State–space Stock Assessment as simple alternative to (semi) deterministic approaches and full parametric stochastic models

Anders Nielsen & Casper W. Berg
an@aqua.dtu.dk

DTU Aqua
National Institute of Aquatic Resources

Problems we wish to solve

- Deterministic approaches
  - Catch at age assumed known without error
  - Procedures not models
  - Convergence of a deterministic procedure
  - Ad-hoc adjustments
- Full parametric statistical models
  - Parametric F–structure (e.g. multiplicative)
  - Trade off between flexible with (too) many parameters and rigid with tractable number of parameters
  - Number of parameters increase with every new year of data added

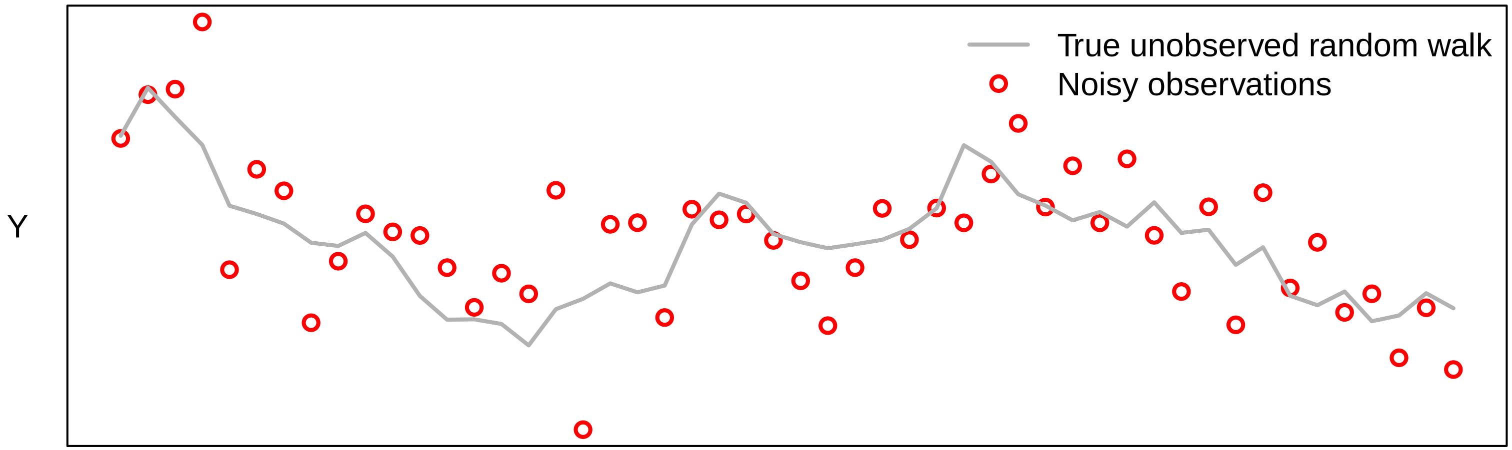
State-space assessment models

- This model class[a] is used in most other quantitative fields
- It is a very useful extension to full parametric statistical models.
- Introduced for stock assessment by Gudmundsson (1987,1994) and Fryer (2001)
- The reason state-space models have not been more frequently used in stock assessment is that software to handle these models has not been available
- Can give very flexible models with low number of model parameters
- For instance we can include things like:

$F_{3,y}$ is a random walk with yearly variance σ^2

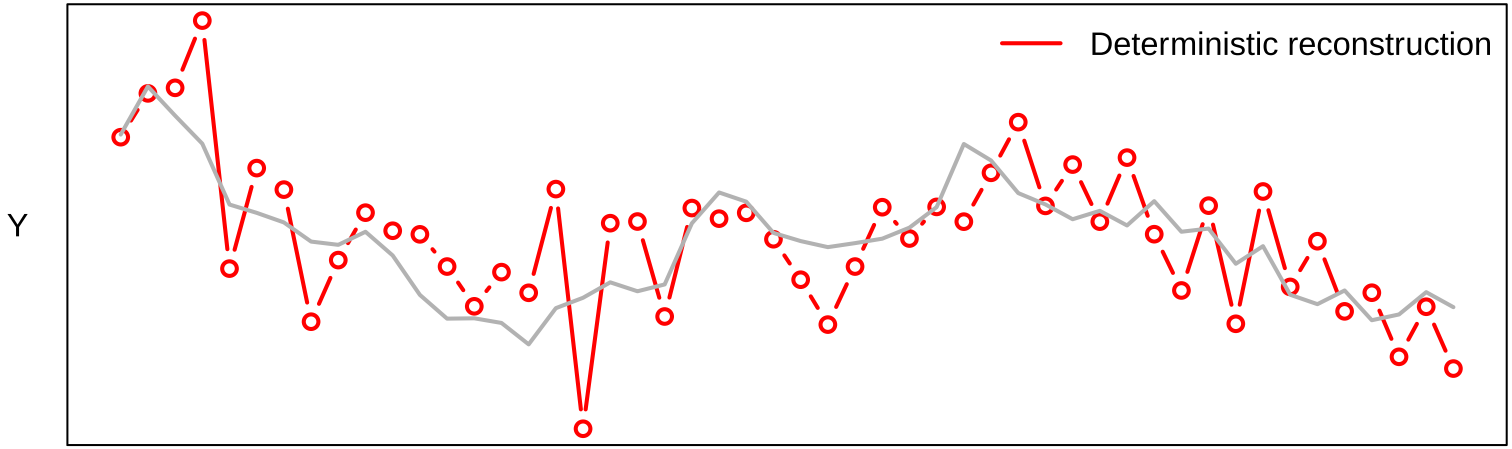
[a] a.k.a. random effects models, mixed models, latent variable models, hierarchical models, ...

Illustration of the three types of models



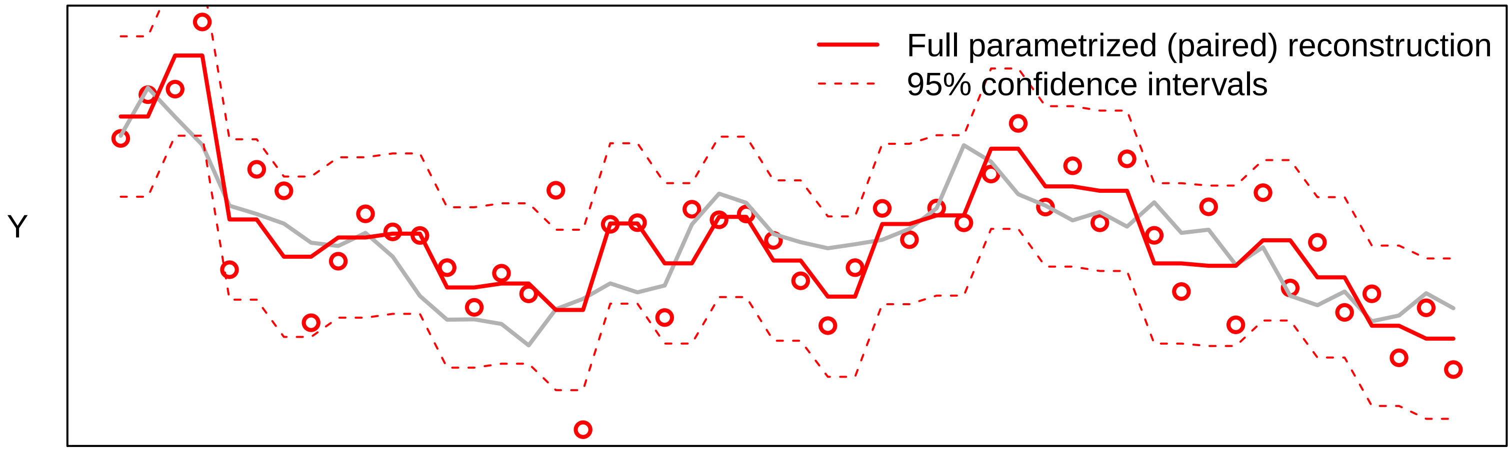
- Consider this example:
  - The true underlying F (here grey) follows a random walk with variance σ_F^2
  - But we only observe Y (here red circles) which is F+'noise' with variance σ_Y^2

