

Do state-space assessment models perform better than traditional statistical  
catch-at-age models?

Timothy J. Miller, Anders Nielsen, Chris Legault, Ernesto Jardim, Olav  
Nikolai Breivik, Jonathan Deroba, Dan Hennen, Arni Magnusson, Vanessa  
Trijoulet, and ???

## Abstract

## Introduction

- General notions on importance of stock assessments for fisheries management
- Evolution of stock assessments: increase in complexity, single or multifleet, state-space vs. deterministic processes
- Numerous models used currently in fisheries management with different modelling assumptions
- Meta-analysis on 13 stocks to investigate how the models differ in performance according to their assumptions
- Quick description of methods: Fit 4 models to 13 stocks and compare the predictability of the models to conclude on the forecast performance

## Methods

- Quick presentation of models (refer mainly to papers)
- Table summarizing difference in assumptions between models (e.g. process error assumptions, measurement error assumptions, age-structured or no, fishing mortality model, etc.)
- Presentation of fish stocks (summary in a table and data in supplementary material?)

## Estimation models

We fit four assessment models to identical data sets for each stock. Two of the models are in the class of statistical catch at age models (ASAP and A4A) and two are state+space models (WHAM and SAM). ASAP is used extensively to assess US fish stocks in the north-west Atlantic ocean. SAM and A4A are used to assess stocks in Europe. WHAM is a more recently developed model that can be configured to estimate models ranging from statistical catch at age to state-space versions with options to treat various aspects as random effects. It has not been used in management but versions have been applied to stocks of yellowtail flounder (Miller et al. 2016), redfish (Miller and Hyun 2018), and Atlantic cod (Miller et al. 2018).

All models treat mortality processes use the Baranov catch equations which treat survival over the annual time interval as

### *WHAM models*

We fitted 4 configurations of the WHAM model to each stock data set. Two configurations are close to traditional statistical catch at age models with deterministic annual transitions in the cohorts. However, annual numbers recruit to the population are treated as independent lognormal random effects. The other two configurations are full state-space models that treat the annual cohort transitions as stochastic. The difference between the two types of each class of models is how the age composition observations are treated. In models 1 and 3

we assumed multinomial distributions with a default effective sample size. For models 2 and 4 we assumed a simple logit normal distribution similar to Miller et al. (2016) except that we treated any unobserved age classes as missing. We fit these models to determine whether there were generalities in model performance (AIC) or retrospective patterns across the stocks. The WHAM model with the lowest AIC was used for comparison with other model frameworks.

The only differences in model configuration between stocks was in configuring selectivity for the age composition observations for the abundance indices and the catch. For many of the stocks, the range of ages used to comprise the aggregate index varied among indices. This led to some difficulties in applying a default logistic function of age, so we often estimated age-specific parameters. The general approach was to initially estimate all age-specific parameters freely to see which ages exhibited peak selectivity. Then selectivity parameters at these ages were fixed at one for final model results. The initial models cannot be used for final results because typically there is confounding of fully selected catchability with other model parameters (e.g., catchability) when all selectivity parameters are free.

### *Parameter Estimation*

All model frameworks use some form of maximum likelihood estimation. ASAP is programmed in AD Model Builder (Fournier et al. 2012) and can be configured to use various penalties to the likelihood, but no penalties were used for any applications here. Estimation for the A4A model is programmed in R (R Core Team 2018) and is also likelihood-based. The WHAM and SAM models are programmed using the Template Model Builder (TMB) package in R (Kristensen et al. 2016) and parameters are estimated by maximizing a Laplace approximation of the marginal likelihood (Skaug and Fournier 2006). To use the TMB package, the joint log-likelihood is written as a C++ program which is compiled and accessed from R. The Laplace approximation of the marginal log-likelihood is returned when called from R and we use the “nlminb” function in R to minimize the negative of the Laplace approximation of the marginal log-likelihood. Empirical Bayes estimates of the state variables (random effects) are provided by the mode of posterior distributions of  $\mathbf{S}$ , conditioned on the fixed effects parameters.

### *Diagnostics*

To get independent residuals from a state-space model the so-called ‘one-observation-ahead’ residuals are computed. The residual for the  $n$ ’th observation is computed by using the first  $n - 1$  observations to predict the  $n$ ’th. Details can be found in Thygesen et al. (2017).

