

The Woods Hole Assessment Model (WHAM): Incorporating environmental covariates into a state-space assessment framework

Brian C. Stock¹, Timothy J. Miller¹

¹brian.stock@noaa.gov, timothy.j.miller@noaa.gov, Northeast Fisheries Science Center, National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543, USA

Abstract

WHAM is great.

Keywords

state-space; stock assessment; mixed effects; random effects; time-varying; Template Model Builder (TMB)

1 Introduction

Grab stuff from NRC and Fish/Climate proposals.

1.1 Context: assessments in the U.S. Northeast

- Long history, high F (pre-data)
- Empirical weight-at-age
- Retrospective patterns
- ASAP3/4
- Operational vs. research-track
- The Northeast U.S. Shelf LME is rapidly changing. Top priority is to “continue development of stock assessment models that include environmental terms” (Hare et al., 2016).

1.2 Motivation #1: advantages of state-space stock assessments

- objective estimation of process errors and data weighting, e.g. σ_R , instead of ad-hoc
- inherently predict unobserved states, so predicting missing data/years and into the future is natural
- allow for time/age variation in demographic processes while estimating fewer parameters
- natural framework to include environmental time-series
- lower retros and AIC, larger (more realistic) uncertainty compared to SCAAs. Cite ICES state-space if in review.

(Aeberhard et al., 2018; Miller et al., 2016; Nielsen and Berg, 2014)

1.3 Motivation #2: allow for environmental effects

- Reduced retrospective patterns
- Lower residual variance

(Miller et al., 2016; Miller and Hyun, 2018; O’Leary et al., 2019)

1.4 How is WHAM different from SAM?

Not sure where to put this... may be more natural after introducing equations in Methods, some in Discussion. Definitely will be a question in readers' minds so may be good to introduce early?

Most assessments in the U.S. assume separability in $F_{a,t}$, estimate F_t and Sel_a . WHAM does this. SAM estimates $F_{a,t}$ directly. WHAM and SAM also make different separability assumptions for the catch/index data (aggregate total + age comps vs. $C_{a,t}$ directly). Should be similar (?) but could test.

Goal is to replicate ASAP assessments in the U.S. Northeast. Can easily turn on/off random effects.

Observation model is natural for landings data that are measured as total weight plus age composition sampling. Age composition sampling often done separately with survey data.

Treating F and Sel separately can be useful for projections. Oftentimes we want to specify F in projections to calculate a reference point, as opposed to continuing a F time-series process.

1.5 Bias correction

- Analytical obs error. (Aldrin et al., 2020).
- Analytical process error.
- TMB epsilon. (Thorson, 2019; Thorson and Kristensen, 2016)

Should these all be used?

1.6 Overview

In summary, the NEFSC wants an assessment framework that i) estimates random effects (i.e. a state-space model), ii) includes environmental effects, and iii) is easy to test against status quo SCAA models (ASAP). The objectives of this manuscript are to introduce the WHAM framework and demonstrate unbiasedness in self- and cross-tests.

2 Methods

2.1 Model description

2.1.1 Unobserved states (random effects)

2.1.1.1 Numbers-at-age (survival)

2.1.1.2 Natural mortality (M)

2.1.1.3 Selectivity

2.1.1.4 Environmental covariate(s)

2.1.1.4.1 Time-series model

2.1.1.4.2 Observation model

2.1.1.4.3 Link to population

2.1.2 Data/observation model

2.1.2.1 Catch (agg, age comp)

2.1.2.2 Index (agg, age comp)

2.2 Simulation tests

We used the stocks in Table 1.

We used R (R Core Team, 2020). WHAM is available as an R package (Miller and Stock, 2020). OSA residuals.

3 Results

Sweet figures.

4 Discussion

4.1 Overview

We described WHAM. Sim tests showed no bias in self-tests (when estimation model matched operating model). Some bias in cross-tests.

4.2 Future work

WHAM will be used in upcoming research track assessments. Could transition to operational. Potential to improve several NEFSC assessments.

- 2D AR(1) selectivity. Most assessments in the U.S. assume separability in $F_{a,t}$, i.e. estimate F_t and Sel_a . WHAM does this. SAM estimates $F_{a,t}$ directly. WHAM and SAM make different separability assumptions for the catch/index data as well (aggregate total + age comps vs. $C_{a,t}$ directly). Should be similar (?) but could test.
- How many time/age-varying random effects can be estimated simultaneously? Stock et al. (n.d.) estimated random effect deviations in survival and M , as well as an environmental covariate effect on recruitment.
- Ecov-Recruitment simulation study. How much information does Ecov need to have to be useful?

