

# State-space models and WHAM Overview

**WHAM Workshop  
NEFSC Woods Hole Lab  
9-13 February 2026**

**Tim Miller and Emily Liljestrand  
Northeast Fisheries Science Center**



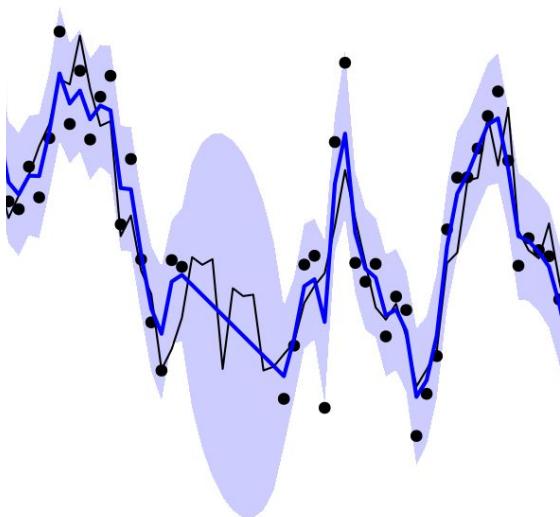
**NOAA  
FISHERIES**

# Outline

- State-space model overview
- WHAM model features
  - Random effects options
  - Environmental covariate effect options
  - Observation likelihood options
  - Biological Reference Point options
  - Projection options
  - Useful features: OSA residuals, auto-generated output, simulations

# What is a “state-space” model?

- $x_t$  **True process**
- $\hat{x}_t$  **Process estimate**
- $y_t$  **Observations**



$$x_t = \mu (1 - \rho) + \rho x_{t-1} + \varepsilon_t$$

$$\varepsilon_t \sim N(0, \sigma^2)$$

$$y_t = x_t + \delta_t$$

$$\delta_t \sim N(0, \tau^2)$$

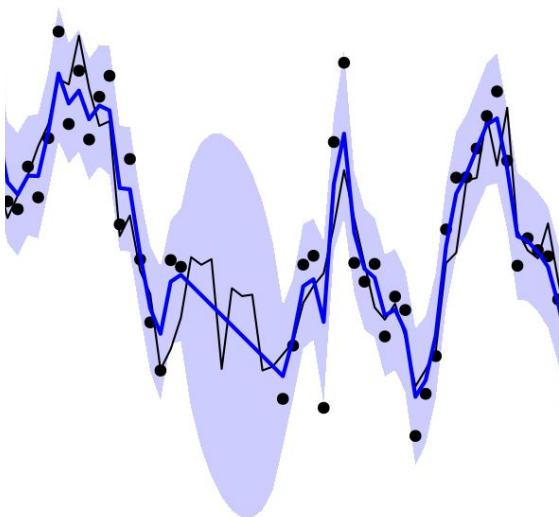
$$l(\boldsymbol{\theta}|\mathbf{y}) = \int \prod_{t=0}^{T-1} f(x_{t+1}|x_t, \boldsymbol{\theta}) g(y_{t+1}|x_{t+1}, \boldsymbol{\theta}) dx$$

$$\hat{\boldsymbol{\theta}} = \operatorname{argmax}_{\boldsymbol{\theta}} l(\boldsymbol{\theta}|\mathbf{y})$$

$$\hat{\mathbf{x}}|\hat{\boldsymbol{\theta}} = \operatorname{argmax}_{\mathbf{x}} \prod_{t=0}^{T-1} f(x_{t+1}|x_t, \hat{\boldsymbol{\theta}}) g(y_{t+1}|x_{t+1}, \hat{\boldsymbol{\theta}})$$

# What is a “state-space” model?

- $x_t$  **True process**
- $\hat{x}_t$  **Process estimate**
- $y_t$  **Observations**



$$x_t = \mu (1 - \rho) + \rho x_{t-1} + \varepsilon_t$$

$$\varepsilon_t \sim N(0, \sigma^2)$$

Process errors

$$y_t = x_t + \delta_t$$

$$\delta_t \sim N(0, \tau^2)$$

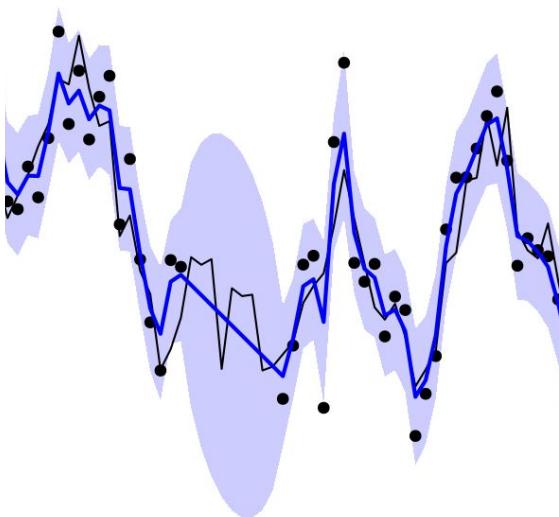
$$l(\boldsymbol{\theta}|\mathbf{y}) = \int \prod_{t=0}^{T-1} f(x_{t+1}|x_t, \boldsymbol{\theta}) g(y_{t+1}|x_{t+1}, \boldsymbol{\theta}) dx$$

$$\hat{\boldsymbol{\theta}} = \operatorname{argmax}_{\boldsymbol{\theta}} l(\boldsymbol{\theta}|\mathbf{y})$$

$$\hat{\mathbf{x}}|\hat{\boldsymbol{\theta}} = \operatorname{argmax}_{\mathbf{x}} \prod_{t=0}^{T-1} f(x_{t+1}|x_t, \hat{\boldsymbol{\theta}}) g(y_{t+1}|x_{t+1}, \hat{\boldsymbol{\theta}})$$

# What is a “state-space” model?

- $x_t$  **True process**
- $\hat{x}_t$  **Process estimate**
- $y_t$  **Observations**