## **Results**

## **Discussion**

## References

- 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): 779–789.
- Berg, C. W. and Nielsen, A. 2016. Accounting for correlated observations in an age-based state-space stock assessment model. *ICES Journal of Marine Science* **73**(7): 1788–1797.
- Berg, C. W., Nielsen, A., and Kristensen, K. 2014. Evaluation of alternative age-based methods for estimating relative abundance from survey data in relation to assessment models. *Fisheries research* **151**: 91–99.
- Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M., Nielsen, A., and Sibert, J. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods and Software* **27**(2): 233–249.
- Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., and Bell, B. M. 2016. TMB: Automatic differentiation and Laplace approximation. *Journal of Statistical Software* **70**(5): 1–21.
- Miller, T. J., Hare, J. A., and Alade, L. 2016. A state-space approach to incorporating environmental effects on recruitment in an age-structured assessment model with an application to Southern New England yellowtail flounder. *Canadian Journal of Fisheries and Aquatic Sciences* **73**(8): 1261–1270, doi:10.1139/cjfas-2015-0339.
- Miller, T. J. and Hyun, S.-Y. 2018. Evaluating evidence for alternative natural mortality and process error assumptions using a state-space, age-structured assessment model. *Canadian Journal of Fisheries and Aquatic Sciences* **75**(5): 691–703, doi: 10.1139/cjfas-2017-0035.
- Miller, T. J., O’Brien, L., and Fratantoni, P. S. 2018. Temporal and environmental variation in growth and maturity and effects on management reference points of georges bank atlantic cod. *Canadian Journal of Fisheries and Aquatic Sciences* **75**(12): 2159–2171, doi: 10.1139/cjfas-2017-0124.
- Nielsen, A. and Berg, C. W. 2014. Estimation of time-varying selectivity in stock assessments using state-space models. *Fisheries Research* **158**: 96–101.
- R Core Team. 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Skaug, H. J. and Fournier, D. A. 2006. Automatic approximation of the marginal likelihood in non-Gaussian hierarchical models. *Computational Statistics & Data Analysis* **51**(2): 699–709.
- Thygesen, U. H., Albertsen, C. M., Berg, C. W., Kristensen, K., and Nielsen, A. 2017. Validation of ecological state space models using the laplace approximation. *Environmental and Ecological Statistics* **24**(2): 317–339.

## Appendix

### WHAM model description

The definitions for probability models describing stochastic changes in abundance at age from one year to another are identical to that given in Miller et al. (2016) and Miller and Hyun (2018). Log-abundance for ages and years greater than 1 are normally distributed conditional on the vector of numbers at age from the previous time step,

$$\log(N_{y,a}) | \mathbf{N}_{y-1} \sim N[f_a(\mathbf{N}_{y-1}), \sigma_{N,j}^2],$$

for  $y > 1$  where  $N(x, y)$  indicates a normal distribution with mean  $x$  and variance  $y$ ,

$$f_a(\mathbf{N}_{y-1}) = \begin{cases} g(N_{y-1,1}) & \text{for } a = 1, \\ \log(N_{y-1,a-1}e^{-Z_{y-1,a-1}}) & \text{for } 1 < a < A, \\ \log(N_{y-1,a-1}e^{-Z_{y-1,a-1}} + N_{y-1,a}e^{-Z_{y-1,a}}) & \text{for } a = A, \end{cases}$$

$Z_{y,a} = M_{y,a} + \sum_{f=1}^G F_{f,y,a}$  is the total mortality, and  $A$  indicates the terminal age class (i.e., the “plus group”). We assume two different variance parameters ( $\sigma_{N,j}$ ) for the abundance at age: one for the variance of annual deviations around mean log-recruitment  $\mu_s$  and one for inter-annual transitions of abundance at older ages. Define  $j = 1, 2$  for variance parameters for recruitment and older ages, respectively. Here we do not consider spawning biomass effects on recruitment, but Beverton-Holt and Ricker assumptions are configured options in the general model. We assume recruitment is a white noise process with  $g(N_{y-1,1}) = \mu$ . We treat the initial numbers at age are treated as fixed effects parameters

Annual fully-selected fishing mortality rates for fleet  $f$  were parameterized as (unpenalized) deviations from the previous year,

$$\log(F_{f,y+1}) = \log(F_{f,y}) + \delta_{f,y}$$

where  $y = 1, \dots, T-1$  and  $\delta_{f,y}$  are the inter-annual, deviations in log-fully selected fishing mortality.

