# Chile Executive Report

# Stock assessment with simulated data "like" Chilean jack mackerel

Cristian Canales R.
(ccanales@ifop.cl)
Instituto de Fomento Pesquero
Chile

## Summary

Two stock assessment models were used to evaluate a stock which mimic the Chilean jack mackerel population. The information available considers three fishing fleets named: Northern, Southern and Offshore.

The assessment framework is based on a statistical catch-at-age approach which considers information under observation error. The population dynamic is modeled forward in time and the differences between predicted and observed values, is modeled using likelihood distributions of the data. Models were fitted using a Bayesian estimation approach that minimizes the negative of the log-posterior distribution of the parameters.

For the given 100 dataset, a similar numbers of assessments were performed under the two models considered. The algorism was code for MATLAB V6.5.

## 1. Datasets and considerations

The details of the 100 datasets are shown bellow:

*Years of information for the different pieces of information*

| | Fleet | | | Survey | | | |
|---|---|---|---|---|---|---|---|
| | Southern | Northern | Offshore | Trawl | Acoustic | Early | Other |
| Catch-at-age proportions<br>sample size | 1975-2007<br>50 | 1975-2007<br>50 | 1979-1991<br>30 | 1997-2001<br>30 | 2002-2006<br>30 | N/A | N/A |
| Landings<br>std error (cv) | 1975-2007<br>0.05 | 1975-2007<br>0.05 | 1979-92 ; 2002-07<br>0.05 | | | | |
| Weigth-at-age | 1975-2007 | 1975-2007 | 1975-2007 | 1975-2007 | 1975-2007 | 1975-2007 | 1975-2007 |
| Abundance index<br>std error | | | | 1997-2001<br>by year | 2002-2007<br>by year | 1981-1983<br>by year | 1984-87 ;1991-95<br>by year |

The sample size and standard errors (std) used for the likelihood distributions were assumed to be known. The indexes of abundance were assumed as proportional to the exploitable biomass at the middle of the year. These indexes have different standard deviations for each year. Thus, the overall coefficient of variation was computed as the ratio between the Standard deviations and their mean which corresponds to the absolute value given for each index. The lognormal distribution was used for this purpose. Gaps of information are also noticed, particularly in age compositions data for some surveys such as "Early" and "Other". This problem was tackled assuming that these indexes are proportional to the total biomass at the middle of the year. In other words, there is not influence of the selectivity in modeling these observations. Another important issue regards with the fact that landings by fleet are identical between replicates, although the informed variation coefficient was 5%.

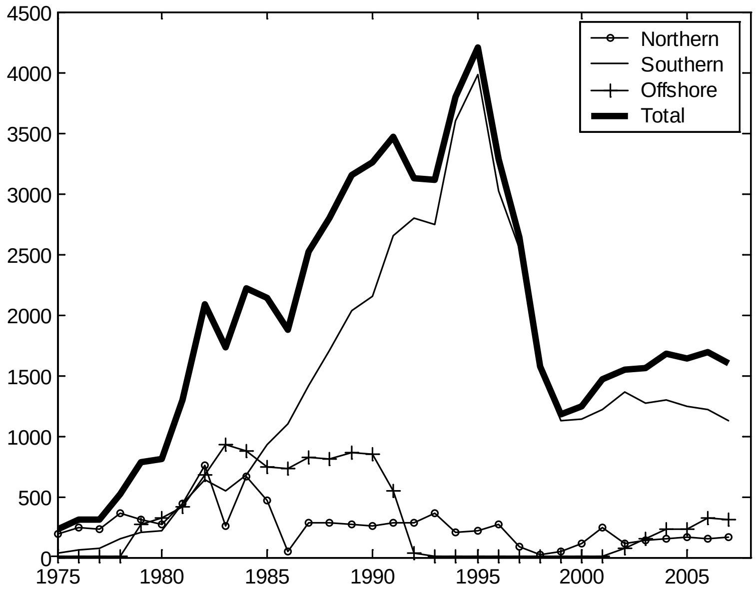


**Fig 1.** Simulated landing by fishing fleet

The maturity at age was considered as absolute and corresponds to the double to the given data. The natural mortality used was M=0,23 $year^{-1}$ and it was assumed constant across ages and years. Finally, an aging error matrix was considered on the observation process as product factor of the expected age composition in the catches by fleet and surveys. This means that the true catch-at-age is being taken out of the population model, but the predicted incorrect ages are compared to the observed ages.

## 2. Models

### a) Base model (U_model)

The model implemented used the Pope's catch equation, reproducing the population dynamics forward in time, considering observation errors in catch-at-age where landings are assumed to have no error. Here, the abundance is projected up to the middle of the year only affected by natural mortality, where the catches occur instantaneously. Afterward, the remaining of the population reaches the end of each year decaying only by natural mortality.

The model depends on a vector of parameter but it does not on the information itself. The model has the ability to predict all variables associated with the chosen assumptions. In this context, each vector of parameters is estimated minimizing the deviates between the predicted and the observed data. The model predicts the exploitable biomass and the available indexes of abundance at any point in time. The model also, predicts the observed age compositions for each fleet and surveys. The age compositions for each fishing fleet are estimated as a result of the abundance, selectivity and exploitation rate. Different selectivity functions across time were implemented. They were divided into four blocks for the southern fleet and into two block for the northern fleet. On the other hand, constant selectivity across years was used for the offshore fleet and surveys.

Regarding the Bayesian approach, the priors established are related to the recruitment process error (around to a known SR relationship) and the initial population structure (around the equilibrium state). The rest of the parameters were assumed to have non-informative priors ( for more details see Appendix 2).