Deterministic model estimates



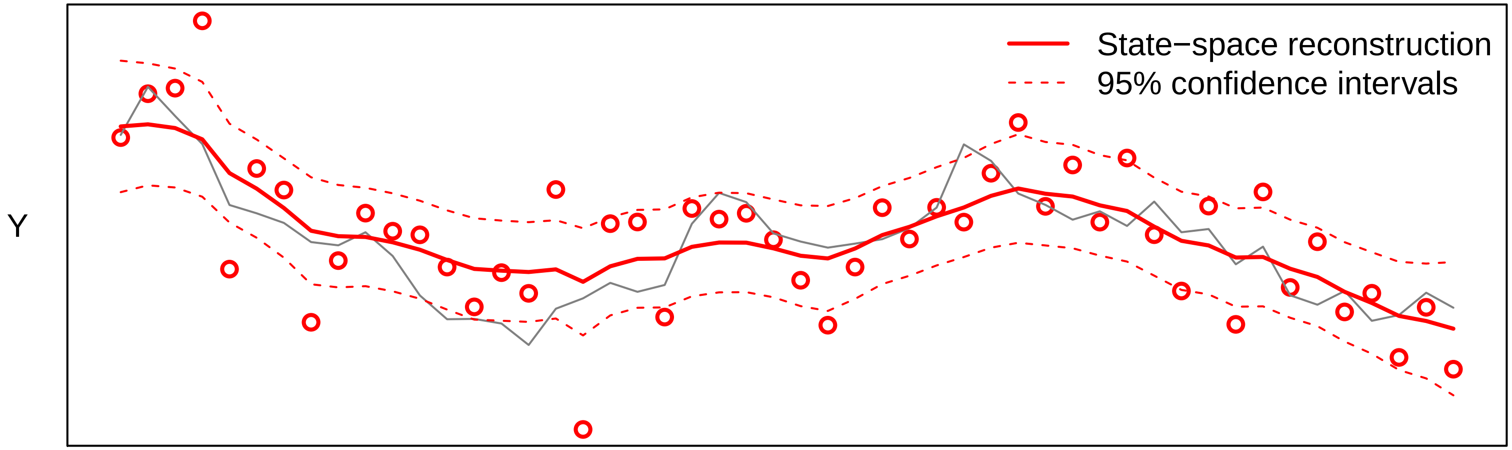
- If we assume no observation error the estimate of F is Y
- Too fluctuating
- No quantification of uncertainties

Fully parametrized statistical model estimates



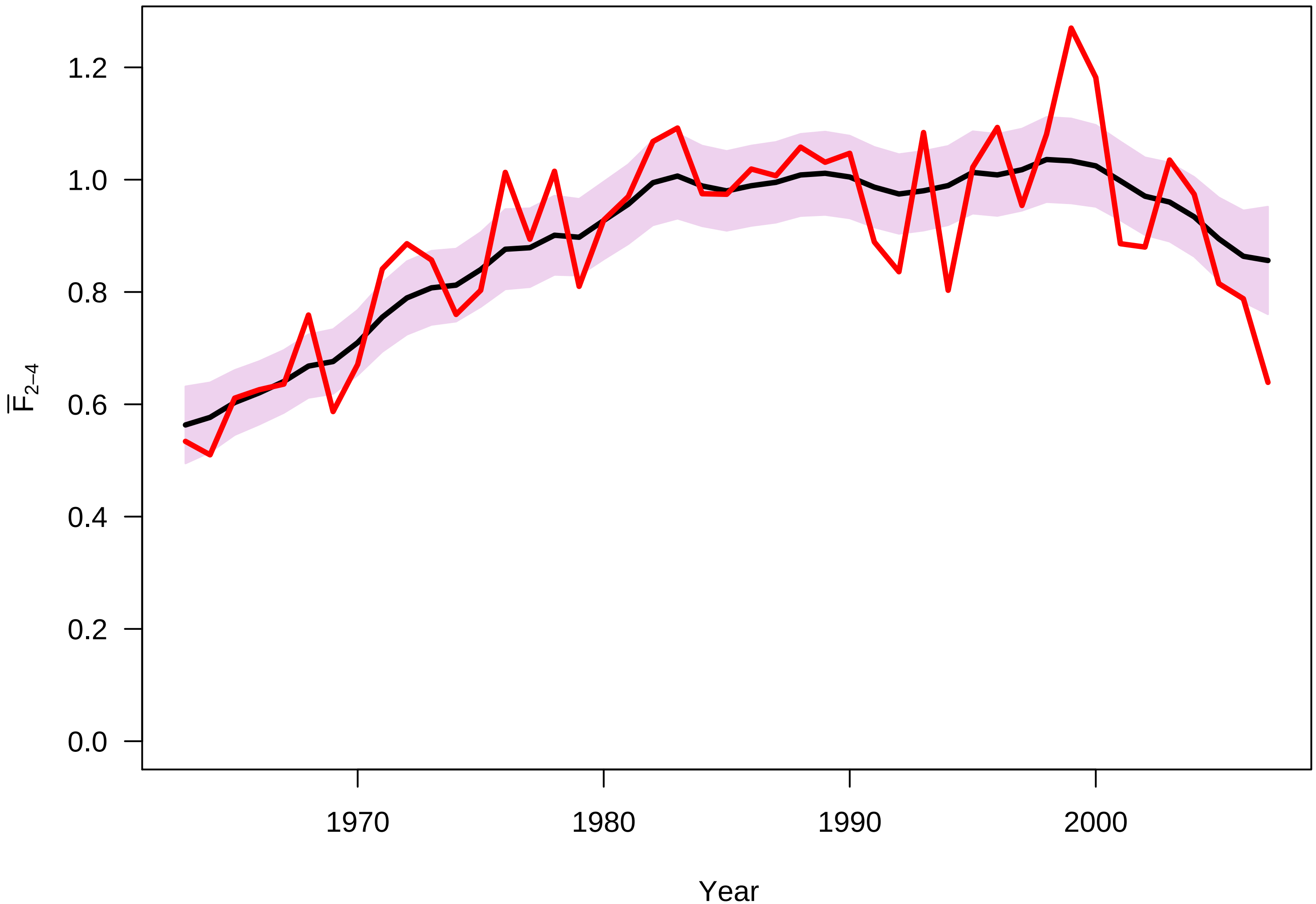
- To use a fully parametrized statistical model we first had to group the observations (here pairs, but choice is arbitrary)
- The reconstructed track appear OK
- The model contain 26 model parameters
- Uncertainties are estimated but the confidence interval seems too wide