To relate population abundance to observed relative abundance indices and corresponding age composition, we estimate a fully-selected catchability  $q_d$  for each index  $d$ . and a selectivity. Then year- and age-specific fishing mortality is parameterized by multiplying age-specific selectivity and annual fishing mortality,  $F_{f,s,y,a} = s_{f,s,y,a}F_{f,y}$ , where  $s_{f,s,y,a}$  is selectivity at age and sex for fishing fleet  $f$ . Similarly, we relate population abundance to observed relative abundance indices and corresponding age composition with a fully-selected catchability  $q_d$  for each index  $d$  and an age-specific selectivity  $s_{d,a}$ . For both fishing fleets and relative abundance indices, we have the same age-based logistic and age-specific parameterization options for selectivity models as the ASAP model.

Observed log-aggregate relative abundance indices from survey  $d$  are also normally distributed,

$$\log(I_{d,y}) \sim N\left[\log(\hat{I}_{d,y}), \sigma_{d,y}^2\right]$$

where observation error variances  $\sigma_{d,y}^2$  are assumed known. The predicted aggregated relative abundance index is just the sum of the predicted relative abundance index at each age,

$$\hat{I}_{d,y} = \sum_{a=1}^A \hat{I}_{d,y,a}.$$

where

$$\hat{I}_{d,y,a} = q_d s_{d,a} N_{y,a} e^{-Z_{y,a} \phi_d},$$

$q_d$  is the fully-selected catchability from survey  $d$ ,  $s_{d,a}$  is the selectivity from survey  $d$ , and  $\phi_d$  is the fraction of the year elapsed when survey  $d$  occurs.

Finally, the observed log-aggregate catches by fishing fleet  $f$  are also normally distributed,

$$\log(C_{f,y}) \sim N \left[ \log(\hat{C}_{f,y}), \tau_{f,y}^2 \right]$$

where observation error variances  $\tau_{f,y}^2$  are assumed to be known. The predicted catch at age given abundances at age is

$$(1) \quad \hat{C}_{f,y,a} = \frac{F_{f,y,a}}{Z_{y,a}} (1 - e^{-Z_{y,a}}) N_{y,a} W_{f,y,a}$$

where  $W_{f,y,a}$  is the weight at age in the catch by fleet  $f$ . The predicted aggregated annual catch is just the sum over ages

$$(2) \quad \hat{C}_{f,y} = \sum_{a=1}^A \hat{C}_{f,y,a}.$$

We assumed a multinomial distribution for the vector of frequencies at age in survey  $d$  in year  $y$ ,

$$(3) \quad \mathbf{n}_{d,y} = E_{d,y} \mathbf{p}_{d,y} \sim \text{Multinomial}(E_{d,y}, \hat{\mathbf{p}}_{d,y})$$

where  $E_{d,y}$  represented the sample size and the age-specific elements of the vector  $\hat{\mathbf{p}}_{d,y}$  are

$$\hat{p}_{d,y,a} = \frac{\hat{I}_{d,y,a}}{\hat{I}_{d,y}}.$$

Similarly, we assume a multinomial distribution with sample size  $E_{f,y}$  for the proportions at age in the catch,

$$\hat{p}_{f,y,a} = \frac{\tilde{C}_{f,y,a}}{\tilde{C}_{f,y}},$$

where  $\tilde{C}_{f,y,a}$  is the predicted number at age (i.e., Eq. 1 with  $W_{f,y,a} = 1$ ) and  $\tilde{C}_{f,y}$  is analogous to Eq. 2. However, there are options for other distributions for the age composition observations.

## The SAM state-space assessment model

The basic state-space assessment model (SAM) is described in Nielsen and Berg (2014). The model has been continuously developed and adapted for different stocks (e.g. to include tagging data and biomass indices). The current implementation, which is available at: <https://github.com/fishfollower/SAM>, is an R-package based on Template Model Builder (TMB) (Kristensen et al. 2016).