### b) Alternative model: F_model

The model implemented corresponds to the Gulland's equation, reproducing the population dynamics forward in time, and admitting error in the catch-at-age and landings observed. Here, the fishing mortality is estimated explicitly and it is considerate as a result of an effect related to the age (Selectivity) and year. In this model, it is assumed that the total mortality (Z=F+M) is continuous inside the year. The catch-at-age is estimated (predicted) considering the Baranov catch equation. In equal terms, the model depends on a vector of parameter but it does not on the information itself. The model has the ability to predict all variables associated with the chosen assumptions. In this context, the vector of parameters is estimated minimizing the deviates between the predicted and the observed data. The model predicts the exploitable biomass and the available indexes of abundance at any point in time. Similarly to the U_model, different selectivity functions across time were implemented. They were divided into four block for the southern fleet and into two block for the northern fleet. On the other hand, constant selectivity across years was used for the offshore fleet and surveys.
( for more details see Appendix 2).

### c) Sensitivity analysis of the selectivity

A sensitivity analysis was carried out for the F_model to evaluate the model behavior assuming constant selectivity across years.

## 3. Results

100 stock assessments were performed based on the two models described above. Both models predict the observations with reasonable accuracy. The model outputs are subjected to the likelihood functions implemented, standard deviations assumed and the given sample sizes. In Appendix1, an example (dataset #1) of the fitting to the abundance indexes is shown, where no important differences between models can be found. It is also shown the fitted U_model to the age composition. In general terms, these models reproduce the data reasonably well, and they seem to be appropriate to describe the main stock and exploitation indicators.

As it was mentioned above, one of the main assumptions in modeling this population regards with the selectivity for each fleet. The results suggest important changes in selectivity, particularly in the northern fleet, where the age at the full exploitation goes down after 1986. This situation is related to important recruitments observed in those years (Fig 2). As a consequence of this, the overall selectivity shows different forms between years. However, it does not seem to cause important differences between both models, where at the beginning of the time series the northern fleet selectivity became important due to its higher landings in those years. (Fig. 3).

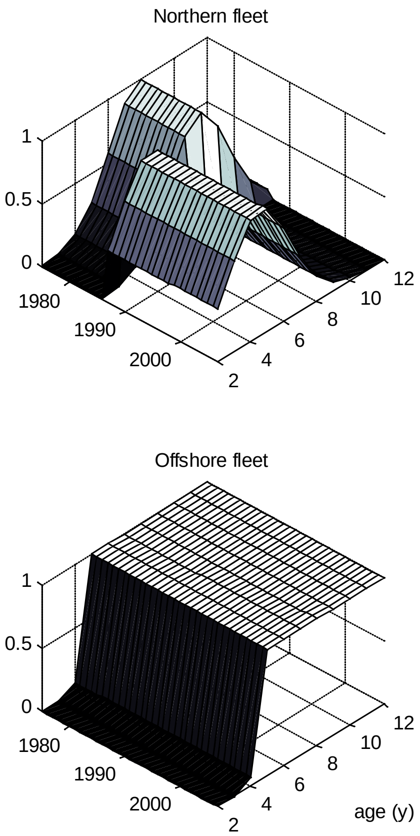


Figure 2. An example of selectivity patterns estimated for different fleets.

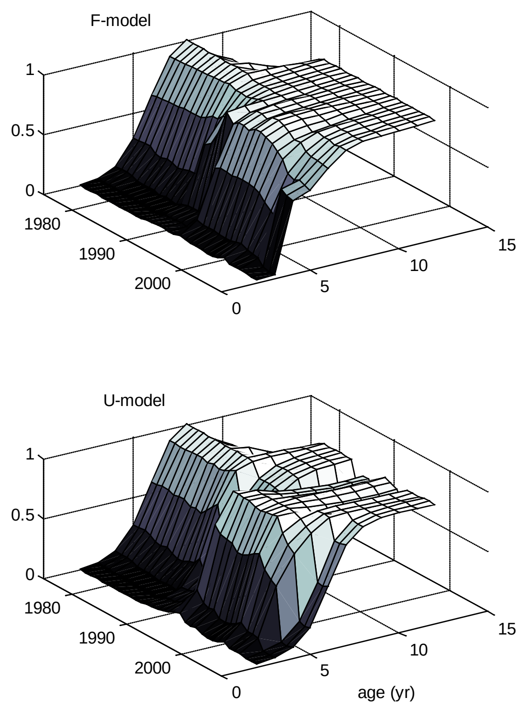


Figure 3. An example of total selectivity patterns estimated for different models.

The result of both models applied for each dataset do not show significant differences in the biomass and recruitments levels. However, the base model (U_model) showed higher variability in their estimates, particularly at the end of the assessment period. The stock shows two large recruitments (1985-1986), which explain the increase of the biomass until 1990. Afterward, high levels of catches and recruitments around the historical average, caused the decrease of the biomass and its current declining trends. (Fig 4 & Fig 5).