$$x_t = \mu (1 - \rho) + \rho x_{t-1} + \varepsilon_t$$

$$\varepsilon_t \sim N(0, \sigma^2)$$

$$y_t = x_t + \delta_t$$

$$\delta_t \sim N(0, \tau^2)$$

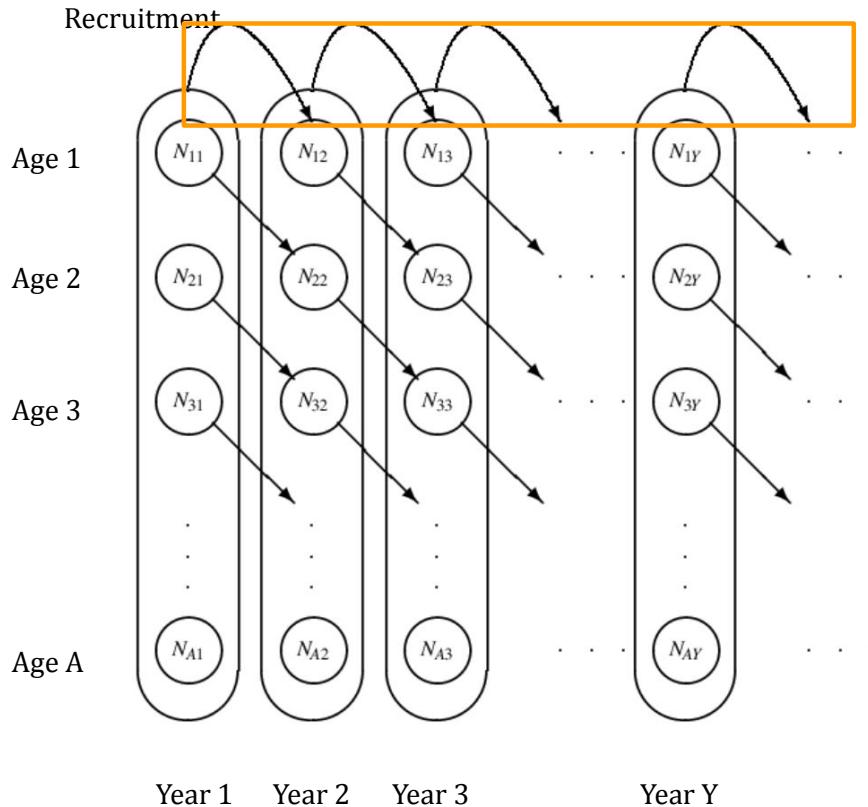
Observation errors

$$l(\boldsymbol{\theta}|\mathbf{y}) = \int \prod_{t=0}^{T-1} f(x_{t+1}|x_t, \boldsymbol{\theta}) g(y_{t+1}|x_{t+1}, \boldsymbol{\theta}) dx$$

$$\hat{\boldsymbol{\theta}} = \operatorname{argmax}_{\boldsymbol{\theta}} l(\boldsymbol{\theta}|\mathbf{y})$$

$$\hat{\mathbf{x}}|\hat{\boldsymbol{\theta}} = \operatorname{argmax}_{\mathbf{x}} \prod_{t=0}^{T-1} f(x_{t+1}|x_t, \hat{\boldsymbol{\theta}}) g(y_{t+1}|x_{t+1}, \hat{\boldsymbol{\theta}})$$

# Traditional age-structured model



## Age-structured population model

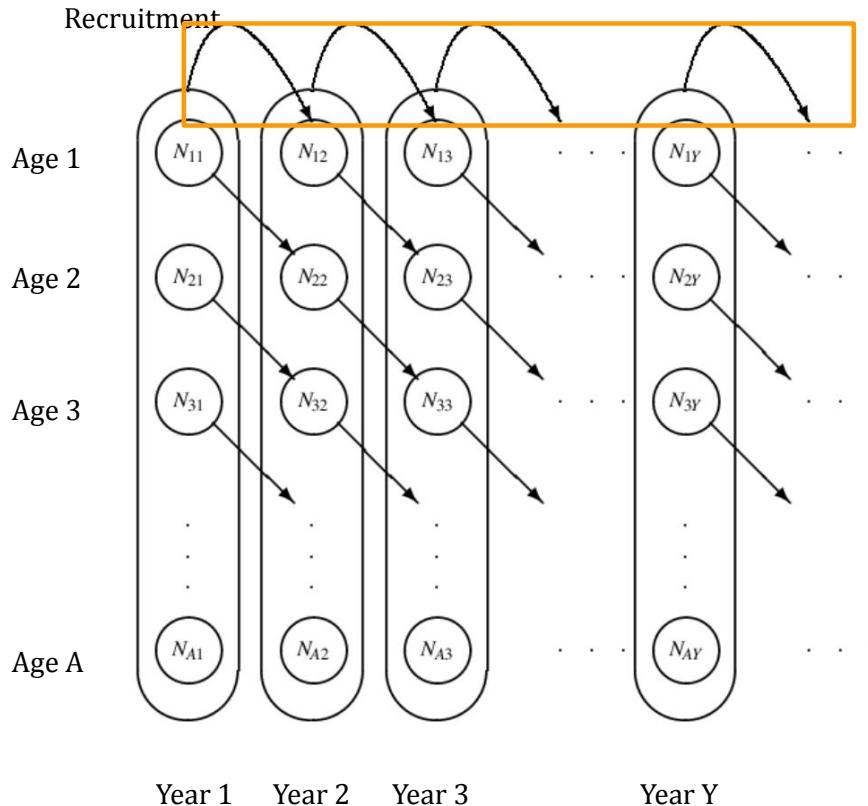
Keep track of the numbers of each age in each year,  $N_{a,y}$

**Recruitment may or may not be explicitly related to spawning population**

$$\log N_{1,y} = \log(f(\text{SSB}_{y-1})) + \varepsilon_{1,y}$$

**Annual recruitment or deviations historically estimated as fixed effects, possibly with a penalty.**

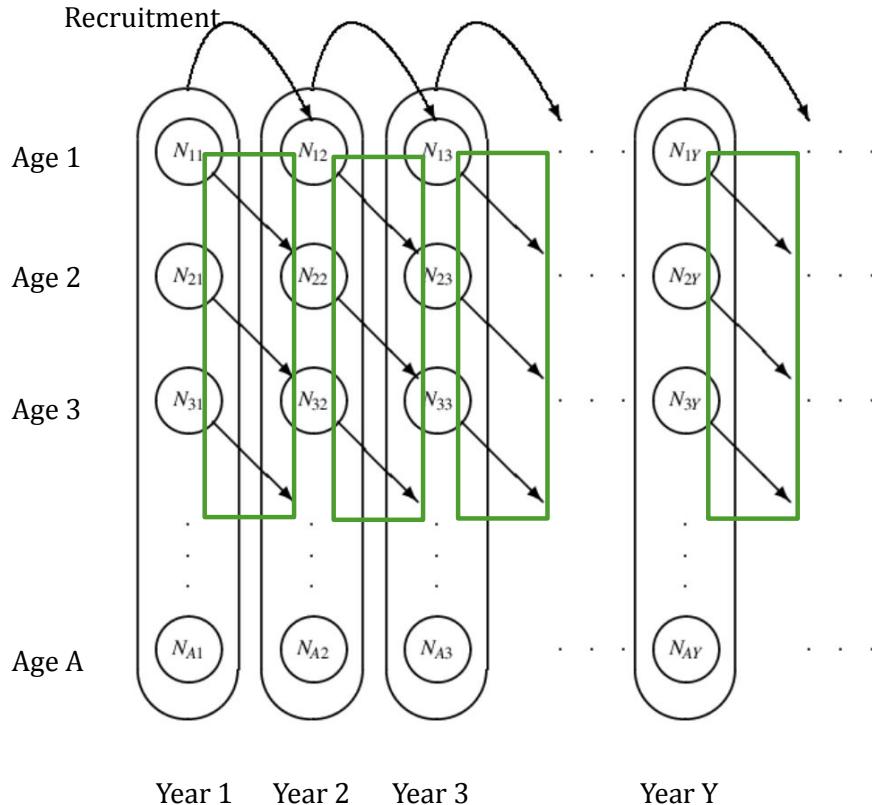
# Simplest state-space model



Treat recruitment or deviations as random effects with some distribution that are integrated out

$$\log N_{1,y} = \log(f(\text{SSB}_{y-1})) + \varepsilon_{1,y}$$

# Traditional age-structured model



**Transitions** between ages/years is a function of

- Natural mortality ( $M_{a,y}$ )
- Fishing mortality ( $F_{a,y}$ )
- Total mortality  $Z_{a,y} = F_{a,y} + M_{a,y}$