The model is a state-space model. The states  $\alpha$  are the log-transformed stock sizes  $\log N_1, \dots, \log N_A$  and fishing mortalities  $\log F_1, \dots, \log F_A$  corresponding to total age specific catches. It is often assumed that some age classes have the same fishing mortality. In any

given year  $y$  the state is the combined vector  $\alpha_y = (\log N_1, \dots, \log N_A, \log F_1, \dots, \log F_A)'$ . The transition equation describes the distribution of the next years state from a given state in the current year. The following is assumed:

$$\alpha_y = T(\alpha_{y-1}) + \eta_y$$

The transition function  $T$  is where the stock equation and assumptions about stock–recruitment enters the model. The equations are:

$$\begin{aligned} \log N_{1,y} &= SR(\log(N_{y-1})) \\ \log N_{a,y} &= \log N_{a-1,y-1} - F_{a-1,y-1} - M_{a-1,y-1} , & 2 \leq a < A \\ \log N_{A,y} &= \log(N_{A-1,y-1} \exp^{-F_{A-1,y-1}-M_{A-1,y-1}} + N_{A,y-1} \exp^{-F_{A,y-1}-M_{A,y-1}}) \\ \log F_{a,y} &= \log F_{a,y-1} , & 1 \leq a \leq A \end{aligned}$$

Here  $M_{a,y}$  is the age and year specific natural mortality parameter, which is assumed known from outside sources.  $F_{a,y}$  is the total fishing mortality. The function ‘ $SR$ ’ is the stock–recruitment relationship assumed (options are: plain random walk on logarithmic scale, Ricker, and Beverton-Holt).

The process noise  $\eta$  is assumed to be Gaussian with zero mean, and separate variance parameters. Typically one for recruitment ( $\sigma_{N_{a=1}}^2$ ), one for survival ( $\sigma_{N_{a>1}}^2$ ), one for fishing mortality at age ( $\sigma_F^2$ ), but these can be flexibly configured. The  $N$ -part of  $\eta$  is assumed uncorrelated, and the  $F$ -part can be assumed correlated according to an AR(1) correlation structure, such that  $\text{cor}(\Delta \log(F_{a,y}), \Delta \log(F_{\tilde{a},y})) = \rho^{|a-\tilde{a}|}$ .

The observation part of the state–space model describes the distribution of the observations for a given state  $\alpha_y$ . Here the vector of all observations from a given year  $y$  is denoted  $x_y$ . The elements of  $x_y$  are age-specific log-catches  $\log C_{a,y}$  and age-specific log-indices from scientific surveys  $\log I_{a,y}^{(s)}$ . The combined observation equation is:

$$x_y = O(\alpha_y) + \varepsilon_y$$

The observation function  $O$  consists of the catch equations for total catches and scientific surveys. The measurement noise term  $\varepsilon_y$  is assumed to be Gaussian. An expanded view of the observation equation becomes:

$$\begin{aligned} \log C_{a,y} &= \log \left( \frac{F_{a,y}}{Z_{a,y}} (1 - e^{-Z_{a,y}}) N_{a,y} \right) + \varepsilon_{a,y}^{(c)} \\ \log I_{a,y}^{(s)} &= \log \left( Q_a^{(s)} e^{-Z_{a,y} \frac{D^{(s)}}{365}} N_{a,y} \right) + \varepsilon_{a,y}^{(s)} \end{aligned}$$

Here  $Z$  is the total mortality rate  $Z_{a,y} = M_{a,y} + F_{a,y}$ ,  $D^{(s)}$  is the number of days into the year where the survey  $s$  is conducted,  $Q_a^{(s)}$  are model parameters describing catchability coefficients. The variance of  $\varepsilon_y$  is setup such that each data source catches, and the four scientific surveys have their own covariance matrix.

Observation uncertainty is important e.g. to get the relative weighting of the different information sources correct, so a lot of effort has been invested in getting the optimal options



into SAM. In Berg and Nielsen (2016) different covariance structures are compared for four ICES stocks. It was found that irregular lattice AR(1) observation correlation structure was optimal for surveys. The covariance structures tested were inspired by a previous study (Berg et al. 2014) of the structures obtained from survey calculations. In the paper Albertsen et al. (2016) 13 different observational likelihood formulations were evaluated for four ICES stocks. It was found that the multivariate log-normal representation was among the optimal in all four cases.