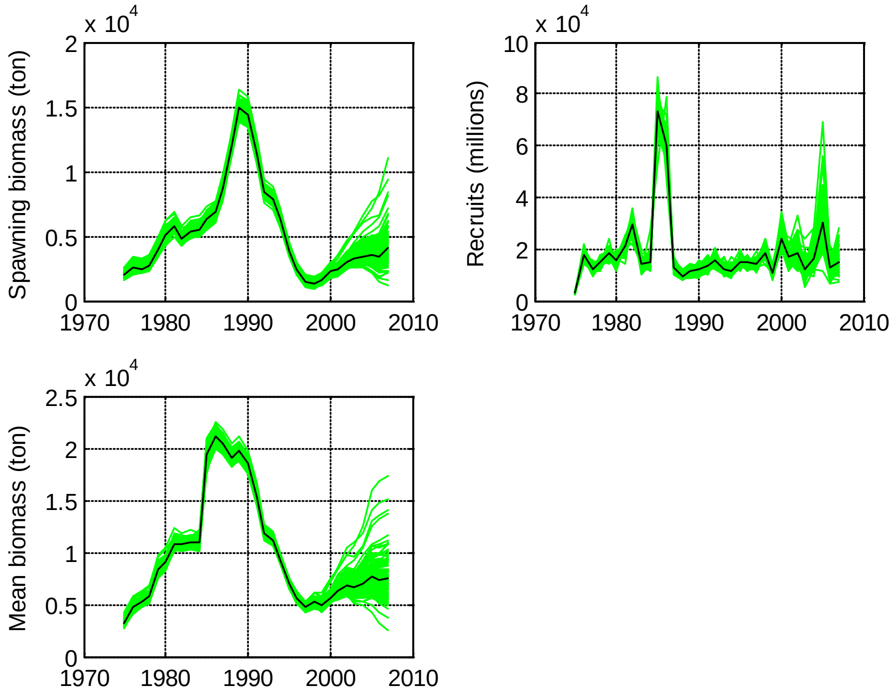


Figure 4. Spawning biomass, recruitment and mean biomass estimated by "U-model". The black line represents the median of 100 estimates.

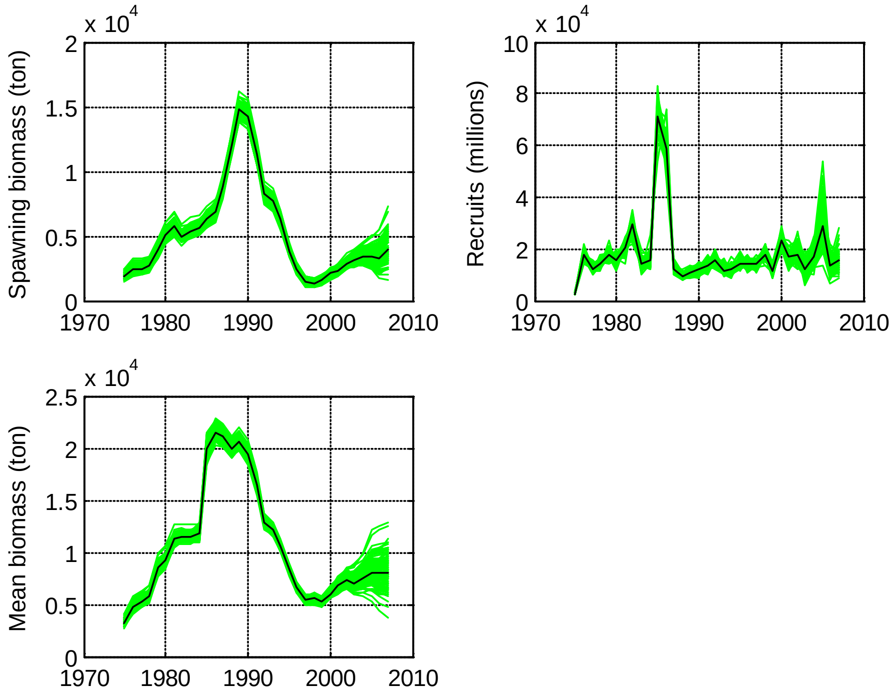


Figure 5. Spawning biomass, recruits and mean biomass estimated by "F-model". The black line represents the median of 100 estimates.

The results of the sensitivity analysis (F_modScte) show that there are important differences in the biomass level, which are mostly dependent on the assumptions chosen (Fig 6). The assumption of constant selectivity implies higher biomass estimates at the end of assessment period. However, this difference seems to be less important if we consider the uncertainty in its estimates.

The behavior of these models suggest that in general terms, all models implemented were able to reproduce the fluctuations of the main variables of the population. However, the model with constant selectivity suggests that the recent recruitments have had higher variability and increase trend. The “F-model” and “U-model” do not show differences in the Spawning biomass levels (Fig. 7). It is important to notice that the recruitment after 1987, has shown erratic variations with no clear trend, excepting for the model that considers constant selectivity across years.

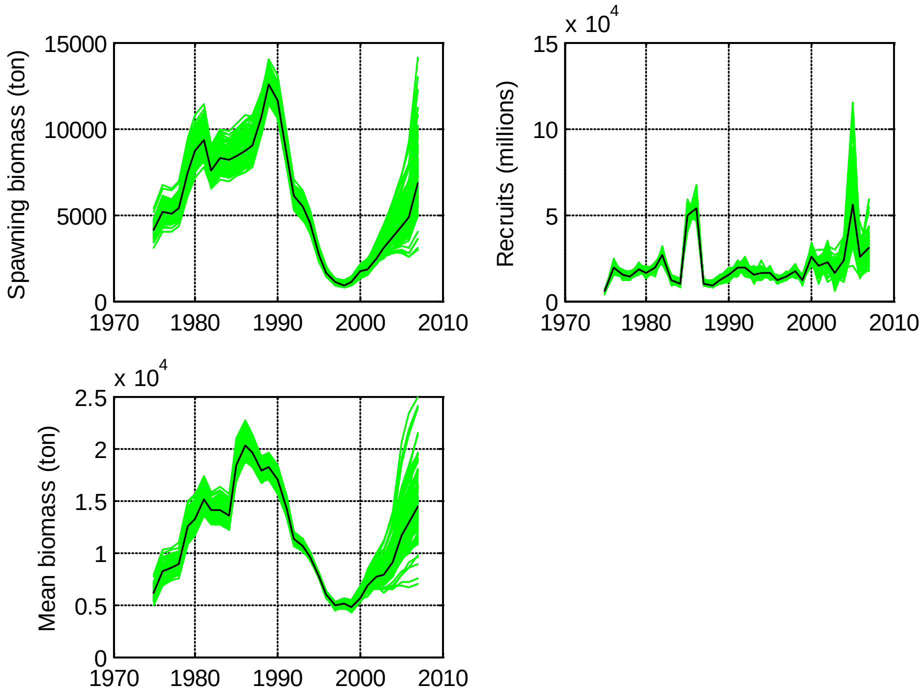


Figure 6. Spawning biomass, recruits and mean biomass estimated by “F-model” with constant selectivity. The central line represents the median of 100 estimates.