State-space model estimates

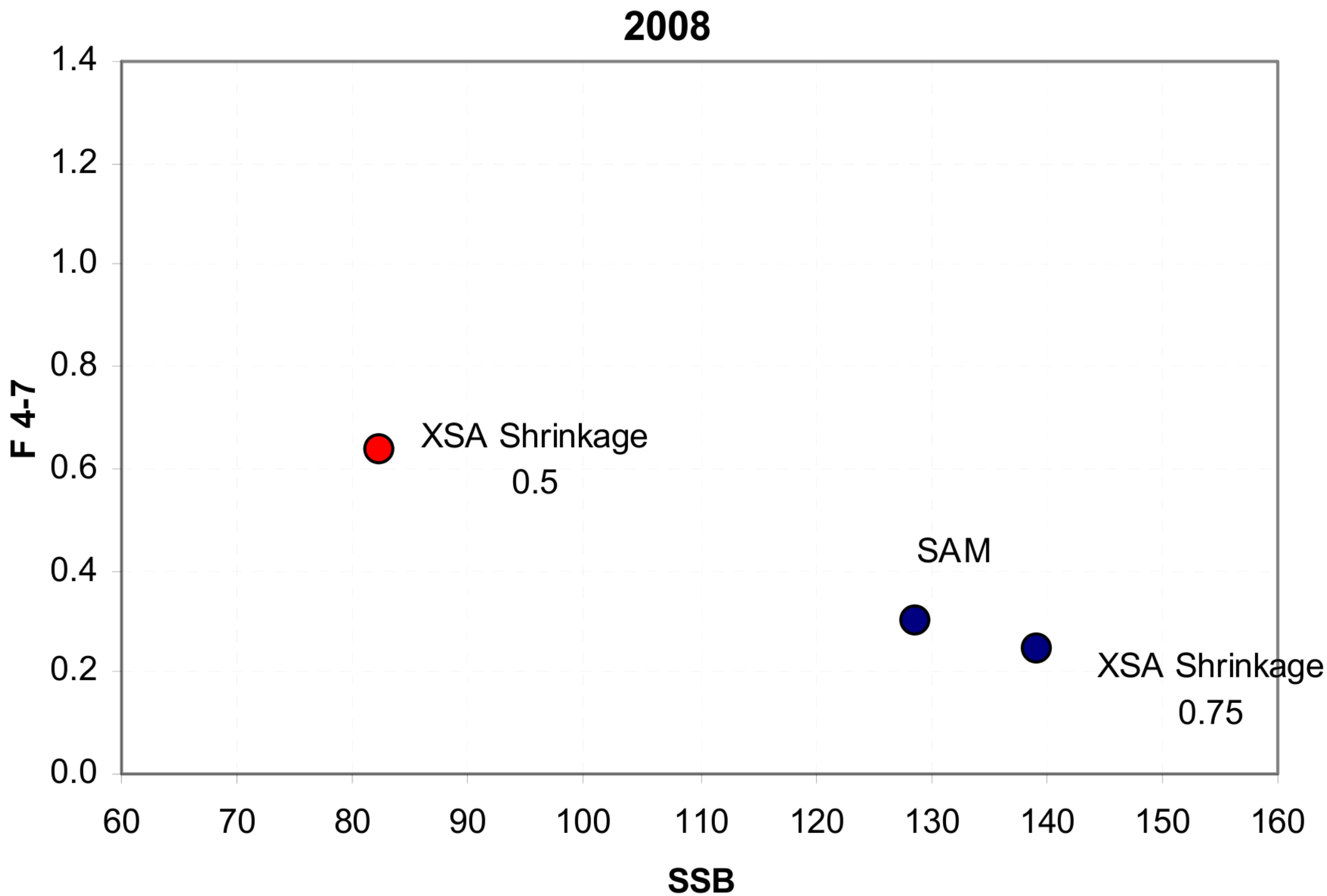


- Consider F as unobserved random variable
  - Estimate model parameters (σ_η and σ_ε) in marginal distribution $\int p(F, Y)dF$
  - Predict F via distribution of $F|Y$
- Closer reconstruction
- No artificial assumptions
- Two model parameters
- Correct coverage of the confidence interval
- Naturally this is just a simulated example, but ...

Example: $\overline{F}_{2-4}$ for North Sea Cod



Avoiding ad-hoc choices — Eastern Baltic Cod



- Using the State-space Assessment Model (SAM) gives us an objective criteria

Model

States are the random variables that we don't observe ($N_{a,y}$, $F_{a,y}$)

$$\begin{pmatrix} \log(N_y) \\ \log(F_y) \end{pmatrix} = T \begin{pmatrix} \log(N_{y-1}) \\ \log(F_{y-1}) \end{pmatrix} + \eta_y$$

Observations are the random variables that we do observe ($C_{a,y}$, $I_{a,y}^{(s)}$)

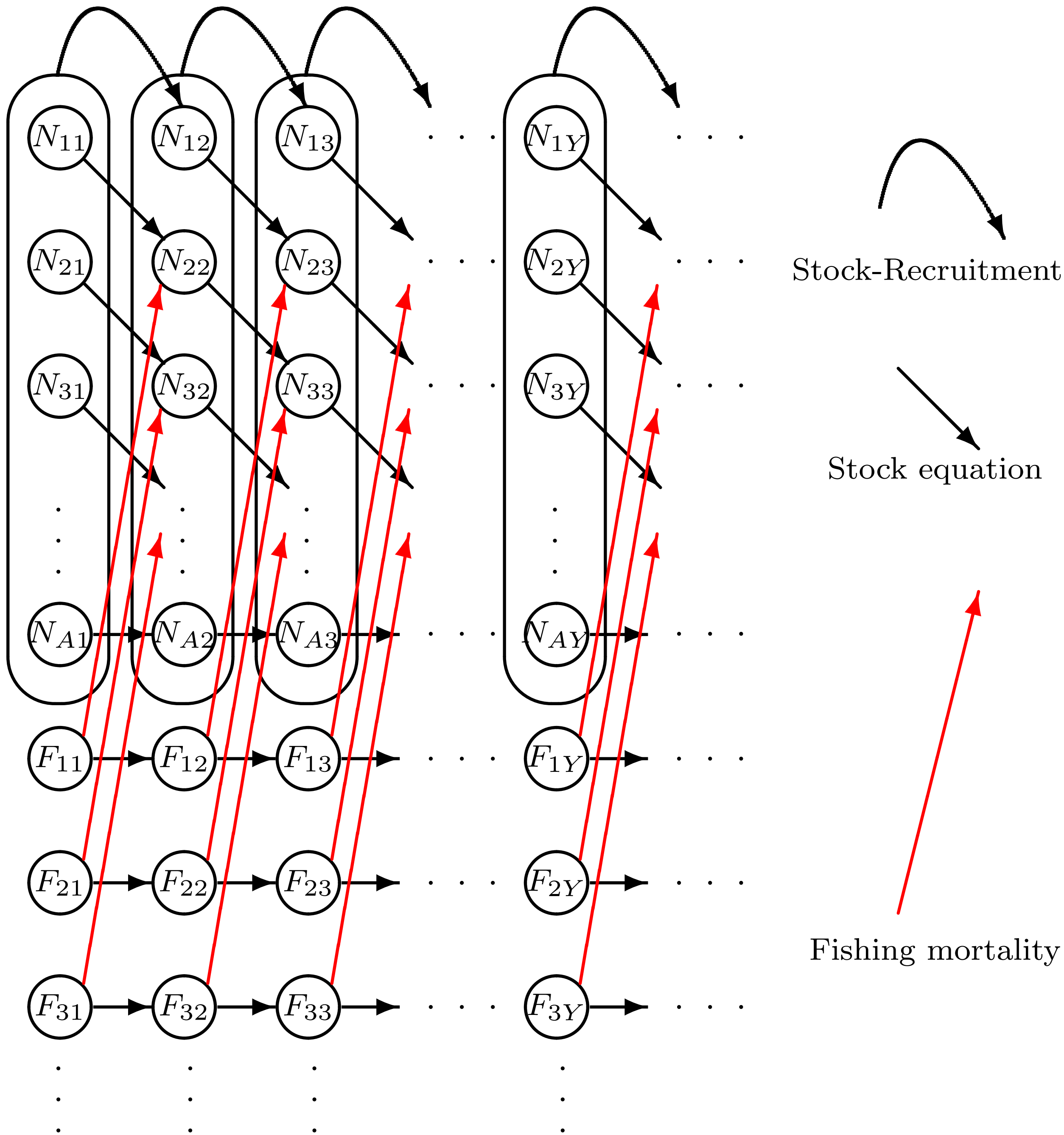
$$\begin{pmatrix} \log(C_y) \\ \log(I_y^{(s)}) \end{pmatrix} = O \begin{pmatrix} N_y \\ F_y \end{pmatrix} + \varepsilon_y$$