To describe the options available consider a yearly vector  $C_y = (C_{a=1,y}, \dots, C_{a=A,y})$  of age specific observations from a fleet (survey or commercial). Assume first that the  $\log(C_y)$  is multivariate Gaussian:

$$\log(C_y) \sim N(\log(\hat{C}_y), \Sigma)$$

where  $\Sigma$  is the covariance matrix, and  $\hat{C}_y$  is the vector of the usual model predictions. The covariance matrix is specified from a vector of standard deviations  $\sigma = (\sigma_1 \dots \sigma_A)$  and a correlation matrix  $\rho$  (by  $\Sigma_{a\tilde{a}} = \sigma_a \sigma_{\tilde{a}} \rho_{a\tilde{a}}$ ). Four options are available for the correlation  $\rho$ : Independent ( $\rho = I$ ), auto-regressive of order 1 ( $\rho_{a\tilde{a}} = 0.5^{\theta|a-\tilde{a}|}$ ,  $\theta > 0$ )\*, irregular auto-regressive of order 1 ( $\rho_{a\tilde{a}} = 0.5^{|\theta_a - \theta_{\tilde{a}}|}$ ,  $\theta_1 = 0 \leq \theta_2 \leq \dots \leq \theta_A$ ), and unstructured (parameterized by the Cholesky of  $\rho$ ). The options for covariance structure can be set for each fleet individually. In addition it is also possible to supply external weights for each individual observation. This option can be used in two ways. To set the relative weighting, or to actually set the fixed variance of each individual observation.

### *Likelihood and approximation*

The likelihood function for this is set up by first defining the joint likelihood of both random effects (here collected in the  $\alpha_y$  states), and the observations (here collected in the  $x_y$  vectors). The joint likelihood is:

$$L(\theta, \alpha, x) = \prod_{y=2}^Y \{\phi(\alpha_y - T(\alpha_{y-1}), \Sigma_\eta)\} \prod_{y=1}^Y \{\phi(x_y - O(\alpha_y), \Sigma_\varepsilon)\}$$

Here  $\theta$  is a vector of model parameters. Since the random effects  $\alpha$  are not observed inference should be obtain from the marginal likelihood:

$$L_M(\theta, x) = \int L(\theta, \alpha, x) d\alpha$$

This integral is difficult to calculate directly, so the Laplace approximation is used. The Laplace approximation is derived by first approximating the joint log likelihood  $\ell(\theta, \alpha, x)$  by a second order Taylor approximation around the optimum  $\hat{\alpha}$  w.r.t.  $\alpha$ . The resulting approximated joint log likelihood can then be integrated by recognizing it as a constant term and a term where the integral is known as the normalizing constant from a multivariate Gaussian. The approximation becomes:

$$\int L(\theta, \alpha, x) d\alpha \approx \sqrt{\frac{(2\pi)^n}{\det(-\ell''_{\alpha\alpha}(\theta, \alpha, x)|_{\alpha=\hat{\alpha}_\theta})}} \exp(\ell(\theta, \hat{\alpha}_\theta, x))$$

---

\*This parametrization is equivalent to the more common  $\phi^{|a-\tilde{a}|}$ , where  $0 < \phi < 1$

Taking the logarithm gives the Laplace approximation of the marginal log likelihood

$$\ell_M(\theta, x) = \ell(\theta, \hat{u}_\theta, x) - \frac{1}{2} \log(\det(-\ell''_{uu}(\theta, u, x)|_{u=\hat{u}_\theta})) + \frac{n}{2} \log(2\pi)$$

Fig. 1. Degree of retrospective pattern in average fishing mortality, recruitment, and spawning stock biomass as measured by Mohn's  $\rho$  for each stock and model type.