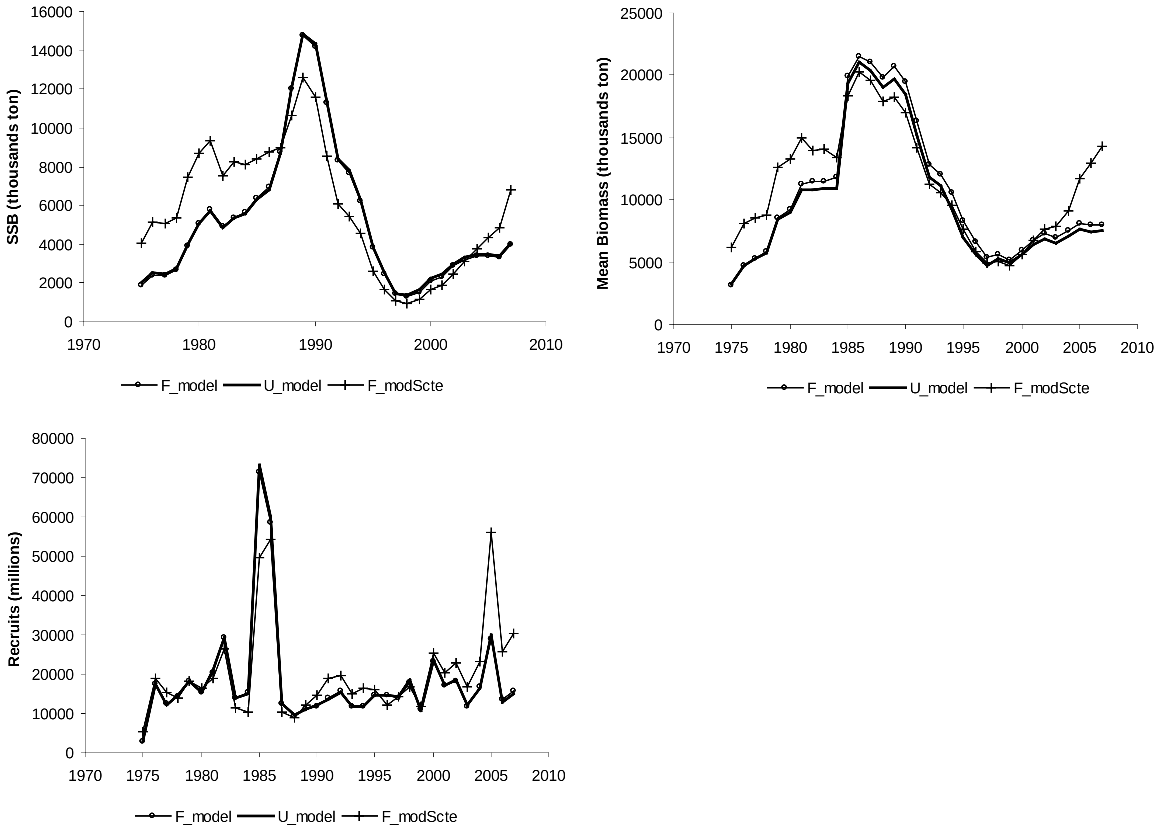


Fig 7. Expected (median) spawning biomass, mean biomass and recruitment estimated between models.

Bayesian Information Criteria (BIC) was used to evaluate the goodness-of-fit of the models implemented. BIC criterion evaluate model selection by weighting the likelihood value and the numbers of parameters fitted in order to identify the best model. Due to its lower BIC value, the base model (U_model) shows the best statistical behaviour. However, the estimated spawning biomass is in the range of 2.1 – 8.3 million tons, which overlap with the confidence intervals of the other models (Table 1). The U_model has less parameters to estimate (more parsimony) in comparison with the other models which might be considered for this purposes.

Table 1

Statistical behavior of different models

| Model | Fleet selectivity | Free parameters (#) | -Log likelihood (1) | BIC | $SSB_{2007}$ (1) (million tons) | CV | $IC_{95\%}$ (million tons) |
|---|---|---|---|---|---|---|---|
| U_mod | Blocks | 63 | 7,494 | 15,424 | 4.04 | 0.37 | [2.1 - 8.3] |
| F_mod | Blocks | 160 | 7,464 | 16,035 | 3.95 | 0.26 | [2.4 - 5.8] |
| | Constant | 152 | 7,759 | 16,570 | 6.82 | 0.28 | [3.5 - 12.2] |

(1) corresponds to the average after 100 runs

## 3. Conclusions.

The behavior of the two models considering simulated data was evaluated. The base model has less parameter to estimate, being more parsimonious, and its statistical behavior was better. This model is based on the Pope's catch equation with a forward dynamics and it requires explicit assumptions related to the selectivity of the different fleets described. The advantages of the statistical catch-at-age models are mainly related with their utility in presence of gaps of information and the flexibility to include and evaluate different hypotheses. These hypotheses are related to the biological processes and observation errors. For example, it could be possible to model explicitly two Stock-Recruitment relationships in relation to different productivity periods (Fig 8).