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

# Full state-space

$Z$  includes natural mortality and fishing mortality

$\varepsilon_{a,y}$  terms may account for

- various time-varying processes: selectivity, natural mortality, migration
- mis-specification of how mortality affects population

Other aspects of the assessment model can be modeled in similar ways

Statistical catch-at-age

$$\log N_{a,y} = f(\log N_{a-1,y-1})$$

Statistical catch-at-age, random recruitment

$$\log N_{1,y} = \log(f(\text{SSB}_{y-1})) + \varepsilon_{1,y}$$

“Full state-space”

$$\log N_{a,y} = \log(N_{a-1,y-1}) - Z_{a-1,y-1} + \varepsilon_{a,y}$$

$$l(\boldsymbol{\theta}|\mathbf{y}) = \int \prod_{t=0}^{T-1} f(\mathbf{N}_{t+1}|\mathbf{N}_t, \boldsymbol{\theta}) g(\mathbf{y}_{t+1}|\mathbf{N}_{t+1}, \boldsymbol{\theta}) d\mathbf{N}$$

# Pros and cons

## Pros

- Reduced retrospective patterns
- More realistic perception of uncertainty in assessment output
- Ability to better model temporal changes in productivity and fisheries
- Allows more statistical rigor in comparing alternative assessment models
- Short-term projections can be done consistently within the assessment model

## Cons

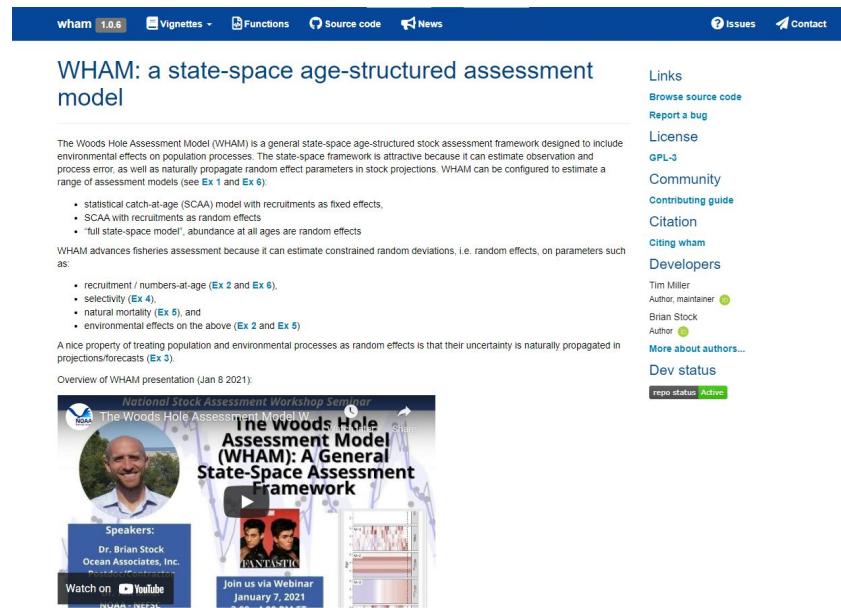
- Greater complexity in fitting and interpretation of results
- Research needed to evaluate
  - improvements over current assessment models
  - new diagnostics
  - reliability of particular configurations

# An open-source state-space assessment framework

- An R package available from Github
- Models can be completely configured using R package functionality
- Several tutorial vignettes
- Automatically produce a variety of output useful for both evaluating models and providing management advice.
- Tests to check package development.
- Several collaborators: **Brian Stock** (IMR) and others

The Woods Hole Assessment Model  
**WHAM:**

11



The screenshot shows the GitHub page for the WHAM repository. At the top, there's a navigation bar with links for 'wham 1.0.6', 'Vignettes', 'Functions', 'Source code', and 'News'. On the right, there are links for 'Issues' and 'Contact'. Below the navigation, the title 'WHAM: a state-space age-structured assessment model' is displayed. A paragraph describes the model as a general state-space age-structured stock assessment framework designed to include environmental effects on population processes. It mentions that WHAM can estimate observation and process error, as well as naturally propagate random effect parameters in stock projections. The text also notes that WHAM can be configured to estimate a range of assessment models (see [Ex 1](#) and [Ex 6](#)). A bulleted list follows, detailing:

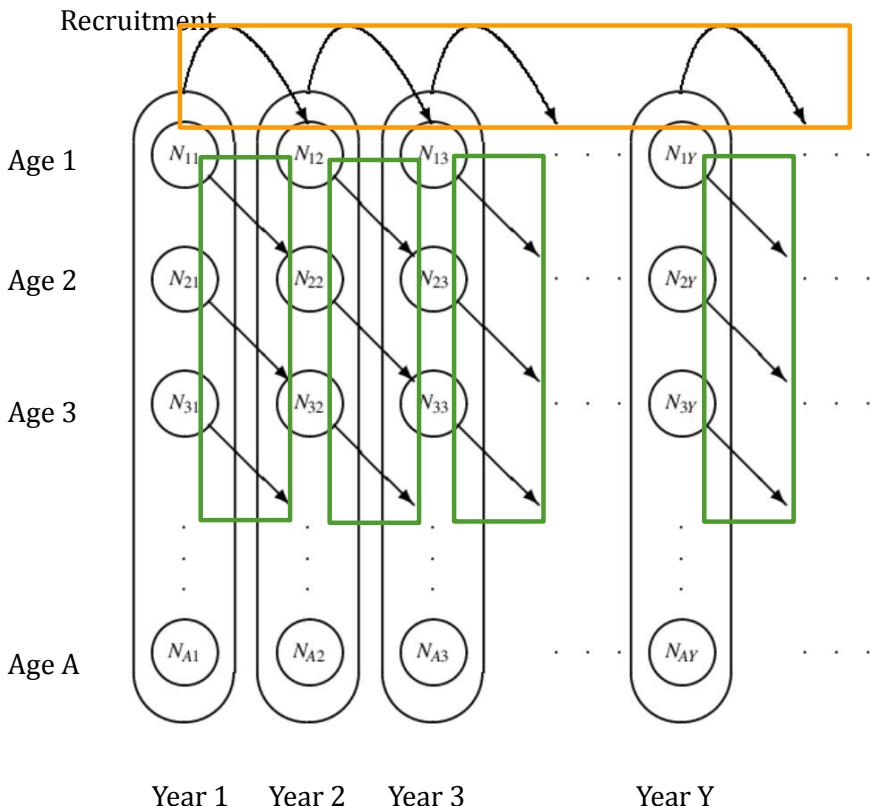
- statistical catch-at-age (SCAA) model with recruitments as fixed effects,
- SCAA with recruitments as random effects
- 'full state-space model', abundance at all ages are random effects