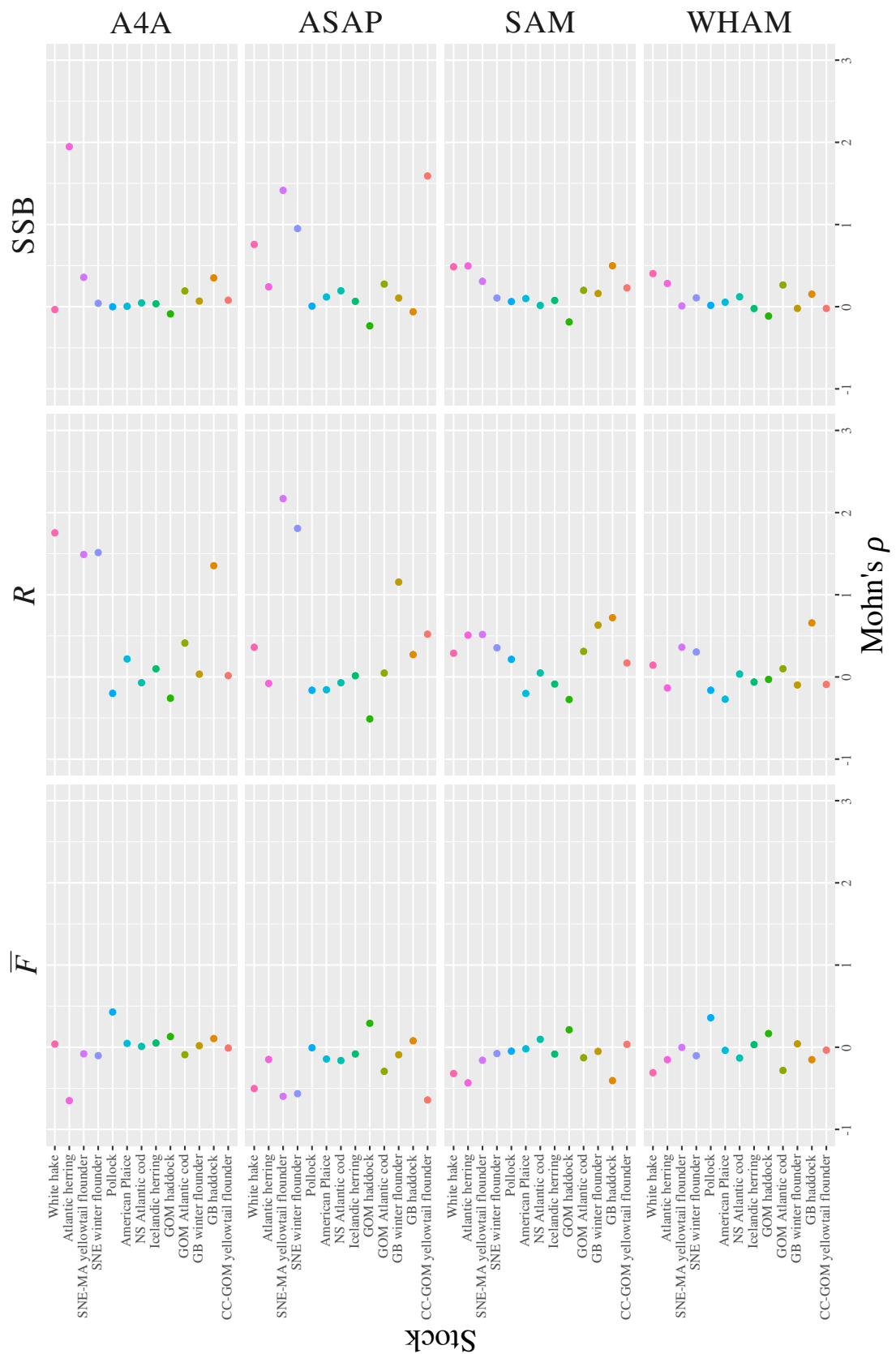


Fig. 2. Relationship between Mohn's  $\rho$  for SSB and average fishing mortality by model.