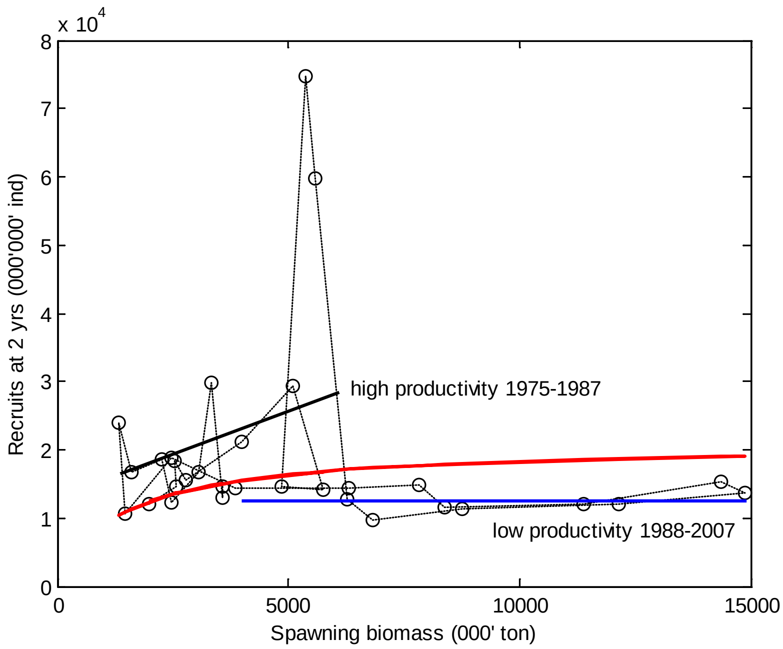


Fig 8. Stock recruitment relationship for different productivity periods.