Further down, it says WHAM advances fisheries assessment because it can estimate constrained random deviations, i.e. random effects, on parameters such as:

- recruitment / numbers-at-age ([Ex 2](#) and [Ex 6](#)),
- selectivity ([Ex 4](#)),
- natural mortality ([Ex 5](#)), and
- environmental effects on the above ([Ex 2](#) and [Ex 5](#))

A note states that a nice property of treating population and environmental processes as random effects is that their uncertainty is naturally propagated in projections/forecasts ([Ex 3](#)). A link to an 'Overview of WHAM presentation (Jan 8 2021)' is provided. To the right, there's a sidebar with links for 'Links', 'License', 'Community', 'Contributing guide', 'Citation', 'Citing wham', 'Developers', 'More about authors...', 'Dev status', and 'repo status Active'. At the bottom, there's a section for the 'National Stock Assessment Workshop Seminar' featuring a video thumbnail for 'The Woods Hole Assessment Model (WHAM): A General State-Space Assessment Framework' and a speaker photo for Dr. Brian Stock.

# WHAM is an age-structured model



Configuration options for abundance at age:

- 1) Statistical catch-at-age (no random effects)

$$\log N_{a,y} = f(\log N_{a-1,y-1})$$

- 2) Statistical catch-at-age, random recruitment

$$\log N_{1,y} = \log(f(SSB_{y-1})) + \varepsilon_{1,y}$$

- 3) "Full state-space" (survival random effects)

$$\log N_{a,y} = \log(N_{a-1,y-1}) - Z_{a-1,y-1} + \varepsilon_{a,y}$$

# Random effects

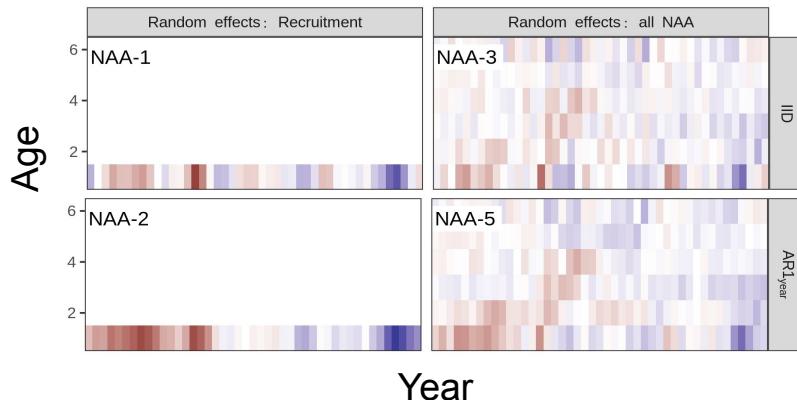
Options for alternative covariance structures (AR1, iid, etc)

- Recruitment (year)
- Interannual transitions in abundance at age ("survival") (year, age)
- Natural mortality (year, age)
- Selectivity (fishery or index) (year, age)
- Catchability (year)
- Hidden (imperfectly observed) environmental/climate variables (year)
- Movement (year,age)(development branch)
- *Growth (development branch)*

# Time- and age-varying processes

Biological processes are often  
correlated by year and age

- Recruitment
- Inter-annual transitions ("Survival")
- Natural mortality
- Selectivity
- Catchability
- Movement (development branch)



```
NAA_re = list(sigma="rec+1", cor="iid"))
```

Code	Description	Parameters
"none"	time-constant (no deviation)	
"iid"	independent, identically-distributed	$\sigma^2$
"ar1"	autoregressive-1 (correlated across ages/parameters)	$\sigma^2, \rho_a$
"ar1_y"	autoregressive-1 (correlated across years)	$\sigma^2, \rho_y$
"2dar1"	2D AR1 (correlated across both years and ages/parameters)	$\sigma^2, \rho_a, \rho_y$

$$\text{Cov}(\varepsilon_{a,y}, \varepsilon_{\tilde{a},\tilde{y}}) = \frac{\sigma_a \sigma_{\tilde{a}} \rho_a^{|a-\tilde{a}|} \rho_y^{|y-\tilde{y}|}}{(1 - \rho_a^2)(1 - \rho_y^2)}$$

# Time- and age-varying processes

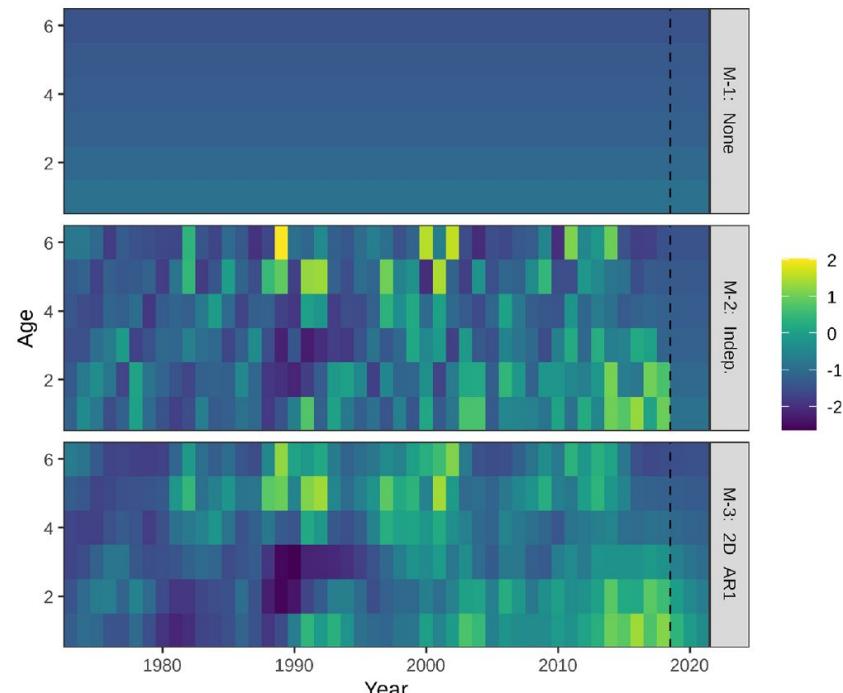
Biological processes are often **correlated by year and age**

- Recruitment
- Inter-annual transitions ("Survival")
- **Natural mortality**
- Selectivity
- Catchability
- Movement (development branch)