Model and parameters are what describes the distribution of states and observations through T, O, η_y, and ε_y.

Parameters: Survey catchabilities, S-R parameters, process and observation variances.

All model equation are as expected:

- Standard stock equation
- Standard stock recruitment (B-H, Ricker, or RW)
- Standard equations for total landings and survey indices



N_{11}
N_{12}
N_{13}
N_{1Y}
N_{21}
N_{22}
N_{23}
N_{2Y}
N_{31}
N_{32}
N_{33}
N_{3Y}
N_{A1}
N_{A2}
N_{A3}
N_{AY}
F_{11}
F_{12}
F_{13}
F_{1Y}
F_{21}
F_{22}
F_{23}
F_{2Y}
F_{31}
F_{32}
F_{33}
F_{3Y}
Stock-Recruitment
Stock equation
Fishing mortality

Numerical Methods

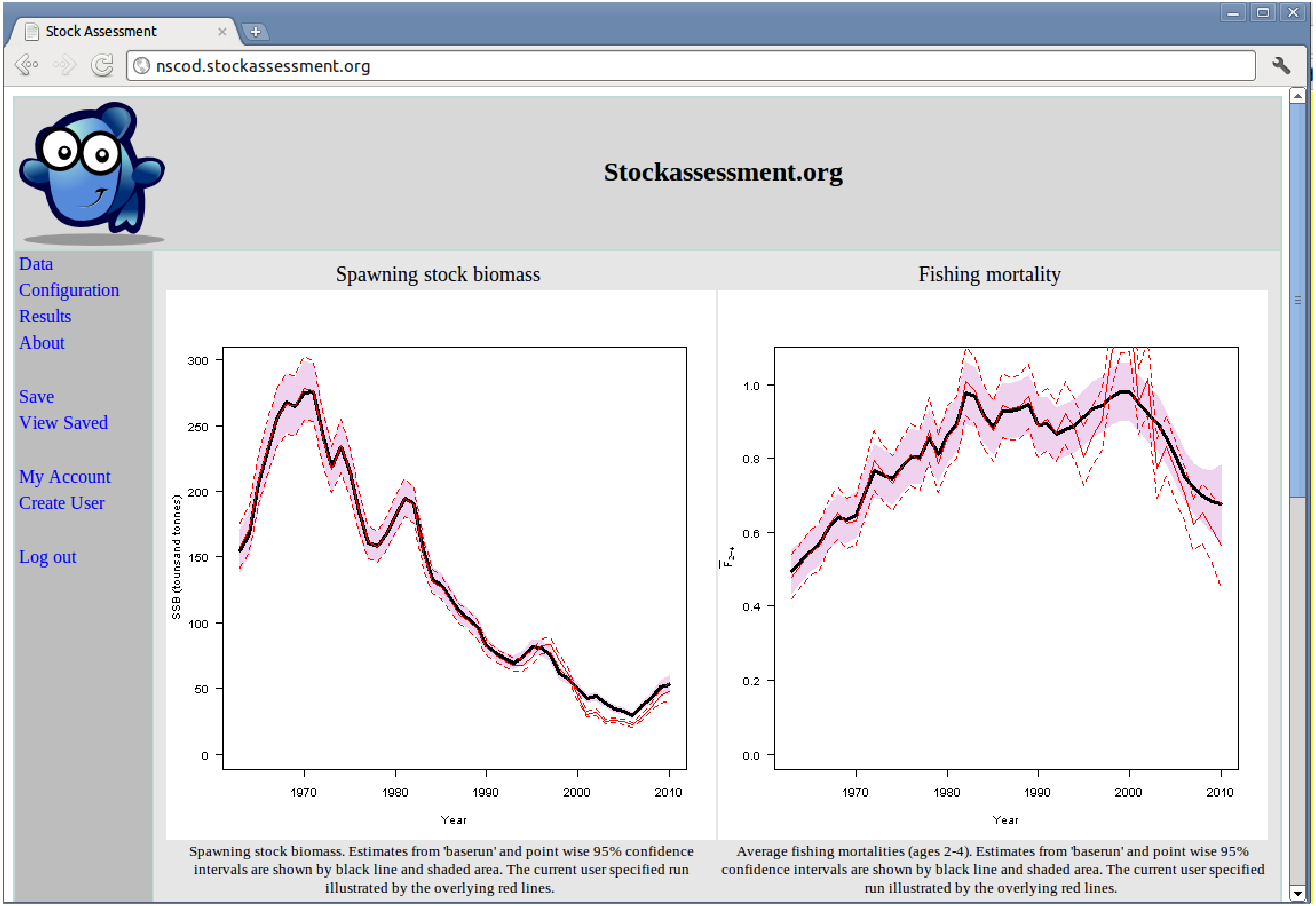
- Kalman Filter
- Extended Kalman Filter
- Unscented Kalman Filter
- Laplace approximation
- Sampling based methods

(Numerical methods are needed to calculate the marginal distribution)

Optimization is done using AD Model Builder

Features of the State-space assessment model

- Statistical model
  - Maximum likelihood estimation of model parameters
  - Estimation of uncertainties are an integrated part of the model
  - Prediction is straight-forward
- Consistent treatment of all $N_{a,y}$
- Allows selectivity to evolve
- Built-in (objective!) 'F-shrinkage' and 'tapered time weights'
- Nicely handles missing observations
- Room for additional features



Stock Assessment
nscod.stockassessment.org
Stockassessment.org
Data
Configuration
Results
About
Save
View Saved
My Account
Create User
Log out
Spawning stock biomass
Fishing mortality
SSB (tounsand tonnes)
300
250
200
150
100
50
0
1970
1980
1990
2000
2010
Year
$\bar{F}_{2-4}$
1.0
0.8
0.6
0.4
0.2
0.0
Spawning stock biomass. Estimates from 'baserun' and point wise 95% confidence intervals are shown by black line and shaded area. The current user specified run illustrated by the overlying red lines.
Average fishing mortalities (ages 2-4). Estimates from 'baserun' and point wise 95% confidence intervals are shown by black line and shaded area. The current user specified run illustrated by the overlying red lines.