APPENDIX 1. Fitted Model

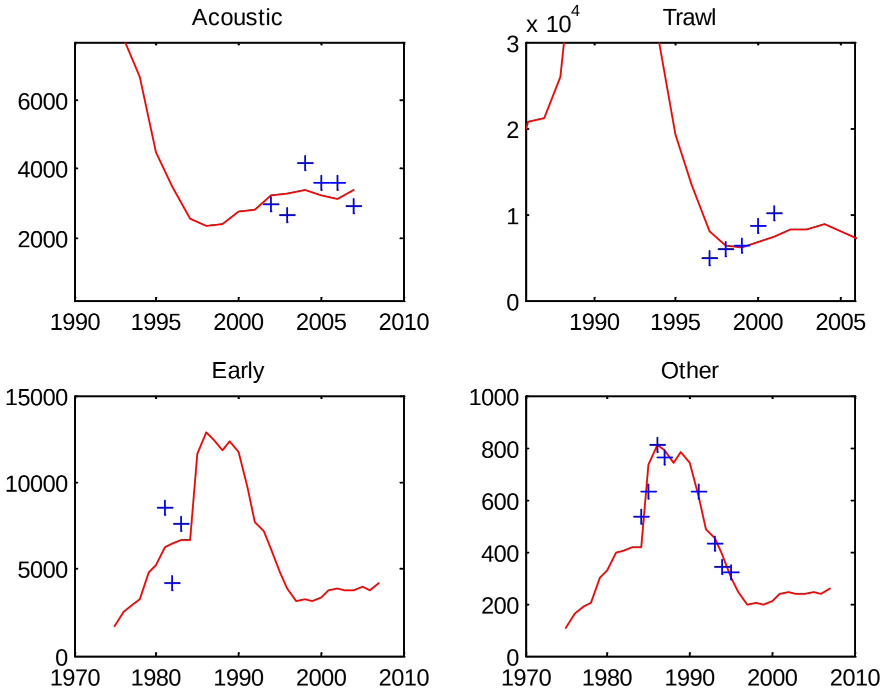


**Figure 1.** Fitted model to the abundance indexes (the line is the model and the crosses represent the data). U_model

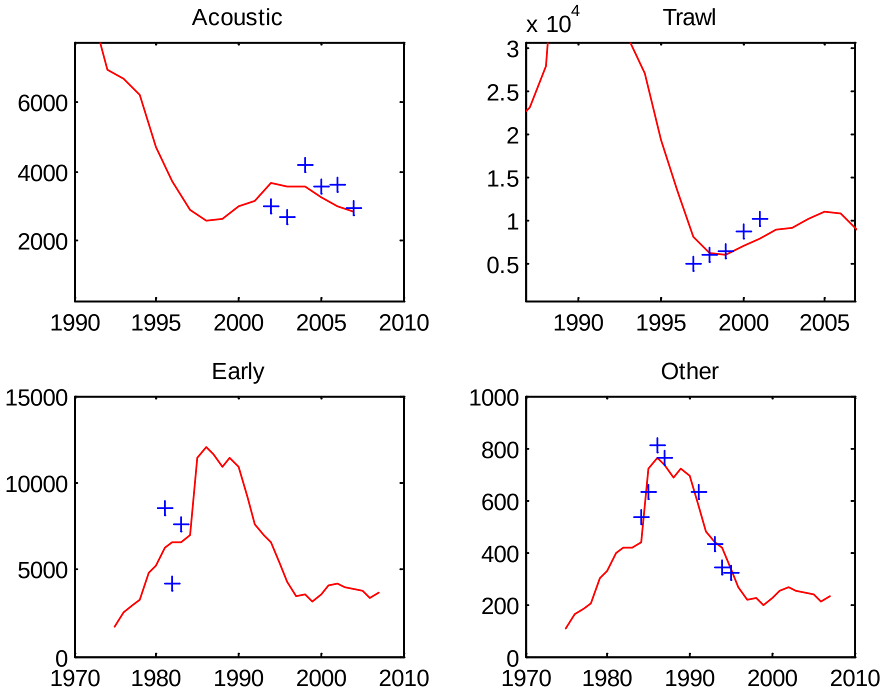


**Figure 2.** Fitted model to the abundance indexes (the line is the model and the crosses the data). F_model

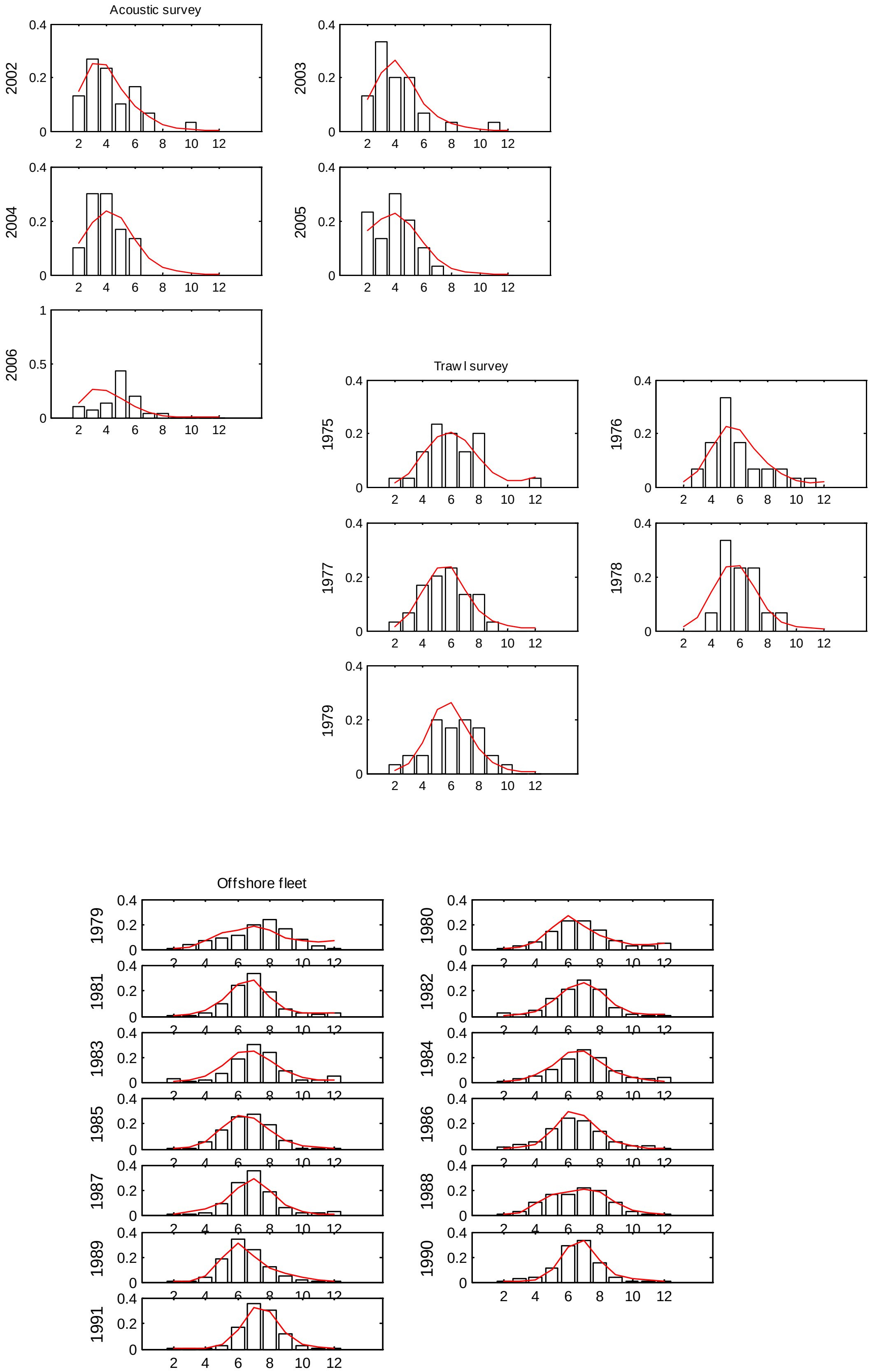


**Figure 3.** Fitted model to the age composition in the catches of surveys and offshore fleet

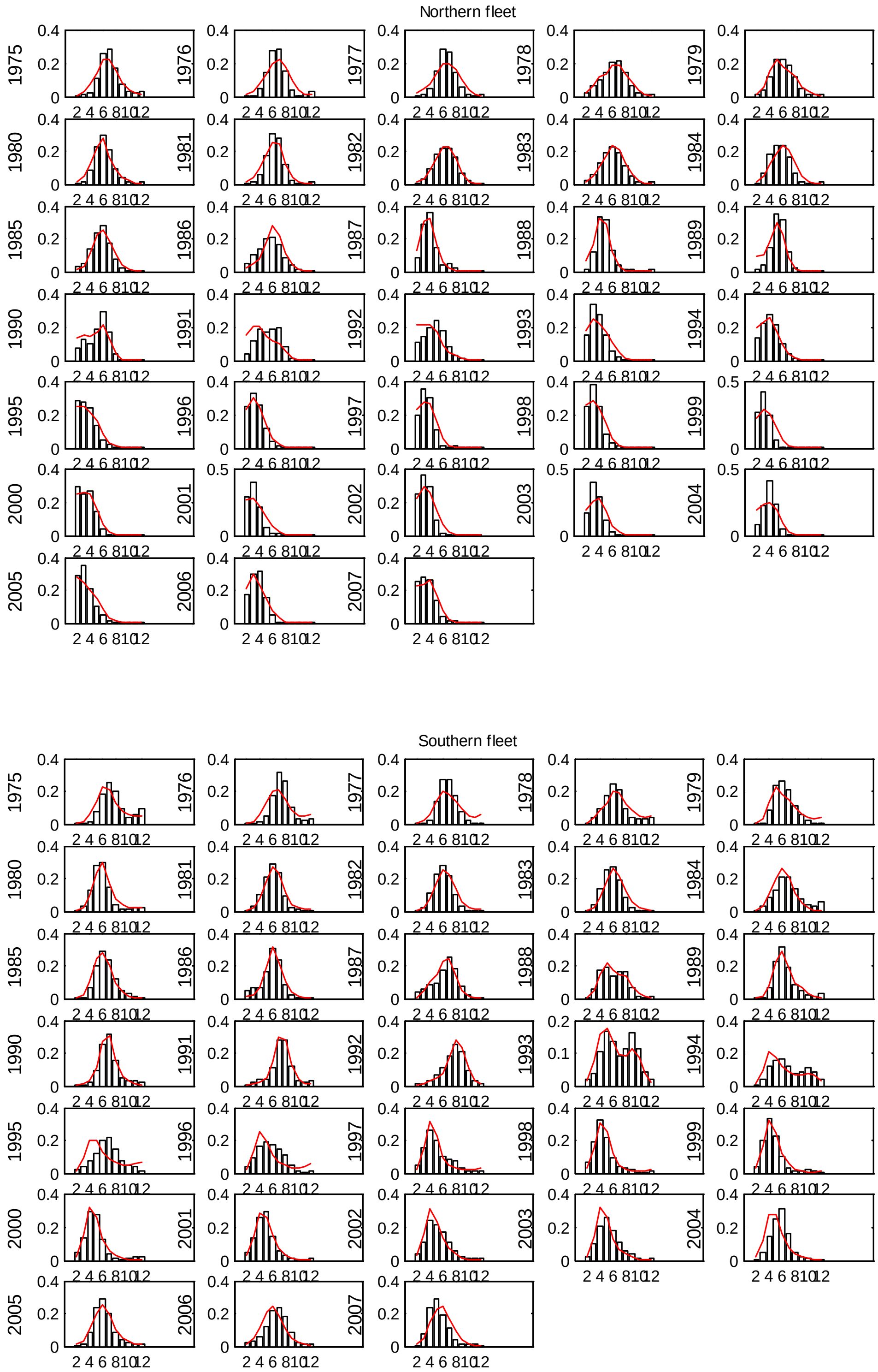


Figure 4. Fitted model to the age composition in the catches of northern and southern fleet.

## Process model

<table>
<tr><th>Base model (U_model)</th><th>Alternative model (F_model)</th></tr>
<tr><td>Population survival<br>$$N_{a,y}=\begin{cases} R_0 e^{-M(a-2)} e^{-\tau_a} & a=2-12; y=1 \\ N_{a-1,y-1}e^{-M}-\hat{C}^{tot}_{a-1,y-1}e^{-0.5M} & y=2-33 \\ \left(N_{a-1,y-1}e^{-M}-\hat{C}^{tot}_{a-1,y-1}e^{-0.5M}\right)+ \\ \left(N_{a,y-1}e^{-M}-\hat{C}^{tot}_{a,y-1}e^{-0.5M}\right) & a=12+; y=2-33 \end{cases}$$<br>$$\tau \sim \mathrm{N}(0,\sigma^2_{No})$$<br>$$\hat{C}^{tot}_{a,y}=\sum_f \hat{C}^f_{a,y}$$<br>$$\hat{C}^f{}_{a,y}=\mu^f{}_{a,y}N_{a,y}e^{-0.5M}$$<br>$$\mu^f{}_{a,y}=\mu^f_y S_{a,y}$$<br>$$\mu^f_y=\frac{D^f_y}{BV^f_y}=\frac{D^f_y}{\sum_a N_{a,y}S^f_{a,y}w_{a,y}e^{-0.5M}}$$</td><td>$$N_{a,y}=\begin{cases} R_0 