$\log(M)$  Gaussian random effects (iid, 2DAR1)

Estimate or fix mean M parameters:

- constant across ages
- age-specific
- function of weight-at-age

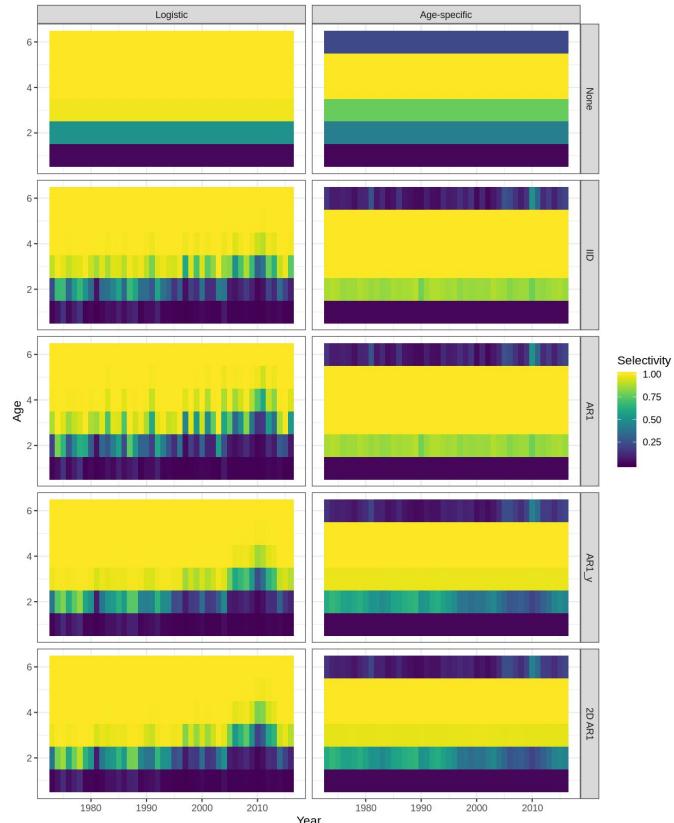


# Time- and age-varying processes

- Recruitment
- Inter-annual transitions ("Survival")
- Natural mortality
- **Selectivity**
- Catchability
- Movement (development branch)

"blocks" indexed to particular years of indices and fleets

- logistic (increasing or decreasing), double logistic, or age-specific
- constant, iid, or 1D or 2D AR1 processes for annual parameter values
- Gaussian on logit scale

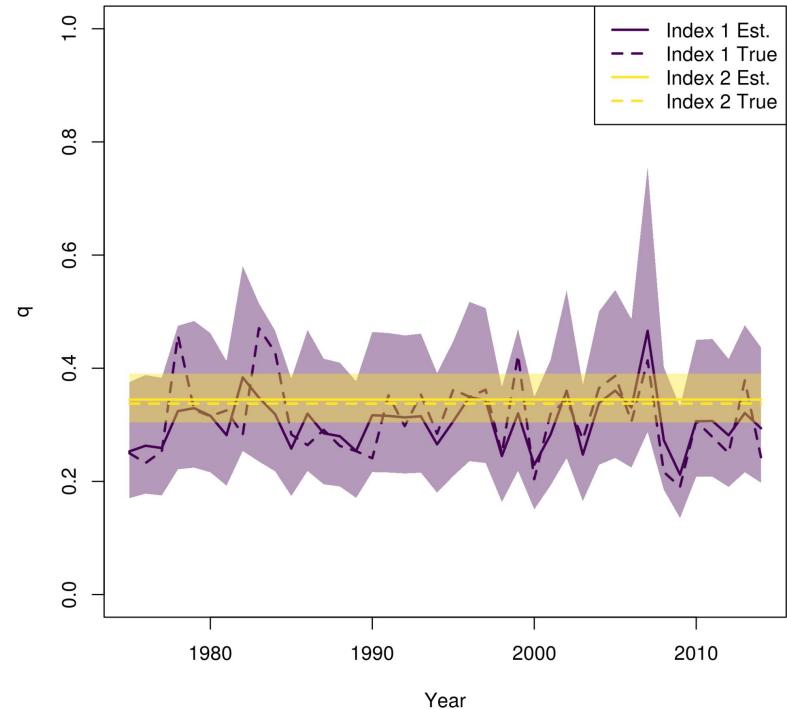


# Time- and age-varying processes

- Recruitment
- Inter-annual transitions ("Survival")
- Natural mortality
- Selectivity
- **Catchability**
- Movement (development branch)

Gaussian iid, or AR1 processes on logit scale

$$\log \left( \frac{q_y - b_l}{b_u - q_y} \right) = \mu_q + \epsilon_{q,y}$$



# Time- and age-varying processes

- Recruitment
- Inter-annual transitions ("Survival")
- Natural mortality
- Selectivity
- Catchability
- Movement (development branch)
  - Fixed effects (mean, variance, correlation parameters) are stock, region->region, season specific
  - random effects by year and/or age
  - movement can be modeled as
    - probabilities sequential to mortality
    - instantaneous rate simultaneous to mortality rates

$$f(\mu_{s,r,r',t,y,a}) = \theta_{s,r,r',t} + \epsilon_{s,r,r',t,y,a} \quad r \neq r'$$

$$Cov(\epsilon_{s,r,r',t,y,a}, \epsilon_{s,r,r',t,y',a'}) = \frac{\rho_{s,r,r',t,A}^{|a-a'|} \rho_{s,r,r',t,Y}^{|y-y'|} \sigma_{s,r,r',t}^2}{(1 - \rho_{s,r,r',t,A}^2) (1 - \rho_{s,r,r',t,Y}^2)}$$

additive logit transformation for probabilities sequential to survival:

$$f(\mu_{s,r,r',t,y,a}) = \log \left( \frac{\mu_{s,r,r',t,y,a}}{1 - \sum_{r'} \mu_{s,r,r',t,y,a}} \right)$$

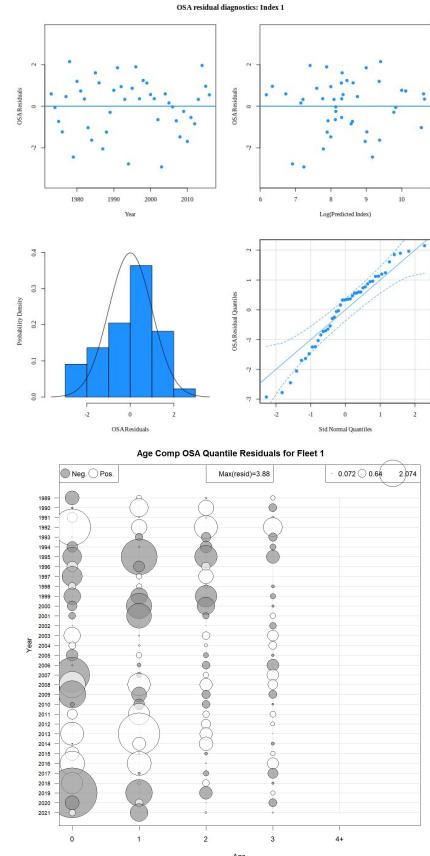
log transformation for instantaneous rates:

$$f(\mu_{s,r,r',t,y,a}) = \log(\mu_{s,r,r',t,y,a})$$

# Data components

All observations have error

- Aggregate catch (fleet-specific)
  - log-normal
- Catch age composition (fleet-specific)
  - Several likelihood options
- Aggregate indices (biomass or numbers)
  - log-normal
- Index age composition (biomass or numbers)
  - Several likelihood options
- **Optional:** Environmental/Climate observations
  - normal
- *Tagging data not yet included*