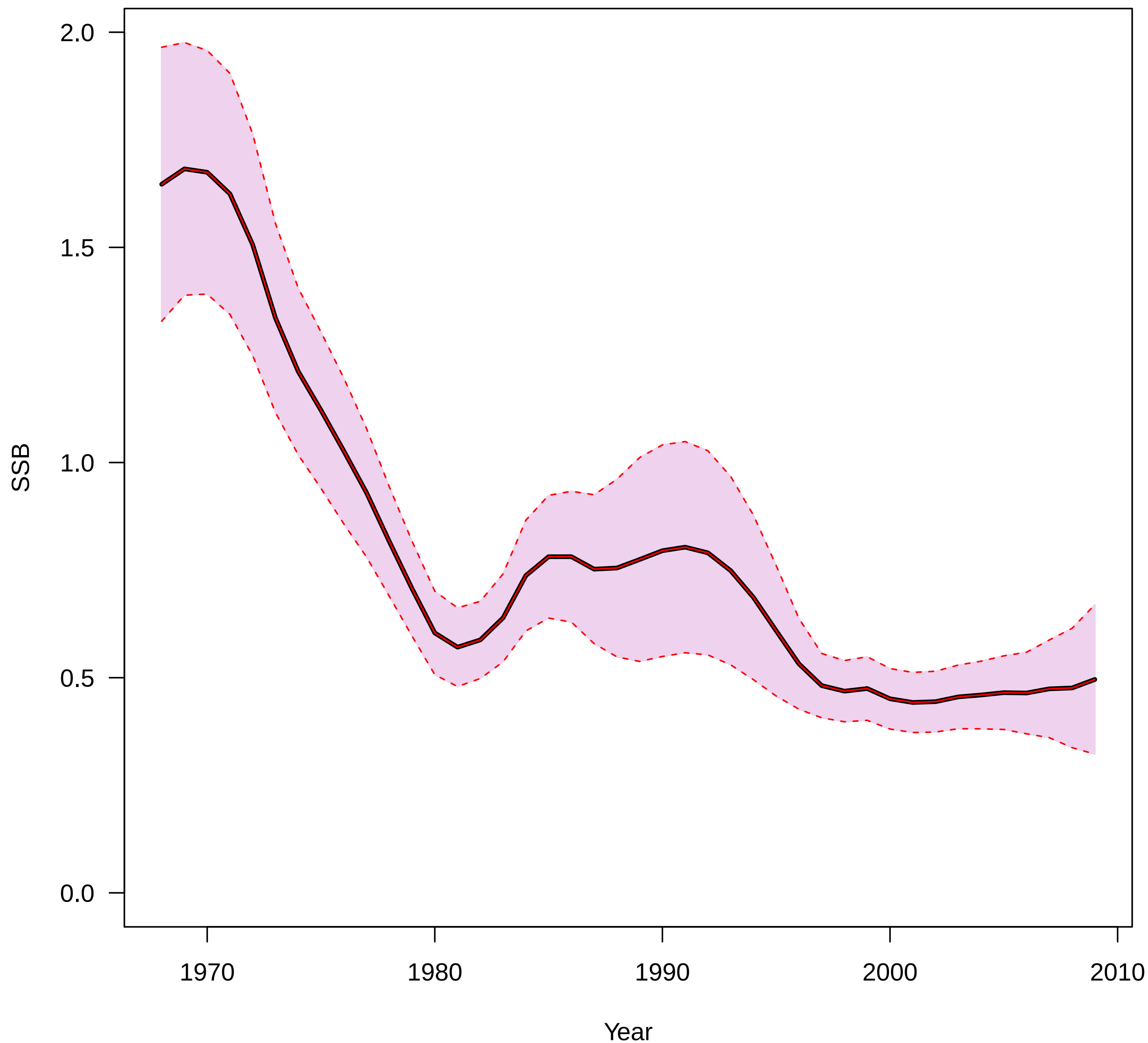
Web interface - Why?

- Scientific software is a way communicate ideas
- We want enable others to benefit from our work
- Peer review process is important
- Should be possible for all involved to:
  - see all details of the implementation
  - run it themselves
  - experiment with data
  - experiment with model assumptions
  - run the same version
- The interface makes it all one step easier