e^{-M(a-2)} e^{-\tau_a} & a=2-12; y=1 \\ N_{a-1,y-1}e^{-Z_{a-1,y-1}} & y=2-33 \\ N_{a-1,y-1}e^{-Z_{a-1,y-1}}+N_{a,y-1}e^{-Z_{a,y-1}} & a=12+; y=2-33 \end{cases}$$<br>$$\tau \sim \mathrm{N}(0,\sigma^2_{No})$$</td></tr>
<tr><td>Mortality<br><br>Implicit</td><td>Mortality<br>$$Z_{a,y}=\sum_f F^f_{a,y}+M$$<br>$$F^f_{a,y}=S^f_a F^f_{cr,y}$$</td></tr>
</table>

| | |
|---|---|
| Recruitments at age 2<br>$R_y = H(SSB_{y-2})e^{-\varepsilon_y}$ ; $y = 3-33$<br>$H(SSB_{y-2}) = E(R_y) = \frac{\alpha\, SSB_{y-2}}{\beta + SSB_{y-2}}$<br>$\varepsilon \sim \mathrm{N}(0, \sigma^2{}_R)$ | Recruitments at age 2<br>Idem |
| Spawning biomasa<br>$SSB_y = \sum_a (N_{a,y}e^{-0.5M} - \hat{C}_{a,t})e^{-4.5/12M} O_a w_{a,y}$ | Spawning biomass<br>$SSB_y = \sum_a N_{a,y} e^{-\frac{10.5}{12} Z_{a,y}} O_a w_{a,y}$ |
| Abundance at mid year (mean abundance)<br>$\overline{N_a} = (N_{a,y}e^{-0,5M} - \hat{C}_{a,y})$ | Abundance at mid year (mean abundance)<br>$\overline{N_a} = N_{a,y}e^{-0,5Z_{a,y}}$ |
| Selectivity by fleet and year<br>$S^{f1}_{a,T} = \left[1 + e^{-\ln(19)\frac{a - a^{f1}_{50\%}}{\Delta^{f1}}}\right]^{-1}$ (Southern fleet)<br>$S^{f2}_{a,y} = \left[1 + e^{-\ln(19)\frac{a - a^{f2}_{50\%}}{\Delta^{f2}}}\right]^{-1}$ (Offshore fleet)<br>$S^{f3}_{a,T} = e^{-\frac{(a - a^{f3}_{50\%})}{2\delta^{f3}}}$ (Northern fleet) | Selectivity by fleet and year<br>Idem |