# State-space models for the covariate

- Imperfectly observed environmental variables can affect
  - Recruitment
  - Natural mortality (by age)
  - Index catchability
  - Movement (development)
- User-defined lag between covariate and population effect
- Effects options are “linear” or orthogonal polynomial
- Each covariate can have multiple effects
- Multiple covariates can be included

Covariate state-space models:

## 1. Random walk

$$\theta = (x_1, \sigma_x^2, \sigma_y^2)$$

$$x_t = x_{t-1} + \mathcal{N}(0, \sigma_x^2)$$

$$y_t = x_t + \mathcal{N}(0, \sigma_y^2)$$

## 2. AR1

$$-1 < \phi < 1$$

$$\theta = (\mu, \sigma_x^2, \sigma_y^2, \phi)$$

$$x_t = \mu + \phi x_{t-1} + \mathcal{N}(0, \sigma_x^2)$$

$$y_t = x_t + \mathcal{N}(0, \sigma_y^2)$$

# Environmental effects on...

## Recruitment models:

1. Random walk (No effects)

$$\hat{R}_t =$$

$$e^{\log R_{y-1} + \epsilon_y}$$

2. Mean (no SRR)

$$e^{\mu_R + \beta E_y + \epsilon_y}$$

3. Beverton-Holt

$$\frac{aS_{y-1}e^{\beta E_y + \epsilon_y}}{1 + bS_{y-1}}$$

Controlling

Limiting

Masking

4. Ricker

$$aS_{y-1}e^{-bS_{y-1} - \beta E_y + \epsilon_y}$$

Controlling

Masking

Iles & Beverton (1998)

## Catchability models:

linear in logit space

$$\log \left( \frac{q_y - b_l}{b_u - q_y} \right) = \mu_q + \beta E_y + \epsilon_{q,y}$$

## M models:

1. log-linear  $\log M_{y,a} = \mu_{M,a} + \beta_a E_y + \epsilon_{y,a}$

2. allometric  $\log M_{y,a} = \log(a) + b \log(W_{y,a}) + \beta_a E_y + \epsilon_{y,a}$

## Movement models:

linear in (additive) logit space or log-space

$$f(\mu_{s,r,r',t,y,a}) = \theta_{s,r,r',t} + \beta E_y + \epsilon_{s,r,r',t,y,a}$$

# Environmental and random effects on...

## Recruitment models:

1. Random walk (No effects)

$$\hat{R}_t =$$

$$e^{\log R_{y-1} + \epsilon_y}$$

2. Mean (no SRR)

$$e^{\mu_R + \beta E_y + \epsilon_y}$$

3. Beverton-Holt

$$\frac{aS_{y-1}e^{\beta E_y + \epsilon_y}}{1 + bS_{y-1}}$$

Controlling

Limiting

Masking

4. Ricker

$$aS_{y-1}e^{-bS_{y-1} + \beta E_y + \epsilon_y}$$

Controlling

Masking

Iles & Beverton (1998)

## Catchability models:

linear in logit space

$$\log \left( \frac{q_y - b_l}{b_u - q_y} \right) = \mu_q + \beta E_y + \epsilon_{q,y}$$

## M models:

1. log-linear  $\log M_{y,a} = \mu_{M,a} + \beta_a E_y + \epsilon_{y,a}$

2. allometric  $\log M_{y,a} = \log(a) + b \log(W_{y,a}) + \beta_a E_y + \epsilon_{y,a}$

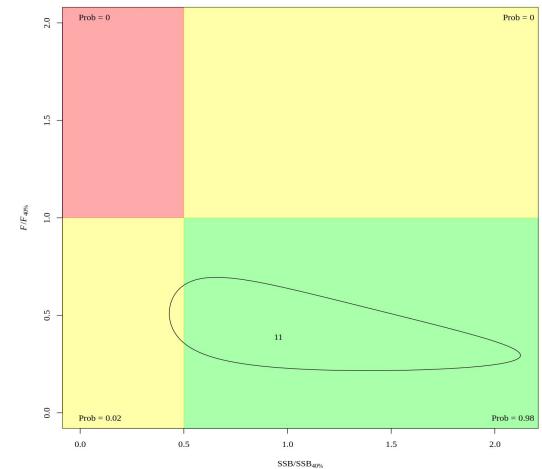
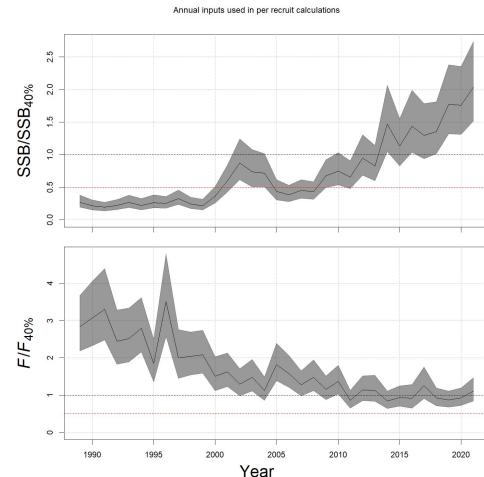
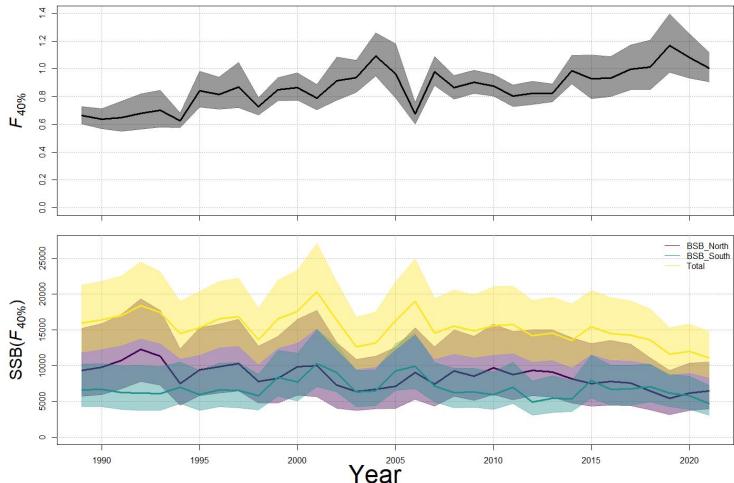
## Movement models:

linear in (additive) logit space or log-space

$$f(\mu_{s,r,r',t,y,a}) = \theta_{s,r,r',t} + \beta E_y + \epsilon_{s,r,r',t,y,a}$$