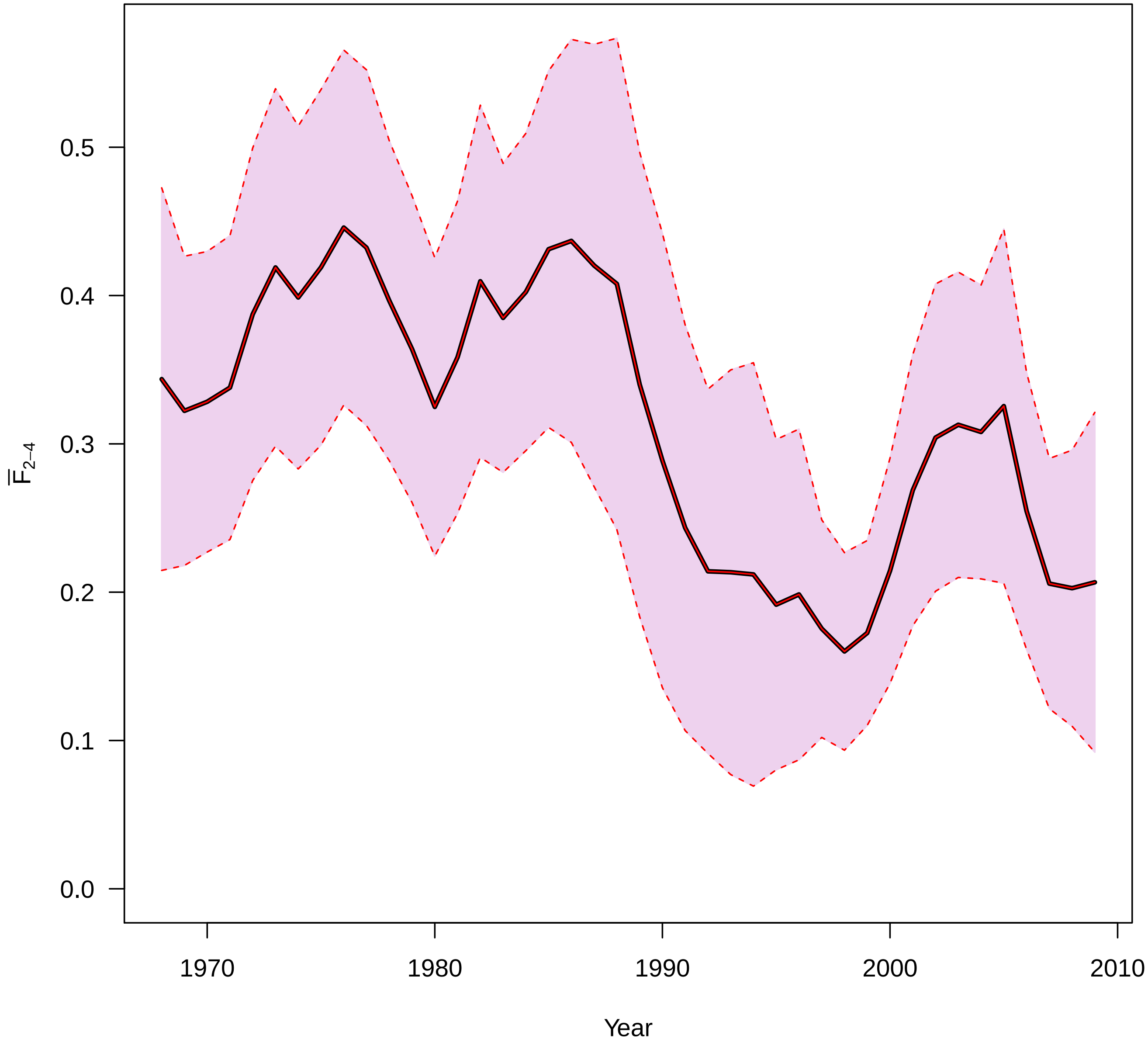
Status

- **Primary model in ICES for:**
  - Western Baltic Cod
  - Kattegat Cod
  - North Sea Cod
  - Skagerrak Sole
- **Exploratory model in ICES for:**
  - Eastern Baltic Cod
  - North Sea Sole
  - North Sea Herring
  - North Sea Haddock
  - Skagerrak Plaice
- **Quick unsystematic tests for some other stocks:**
  Western Baltic herring, 3PS Cod, 4VWX Herring, Greenland Halibut SA2+3KLMNO, American Plaice, Namibian Hake, Georges Bank Yellowtail Flounder, Bothnian Sea Herring, ...