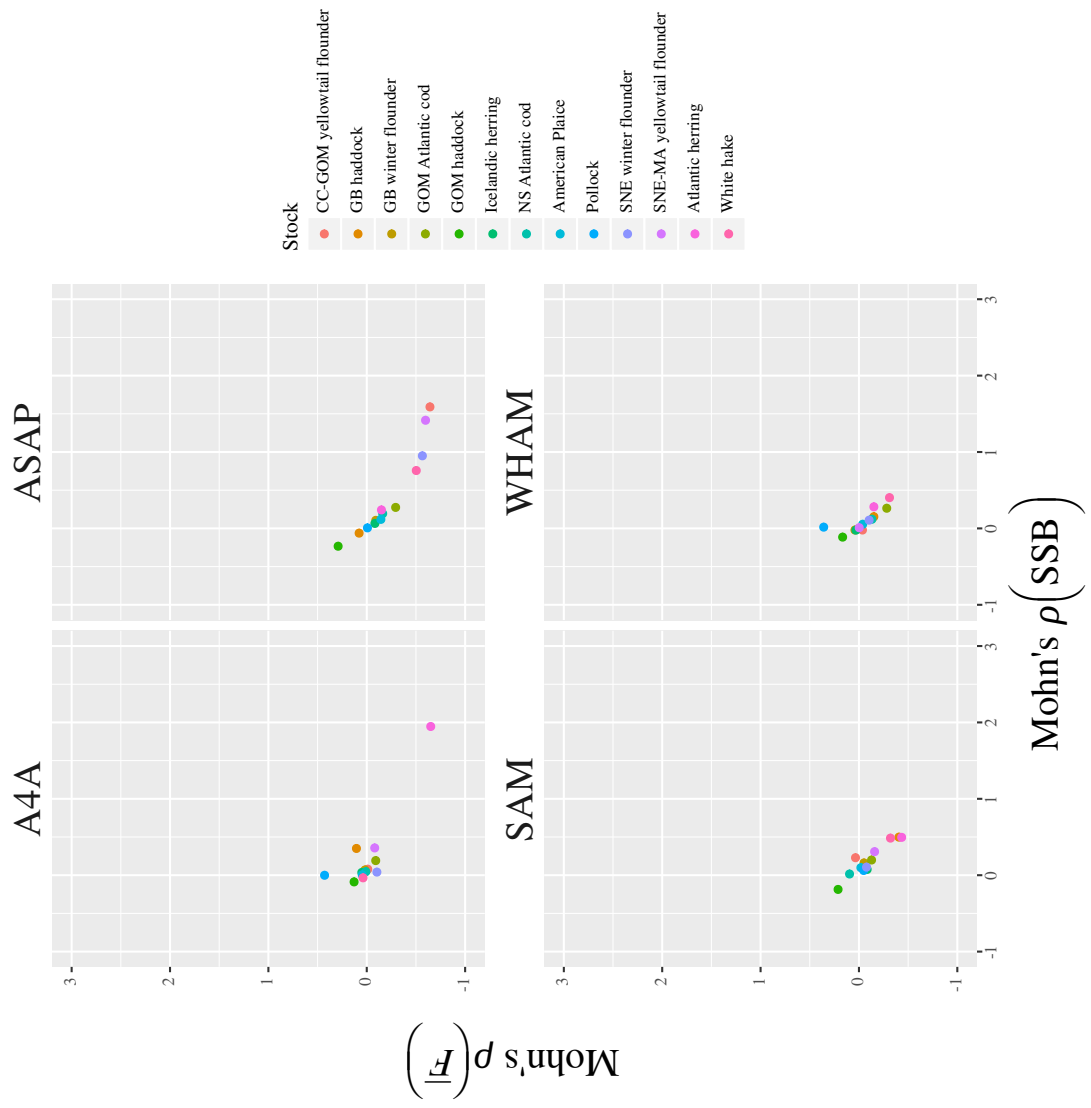


Fig. 3. Degree of retrospective pattern in average fishing mortality, recruitment, and spawning stock biomass as measured by Mohn's  $\rho$  for each stock and WHAM model configuration.