# Annual and prevailing BRPs and status

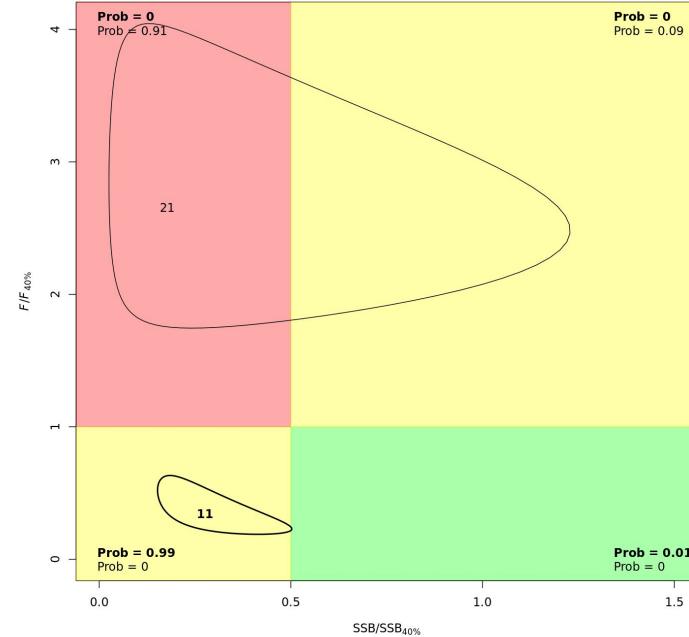
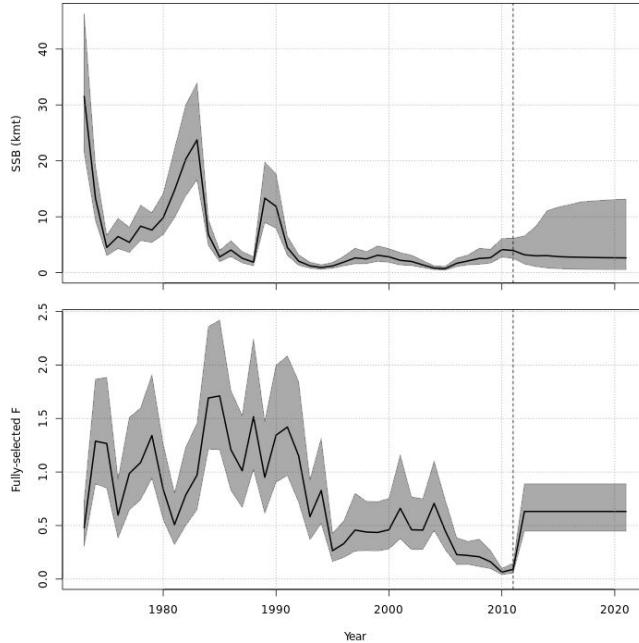
- Internally calculated reference points and status
- Allows uncertainty in parameters to be propagated



# Projections

Random effects (and uncertainty) can be projected

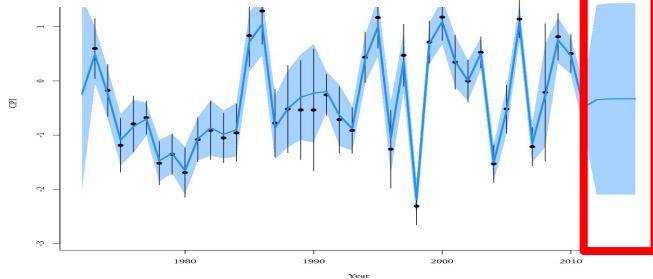
Can specify catch, status quo F, average F,  $F(X\%SPR)$ , FMSY



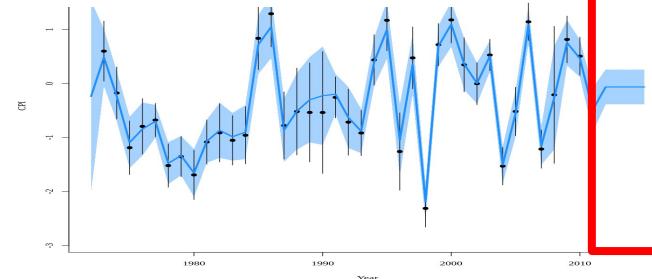
# Projections

Several options for treating environmental covariates

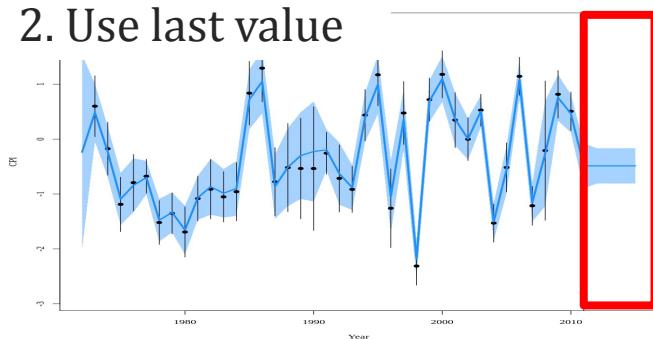
1. Continue RW/AR1



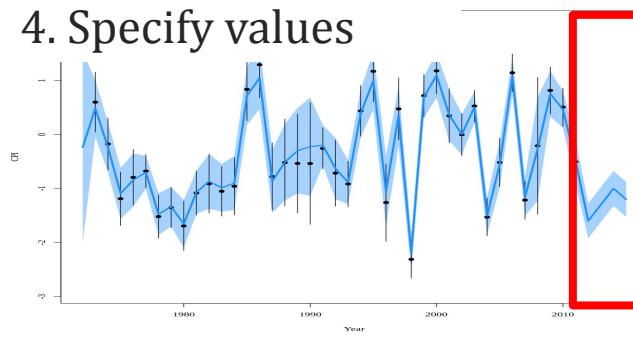
3. Use average value



2. Use last value



4. Specify values



# OSA residual diagnostics

## One step ahead (OSA) residuals

- provides independent residuals for correlated observations
- available for all observation types: aggregate catch and indices, age composition, environmental covariates

Environ Ecol Stat (2017) 24:317–339  
DOI 10.1007/s10651-017-0372-4

Fisheries Research 257 (2023) 106487

## Validation of ecological state space models using the Laplace approximation

Uffe Høgsbro Thygesen<sup>1</sup> · Christoffer Moesgaard Albertsen<sup>1</sup> ·  
Casper Willesøfte Berg<sup>1</sup> · Kasper Kristensen<sup>1</sup> · Anders Nielsen<sup>1</sup>