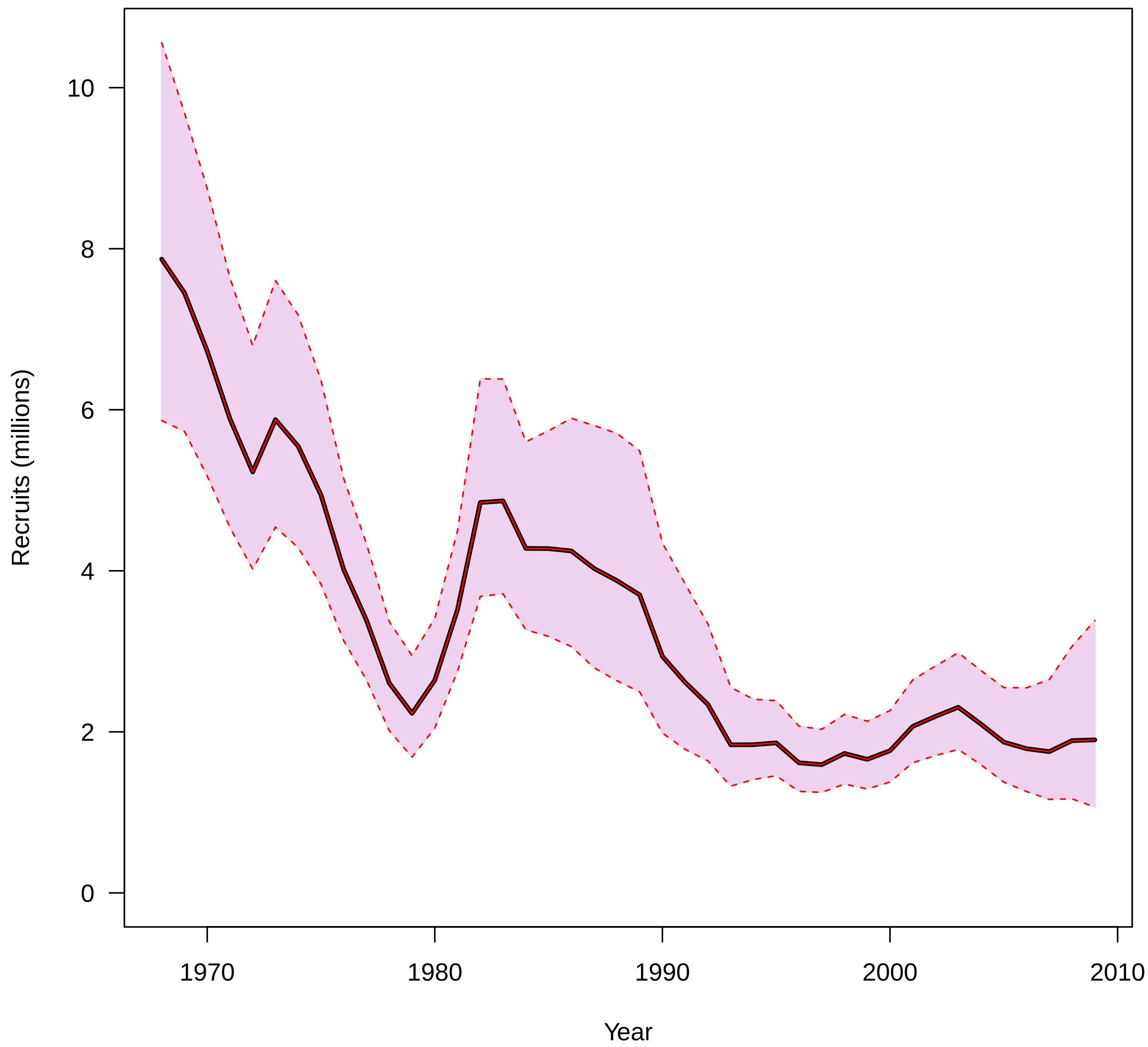
Preliminary results for Namibian Hake



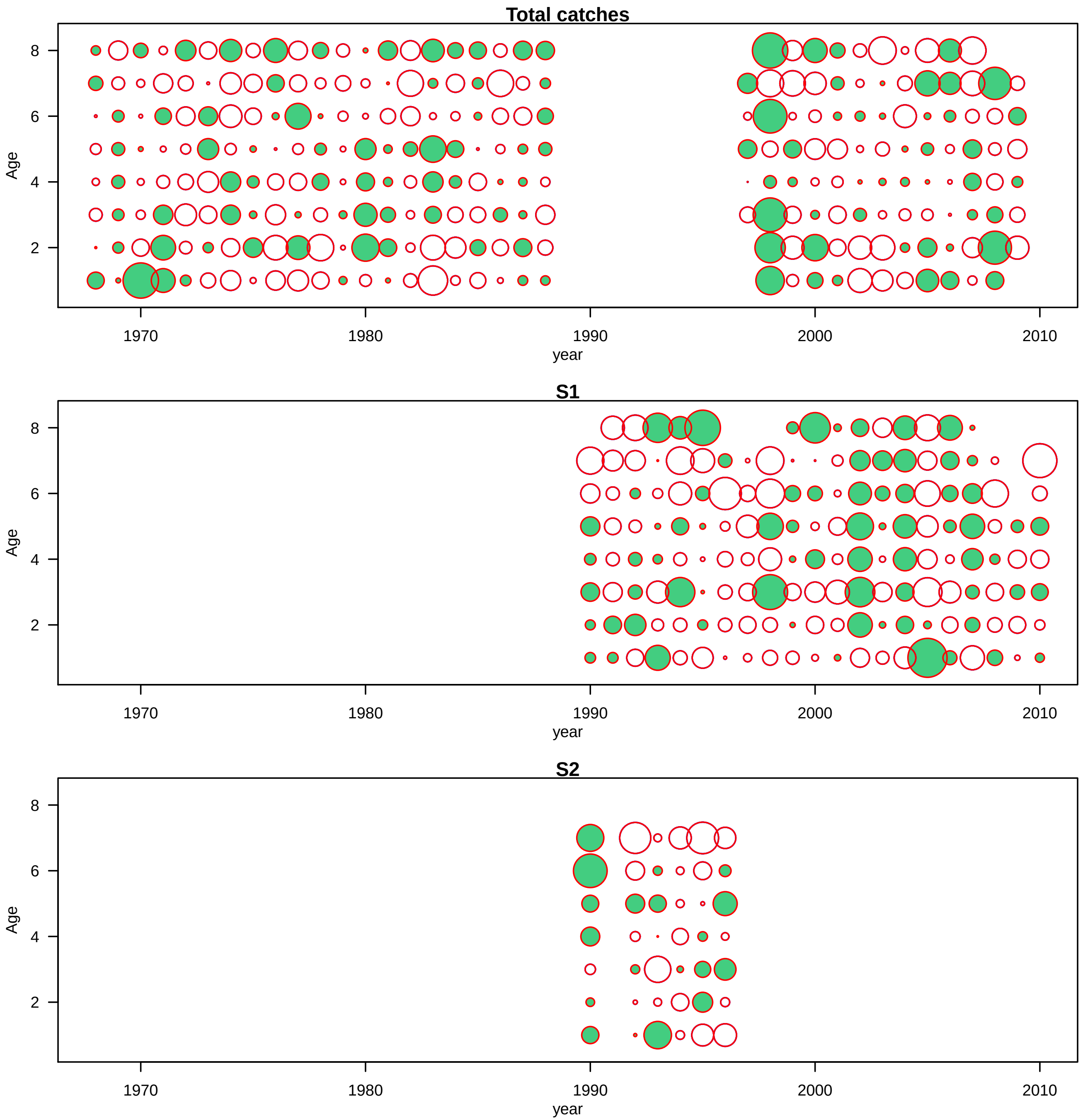
2.0
1.5
1.0
0.5
0.0
SSB
1970
1980
1990
2000
2010
Year



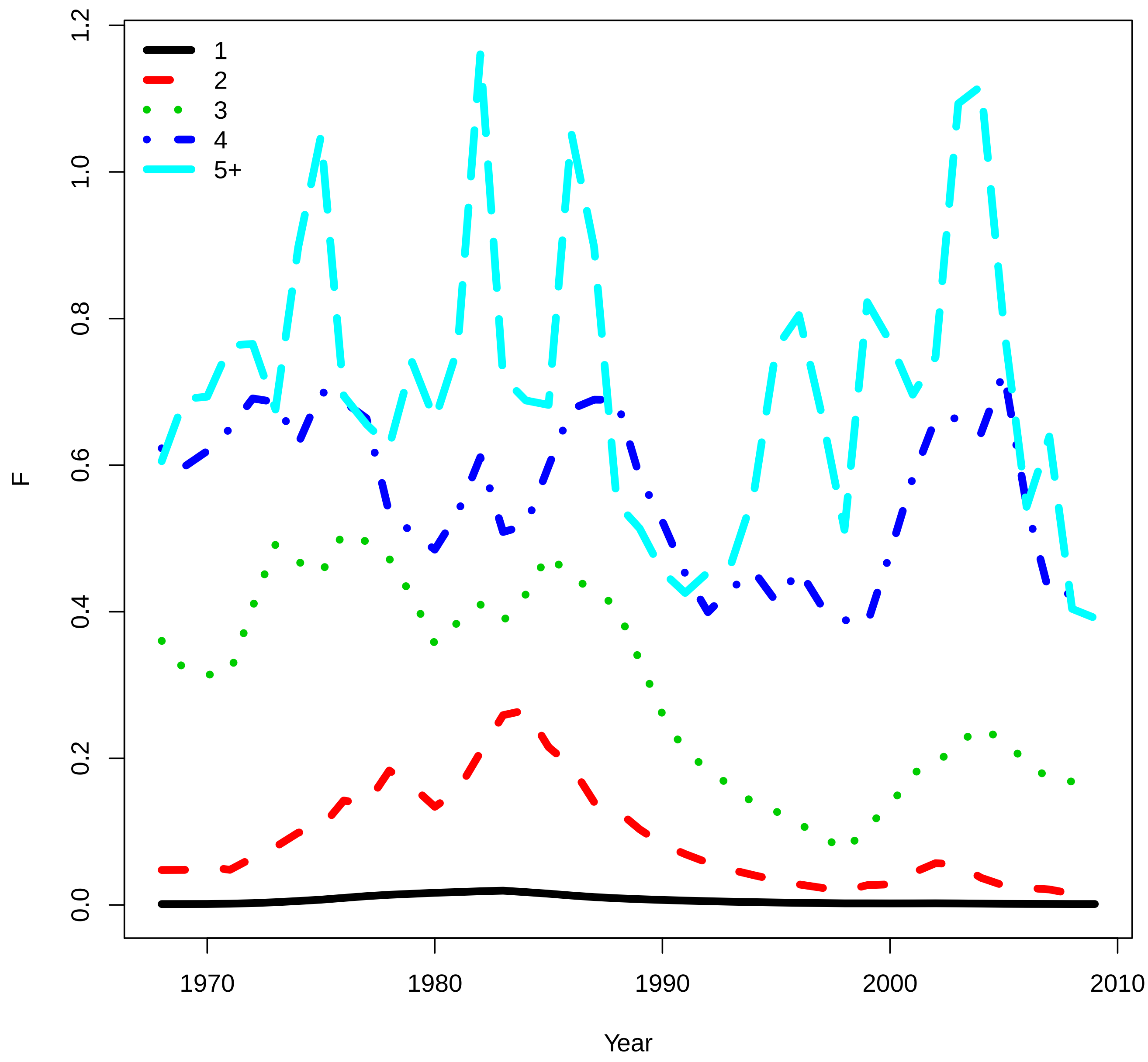
0.0
0.1
0.2
0.3
0.4
0.5
$\bar{F}_{2-4}$
1970
1980
1990
2000
2010
Year