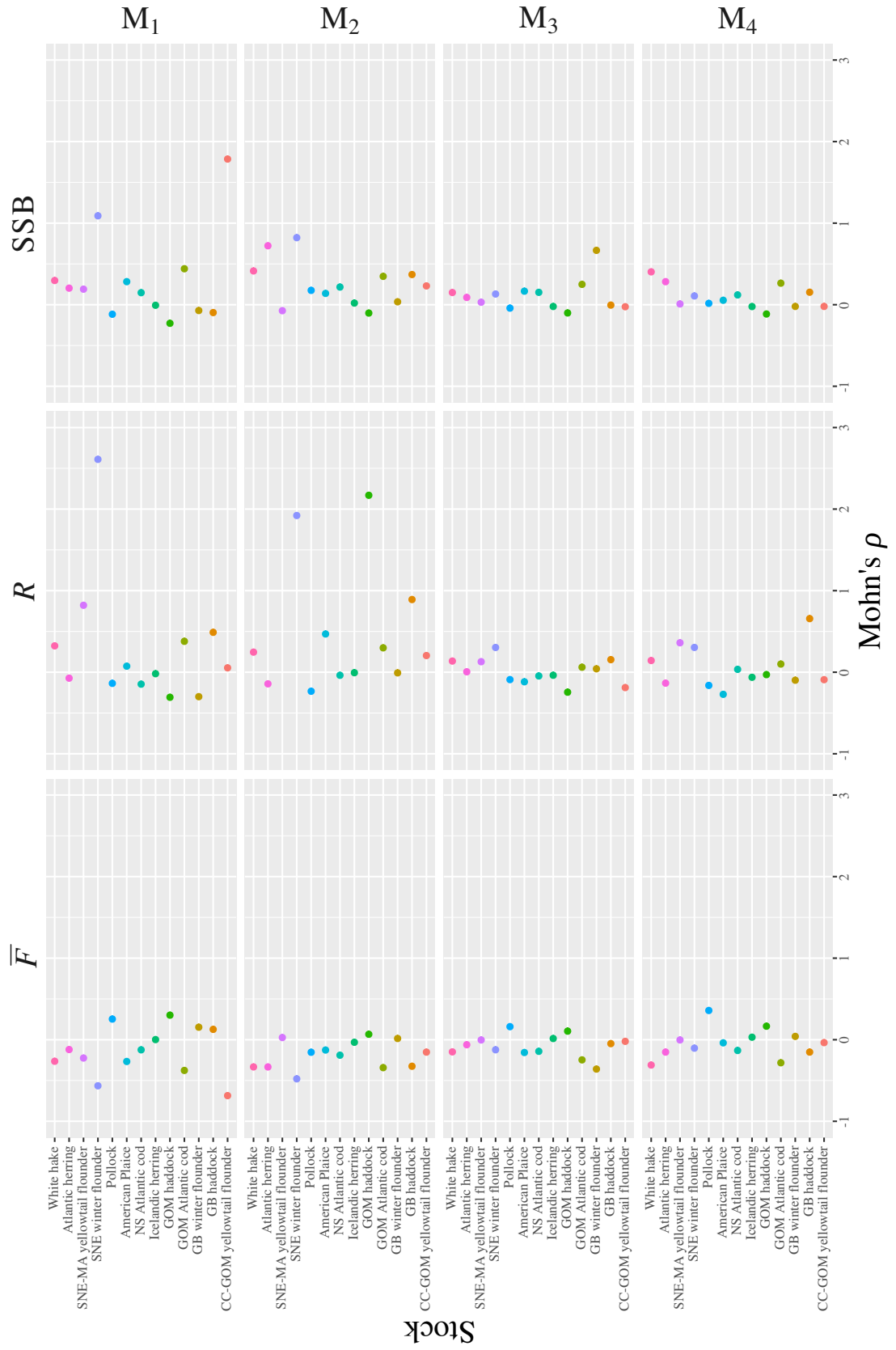


Table 1. Modelling assumptions for the four estimation models.

Assumptions	A4A	ASAP	WHAM	SAM
Fit to total catch				
Fit to catch age composition				
Fit to catch at age matrix				
Fit to total survey index				
Fit to survey index age composition				
Fit to survey index at age matrix				
Recruitment				
Fishing mortality				
Numbers at age				
Comments				

Table 2. Fish stocks considered in this study.

Fish stock	Name code	Reference
Cape Cod-Gulf of Maine yellowtail	CC-GOM yellowtail flounder	
Georges Bank haddock	GB haddock	
Georges Bank winter flounder	GB winter flounder	
Gulf of Maine Atlantic cod	GOM Atlantic cod	
Gulf of Maine haddock	GOM haddock	
Icelandic herring	Icelandic herring	
North Sea Atlantic cod	NS cod	
American plaice	American plaice	
US Pollock	Pollock	
Southern New England-Mid Atlantic winter flounder	SNE winter flounder	
Southern New England-Mid Atlantic yellowtail	SNE-MA yellowtail flounder	
Atlantic herring	Atlantic herring	
White hake	White hake	

Table 3. Mean and Standard Deviation of Mohn's  $\rho$  for SSB,  $\overline{F}$ , and recruitment across all stocks by model type.

Model	Mean	SD
$\overline{F}$		
A4A	-0.01	0.24
ASAP	-0.22	0.28
SAM	-0.11	0.19
WHAM	-0.05	0.18
$R$		
A4A	1.13	2.30
ASAP	0.41	0.81
SAM	0.25	0.31
WHAM	0.06	0.26
SSB		
A4A	0.23	0.53
ASAP	0.42	0.58
SAM	0.20	0.21
WHAM	0.09	0.15