Contents lists available at ScienceDirect

Fisheries Research

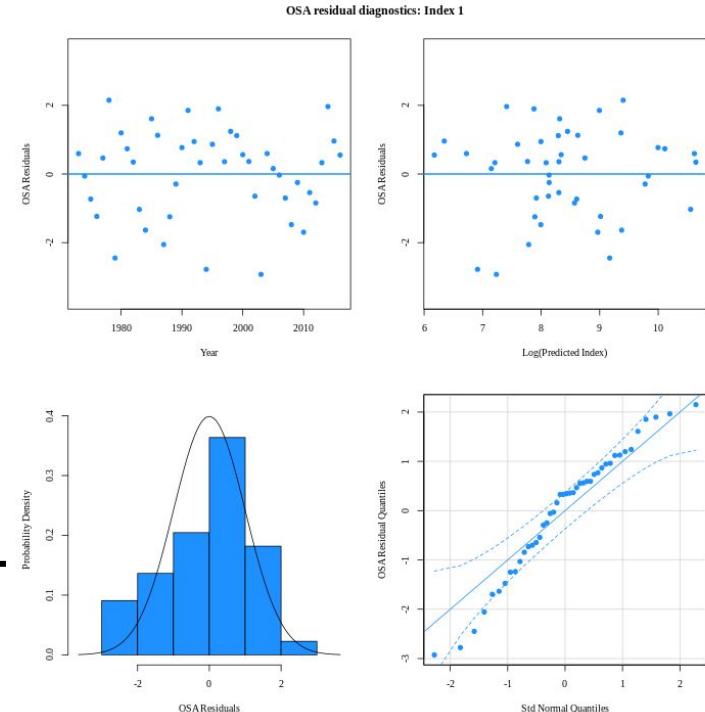
Journal homepage: [www.elsevier.com/locate/fishes](http://www.elsevier.com/locate/fishes)

## Model validation for compositional data in stock assessment models: Calculating residuals with correct properties

Vanessa Trijoulet <sup>a,\*</sup>, Christoffer Moesgaard Albertsen <sup>a</sup>, Kasper Kristensen <sup>a</sup>,  
Christopher M. Legault <sup>b</sup>, Timothy J. Miller <sup>b</sup>, Anders Nielsen <sup>a</sup>

<sup>a</sup> National Institute of Aquatic Resources, Technical University of Denmark, Kemitorvet 201, DK-2800 Kgs. Lyngby, Denmark

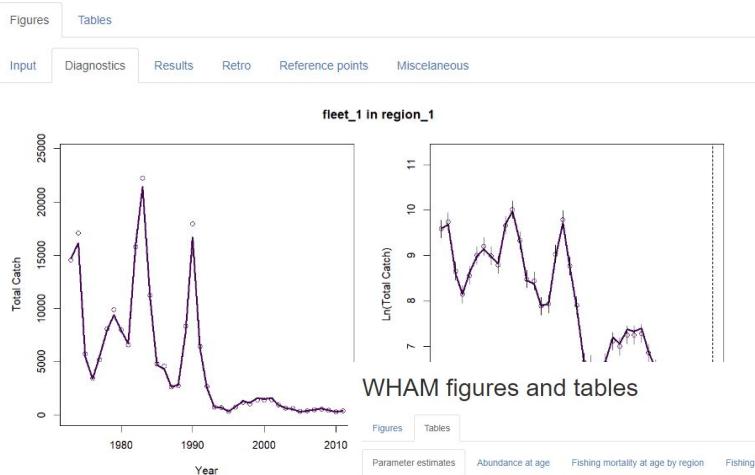
<sup>b</sup> Northeast Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 166 Water Street, Woods Hole, MA 02543, USA



# Automatically generated outputs

```
plot_wham_output(mod=m4, out.type='html')
```

## WHAM figures and tables



Parameter estimates, standard errors, and confidence intervals. Rounded to 3 decimal places.

	Estimate	Std. Error	95% CI lower	95% CI upper
index 1 fully selected q	2.441	0.272	1.962	3.037
index 2 fully selected q	0.490	0.038	0.421	0.571
index 3 fully selected q	0.297	0.042	0.225	0.393
index 4 fully selected q	0.006	0.001	0.005	0.009
index 5 fully selected q	0.051	0.008	0.037	0.070
Block 1: fleet 1 $a_{50}$	2.473	0.051	2.374	2.574
Block 1: fleet 1 slope (increasing)	0.417	0.017	0.385	0.452
Block 2: index 1 $a_{50}$	1.942	0.092	1.766	2.127

```
check_convergence(m1)
```

```
#> stats:nlminb thinks the model has converged: mod$opt$conve
#> Maximum gradient component: 1.01e-07
#> Max gradient parameter: log_F1
#> TMB:sdreport() was performed successfully for this model
```

```
res <- compare_wham_models(mods, fname=
```

```
#>          AIC rho_R rho_SSB rho_Fbar
#> m4 -1466.9 0.3610 0.0091 -0.0106
#> m2 -1172.7 3.1589 -0.0735 -0.0167
#> m3 4107.1 0.1287 0.0304 -0.0162
#> m1 4846.5 0.8207 0.1905 -0.2322
```

# Online Tutorials

The screenshot shows a web browser displaying the "wham 1.0.6" online tutorials. The top navigation bar includes links for "Vignettes", "Functions", "Source code", and "News". A dropdown menu from the "Vignettes" link lists 11 examples (Ex 1 to Ex 11) and a section on operating models. The main content area discusses WHAM's ability to estimate random effects and provides a list of four key applications.

wham 1.0.6

Vignettes ▾ Functions Source code News

Overview

Ex 1: The basics

Ex 2: Recruitment linked to an environmental covariate (Cold Pool Index)

Ex 3: Projecting / forecasting random effects

Ex 4: Selectivity with time- and age-varying random effects

Ex 5: Time-varying natural mortality linked to the Gulf Stream Index

Ex 6: Numbers-at-age / survival deviations as random effects

Ex 7: Debugging WHAM models

Ex 8: Compare ASAP and WHAM model results

Ex 9: Retrospective predictions

Ex 10: Operating models and MSE

Ex 11: Catchability configurations

WHAM advances fisheries assessment because it can estimate constrained random deviations, i.e. random effects, on parameters such as:

- recruitment / numbers-at-age ([Ex 2](#) and [Ex 6](#)),
- selectivity ([Ex 4](#)),
- natural mortality ([Ex 5](#)), and
- environmental effects on the above ([Ex 2](#) and [Ex 5](#))

A nice property of treating population and environmental processes as random effects is that their uncertainty is naturally propagated in projections/forecasts ([Ex 3](#)).

<https://timjmiller.github.io/wham/articles/index.html>

# Simulations, including MSE

## Operating model/MSE usage

- Can be used for simulating populations and data as well as estimation
- Used this way in Index-based Methods Research Track and state-space Research Track
- Used for testing reliability of models in stock-specific research tracks.