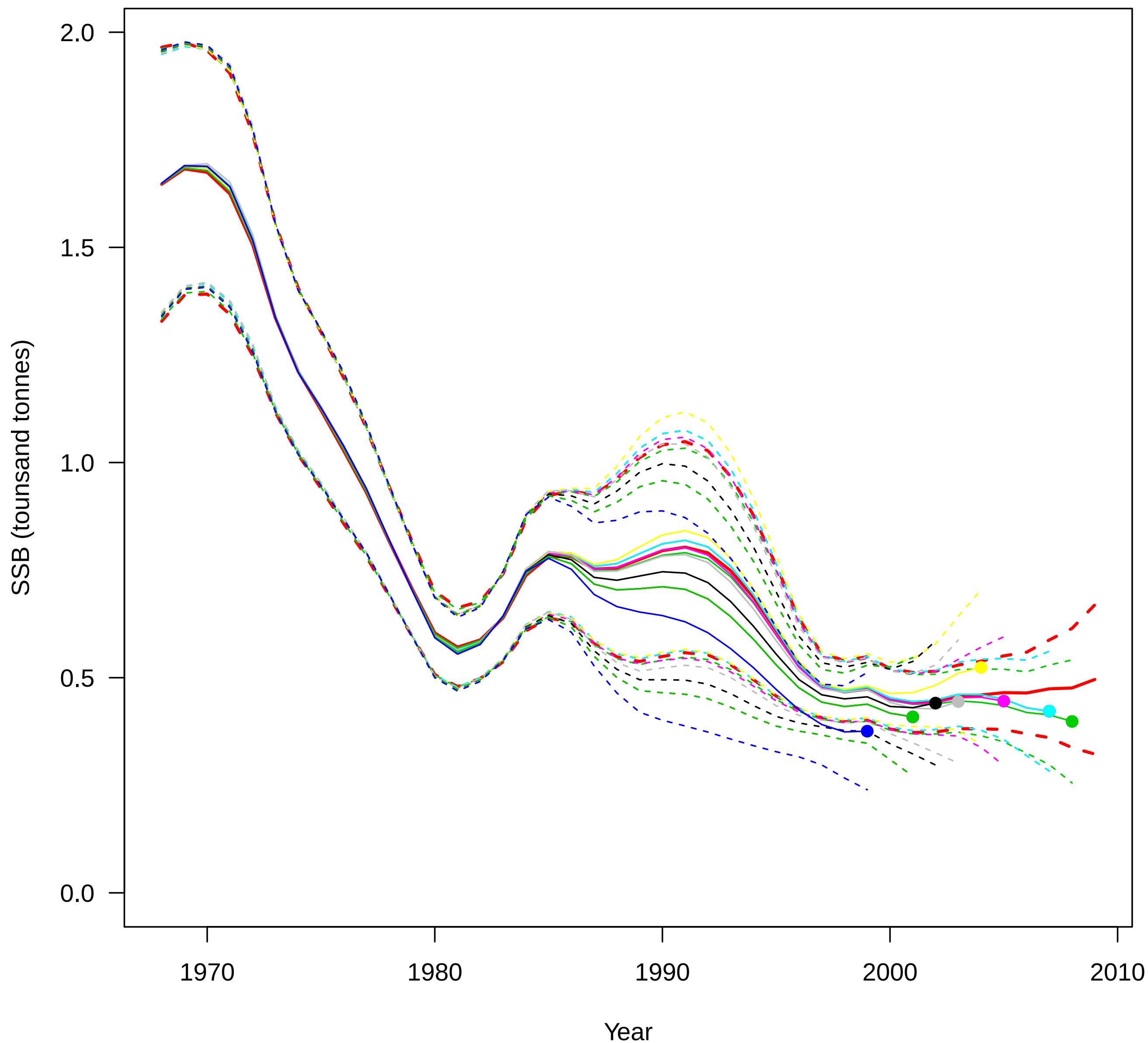
10
8
6
4
2
0
Recruits (millions)
1970
1980
1990
2000
2010
Year



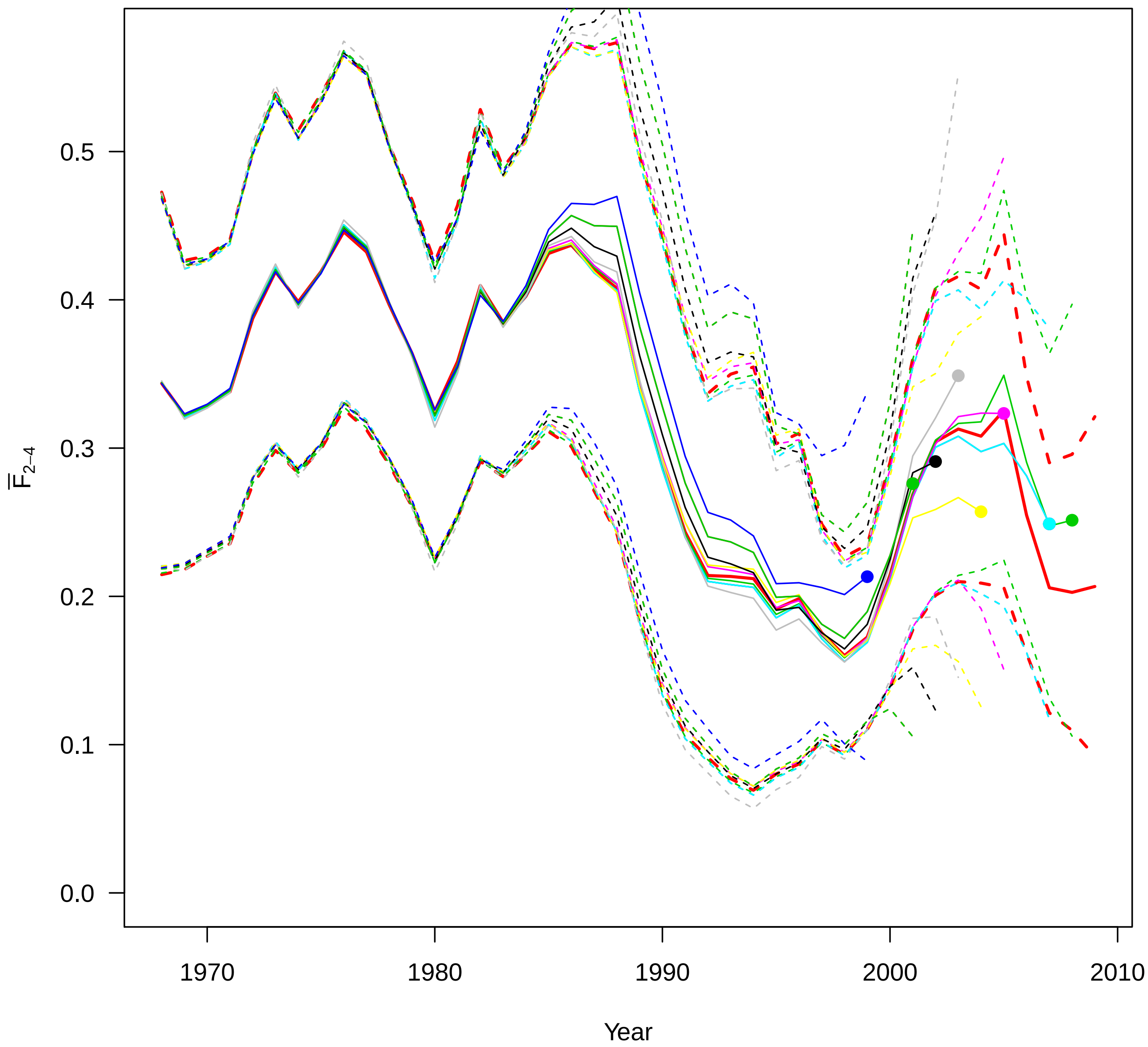
Total catches
S1
S2
Age
year
1970
1980
1990
2000
2010
2
4
6
8



1
2
3
4
5+
F
Year
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1970
1980
1990
2000
2010



SSB (tounsand tonnes)
2.0
1.5
1.0
0.5
0.0
1970
1980
1990
2000
2010
Year



0.0
0.1
0.2
0.3
0.4
0.5
$\bar{F}_{2-4}$
1970
1980
1990
2000
2010
Year