## Observation models

| Base model (U_model) | Alternative model (F_model) |
| --- | --- |
| Catches<br>$E(C^{f}_{a,y}) = E_{a,a}\hat{C}^{f}_{a,y}$ | Catches and landings<br>$E(C^{f}{}_{a,y}) = E_{a,a}\frac{F^{f}_{a,y}}{Z_{a,y}}N_{a,y}e^{-Z_{a,y}}$<br>$D^{f}_{y} = \sum_{a}\hat{C}^{f}_{a,y}w_{a}$ |
| Abundance indexes<br>$\hat{B}^{s}_{y} = q^{s}_{y}\sum_{a}\overline{N}_{a}S^{s}_{a}\overline{w}_{a,y}$<br>$q^{s} = \exp\left\{\frac{1}{n}\sum_{i}\log\left(\frac{\dot{B}^{s}_{y}}{\hat{B}^{s}_{y}}\right)\right\}$<br>$S^{s}_{a} = \left[1+e^{-\ln(19)\frac{a-a^{s}_{50\%}}{\Delta^{s}}}\right]^{-1}$ | Abundance indexes<br>$\hat{B}^{s}_{y} = q^{s}_{y}\sum_{a}\overline{N}_{a}S^{s}_{a}\overline{w}_{a,y}$<br>$q^{s} = \exp\left\{\frac{1}{n}\sum_{i}\log\left(\frac{\dot{B}^{s}_{y}}{\hat{B}^{s}_{y}}\right)\right\}$<br>$S^{s}_{a} = \left[1+e^{-\ln(19)\frac{a-a^{s}_{50\%}}{\Delta^{s}}}\right]^{-1}$ |

# Error models

| Base model (U_model) | Alternative model (F_model) |
|---|---|
| Joint log-posterior<br>$z = \sum_f \ln L_{p^f} + \sum_s \ln L_{p^s} + \sum_s \ln L_{B^s} + \ln p(\tau) + \ln p(\varepsilon)$ | Joint log-posterior<br>$z = \sum_f \ln L_{p^f} + \sum_s \ln L_{p^s} + \sum_s \ln L_{B^s} + \sum_f \ln L_{D^f} + \ln p(\tau) + \ln p(\varepsilon)$ |
| Log-likelihood for proportions by fleet (multinomial distribution)<br>$\ln L_{p^f} = -n^f p^f_{a,y} \ln(\hat{p}^f_{a,y})$<br>$\hat{p}^f_{a,y} = \frac{E(\hat{C}^f_{a,y})}{\sum_a E(\hat{C}^f_{a,y})}$ ; $p^f_{a,y} = \frac{C^f_{a,y}}{\sum_a C^f_{a,y}}$ | Log-likelihood for proportions by fleet (multinomial distribution)<br>Idem |
| Log-likelihood for proportions by survey (multinomial distribution)<br>$\ln L_{p^s} = -n^s p^s_{a,y} \ln(\hat{p}^s_{a,y})$<br>$\hat{p}^s_{a,y} = \frac{E_{a,a} \overline{N}_{a,y} S^s_a}{\sum_a E_{a,a} \overline{N}_{a,y} S^s_a}$ ; $p^S_{a,y} = \frac{N^s_{a,y}}{\sum_a N^S_{a,y}}$ | Log-likelihood for proportions by survey (multinomial distribution)<br>Idem |
| Log-likelihood for abundance indexes (log-normal distribution)<br>$\ln L_{B^s} = \frac{1}{2\sigma_s^2} \sum_y \ln\left(\frac{B^s_y}{\hat{B}^s_y}\right)^2 + c$ | Log-likelihood for abundance indexes (log-normal distribution)<br>Idem |
| | Log-likelihood for landings (log-normal distribution)<br>$\ln L_{D^f} = \frac{1}{2\sigma_D^2} \sum_y \ln\left(\frac{D^f_y}{\hat{D}^f_y}\right)^2 + c$ |

Nomenclature

a = age

y= year

$\hat{C}_{a,y}$ = Expected catch-at-age

$E(\hat{C}_{a,y})$ = Corrected catch-at-age

$C_{a,y}$ = Observed catch-at-age

$f$ =fleet

M = natural mortality

$F^{f}_{a,y}$ = Fishing mortality at age by fleet

$S^{f}_{a}$ = Selectivity at age by fleet

$Z_{a,y}$ = Total mortality at age

Ro=Initial recruitment

$\tau$= log-deviate of initial abundance, $\sigma^{2}_{No}$ =standard deviation of $\tau$ (=0.5)

H = S/R relationship (B&H by default h=0,75)

$\varepsilon$= log-deviate on recruitments, $\sigma^{2}_{R}$ = standard deviation of $\varepsilon$ (=0.6)

w= mean weigth-at-age

O= Sexual maturity at age, T=Year blocks:

$a_{50\%}$= age at 50% of selectivity, $\Delta$ = range between $a_{50\%}$ and $a_{95\%}$, T= Selectivity by blocks. Two blocks in the North (1975-1986 and 1987-2008) and fourth blocks in the South (1975-1987; 1988-1992; 1993-2004 and 2005-2007)

$\delta$=Std. deviation

$\mu$ = Exploitation rate

$D$ = Observed landings (ton)

$\hat{D}$ = Expected landings (ton)

BV=Exploitable biomass

n = number of positive observations

$\dot{B}^{s}_{y}$ = Biomass observed in surveys

$S^{s}_{a}$ = Survey selectivity (availability)

s= Survey type: Acoustic, Trawl, Early, Other.

lnL=log likelihood

p( )=log-priors on $\tau$ and $\varepsilon$

$n_f$=sample size by fleet

$p^f$= proportions of catch-at-age by fleet

$n^s$=sample size by survey

$p^s$= proportions of catch-at-age by survey

$\sigma^{2}_{s}$ = standard deviation of log-B

$\sigma^{2}_{D}$ = standard deviation of log-D

c=constant term

$E_{a,a}$ = ageing error matrix

$N^{s}$ = abundance age composition in surveys

$N$ = population abundance at age