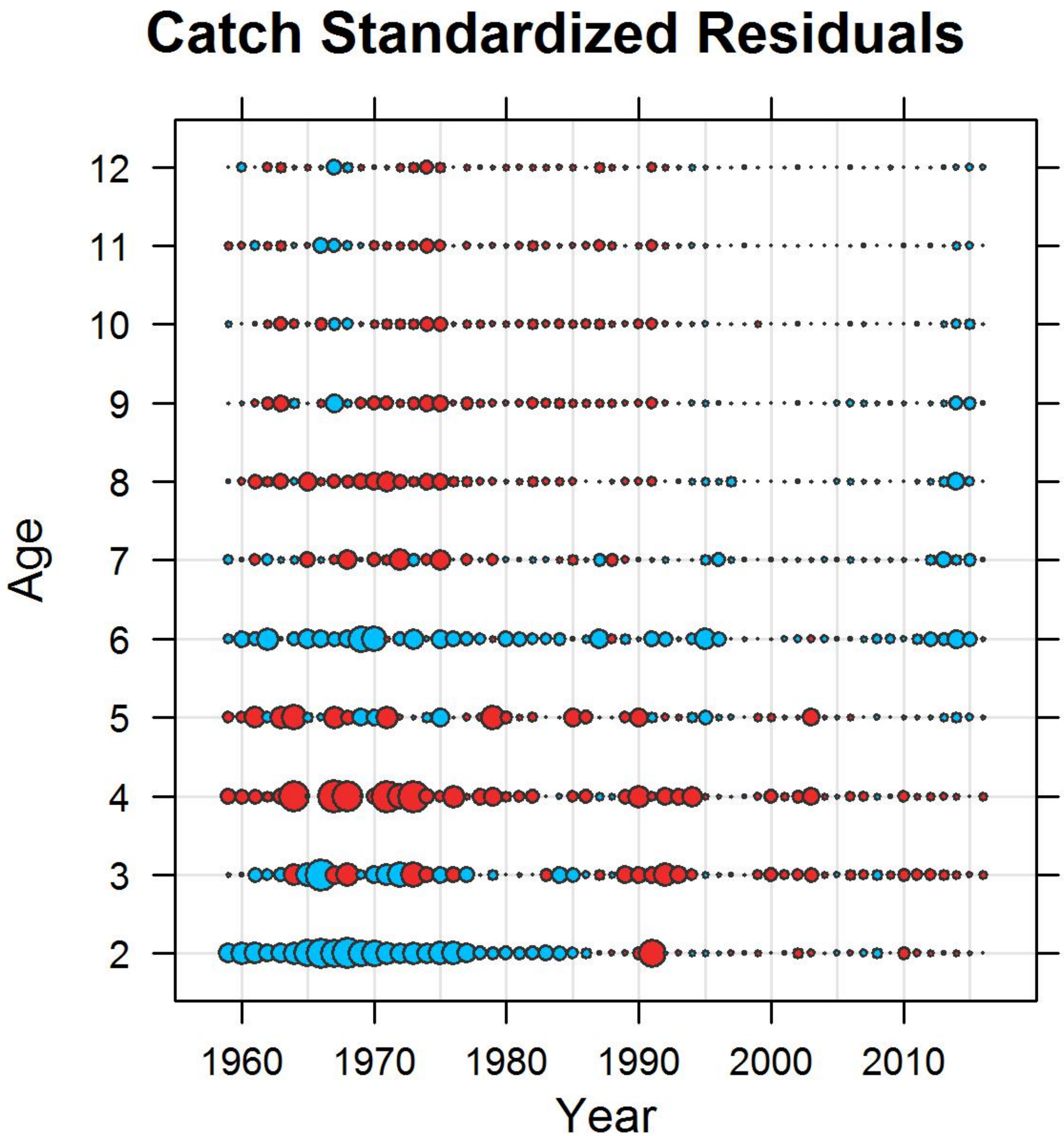


3NO cod SCA fits

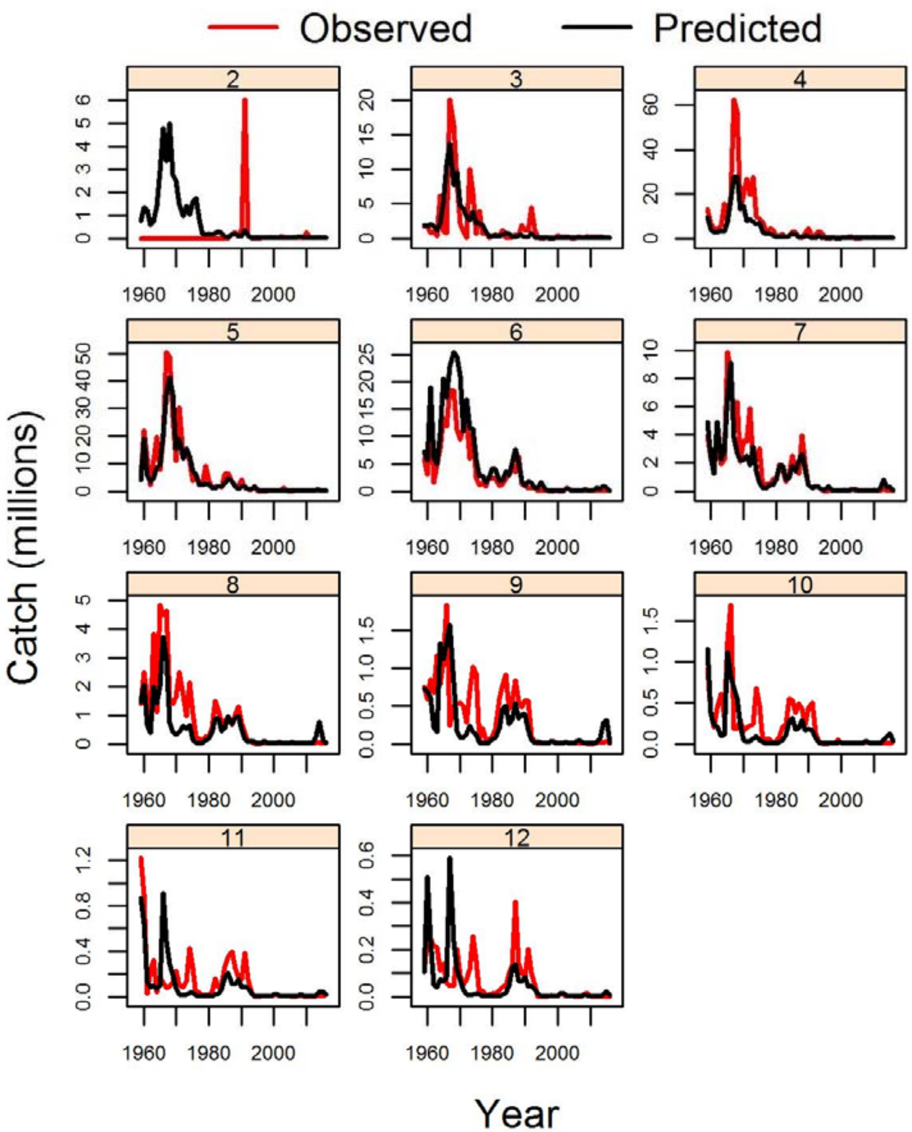


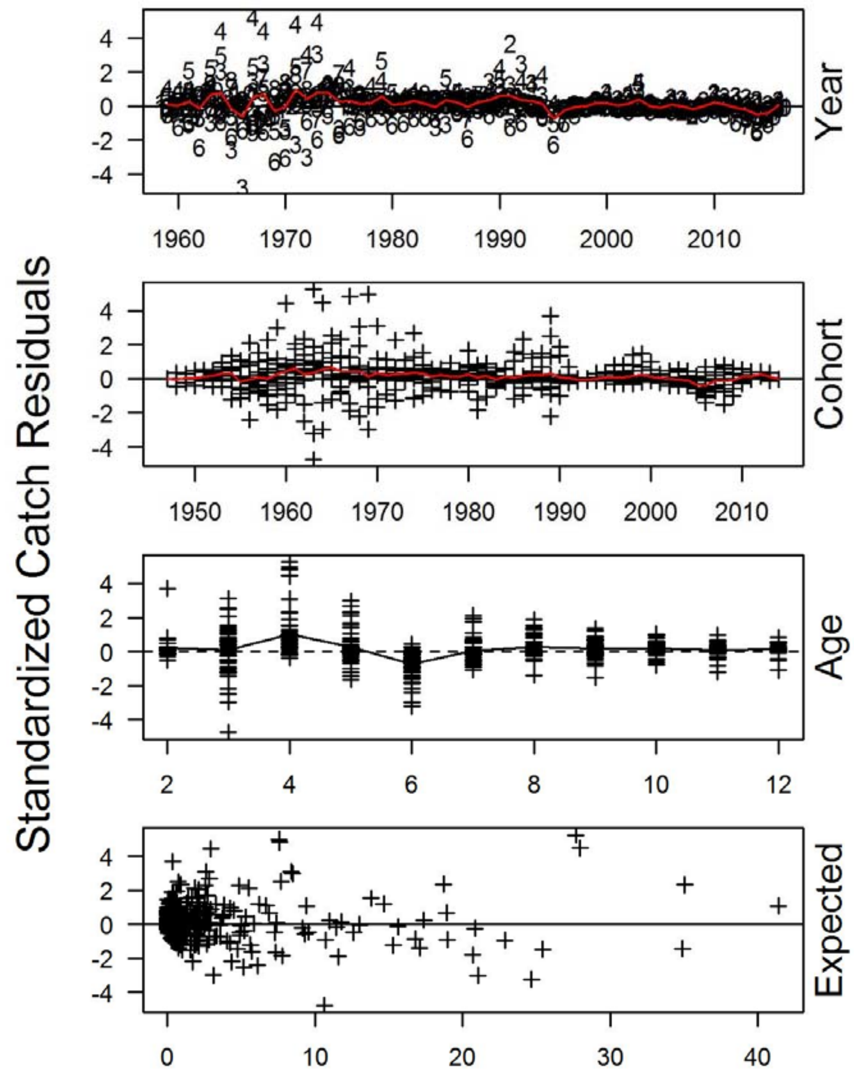


3NO cod SCA fits



3NO cod SCA C@A fits





3No cod SCA

- Separable *F* model unrealistic
- Catch fit function needs improvement
- How accurate is total catch?
- Index fit function too?? Year effects?
- There are model process errors we are not accounting for
- Not much data to fit initial numbers at age
- Too many recruitment parameters??

F6004 – The End!

- Separable *F* model unrealistic

- Too many recruitment parameters??