4.3 Extensions

4.3.1 Multivariate spatiotemporal environmental data

4.3.2 Length/growth estimation

4.3.3 Ecov models

- AR(k)
- splines

- Gaussian process/EDM/Munch/Sugihara

4.4 Conclusion

Development of TMB has facilitated significant advancement in fisheries assessment, allowing us to treat population processes as random effects. A grand challenge in fisheries is to assess and manage stocks in a changing environment. Increasingly have the environmental data. Population time-series are lengthening. WHAM is a step in this direction.

Acknowledgements

This research was performed while BCS held an NRC Research Associateship award at the NEFSC. NOAA Fish & Climate grant number.

¹⁰³ **Supplementary material**

¹⁰⁴ More figures.

References

- Aeberhard, W.H., Mills Flemming, J., Nielsen, A., 2018. Review of State-Space Models for Fisheries Science. *Annu. Rev. Stat. Appl.* 5, 215–235. <https://doi.org/10.1146/annurev-statistics-031017-100427>
- Aldrin, M., Tvete, I., Aanes, S., Subbey, S., 2020. The specification of the data model part in the SAM model matters. *Fisheries Research* 229, 105585. <https://doi.org/10.1016/j.fishres.2020.105585>
- Hare, J.A., Borggaard, D.L., Friedland, K.D., Anderson, J., Burns, P., Chu, K., Clay, P.M., Collins, M.J., Cooper, P., Fratantoni, P.S., Johnson, M.R., Manderson, J.P., Milke, L., Miller, T.J., Orphanides, C.D., Saba, V.S., 2016. Northeast Regional Action Plan - NOAA Fisheries Climate Science Strategy (No. NMFS-NE-239). NOAA Fisheries, Northeast Fisheries Science Center, Woods Hole, MA.
- Miller, T.J., Hare, J.A., Alade, L.A., 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, 1261–1270. <https://doi.org/10.1139/cjfas-2015-0339>
- Miller, T.J., 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, 691–703. <https://doi.org/10.1139/cjfas-2017-0035>
- Miller, T.J., Stock, B.C., 2020. The Woods Hole Assessment Model (WHAM).
- Nielsen, A., Berg, C.W., 2014. Estimation of time-varying selectivity in stock assessments using state-space models. *Fisheries Research* 158, 96–101. <https://doi.org/10.1016/j.fishres.2014.01.014>
- O’Leary, C.A., Miller, T.J., Thorson, J.T., Nye, J.A., 2019. Understanding historical summer flounder (*Paralichthys Dentatus*) abundance patterns through the incorporation of oceanography-dependent vital rates in Bayesian hierarchical models. *Can. J. Fish. Aquat. Sci.* 76, 1275–1294. <https://doi.org/10.1139/cjfas-2018-0092>
- R Core Team, 2020. R: A Language and Environment for Statistical Computing.
- Stock, B.C., Xu, H., Miller, T.J., Thorson, J.T., Nye, J.A., n.d. Implementing a 2-dimensional smoother on either survival or natural mortality improves a state-space assessment model for Southern New England-Mid Atlantic yellowtail flounder.
- Thorson, J.T., 2019. Perspective: Let’s simplify stock assessment by replacing tuning algorithms with statistics. *Fisheries Research* 217, 133–139. <https://doi.org/10.1016/j.fishres.2018.02.005>

134 Thorson, J.T., Kristensen, K., 2016. Implementing a generic method for bias correction in statistical
135 models using random effects, with spatial and population dynamics examples. *Fisheries Research* 175, 66–74.
136 <https://doi.org/10.1016/j.fishres.2015.11.016>

Table 1: Stocks used in simulation tests.

Stock	Modules tested				Model dim		Biol. par.		Stock status	
	NAA	M	Sel	Ecov	# Ages	# Years	M	σ_R	$\frac{B}{B_{40}}$	$\frac{F}{F_{40}}$
SNEMA yellowtail flounder	x	x		x	6	49	0.2-0.4	1.67	0.01	0.44
Butterfish	x	x			5	31	1.3	0.23	2.57	0.03
North Sea cod	x	x			6	54	0.2-1.2	0.87	0.14	2.00
Icelandic herring	x				11	30	0.1	0.55	0.40	1.81
Georges Bank haddock	x		x		9	86	0.2	1.65	5.16	0.12

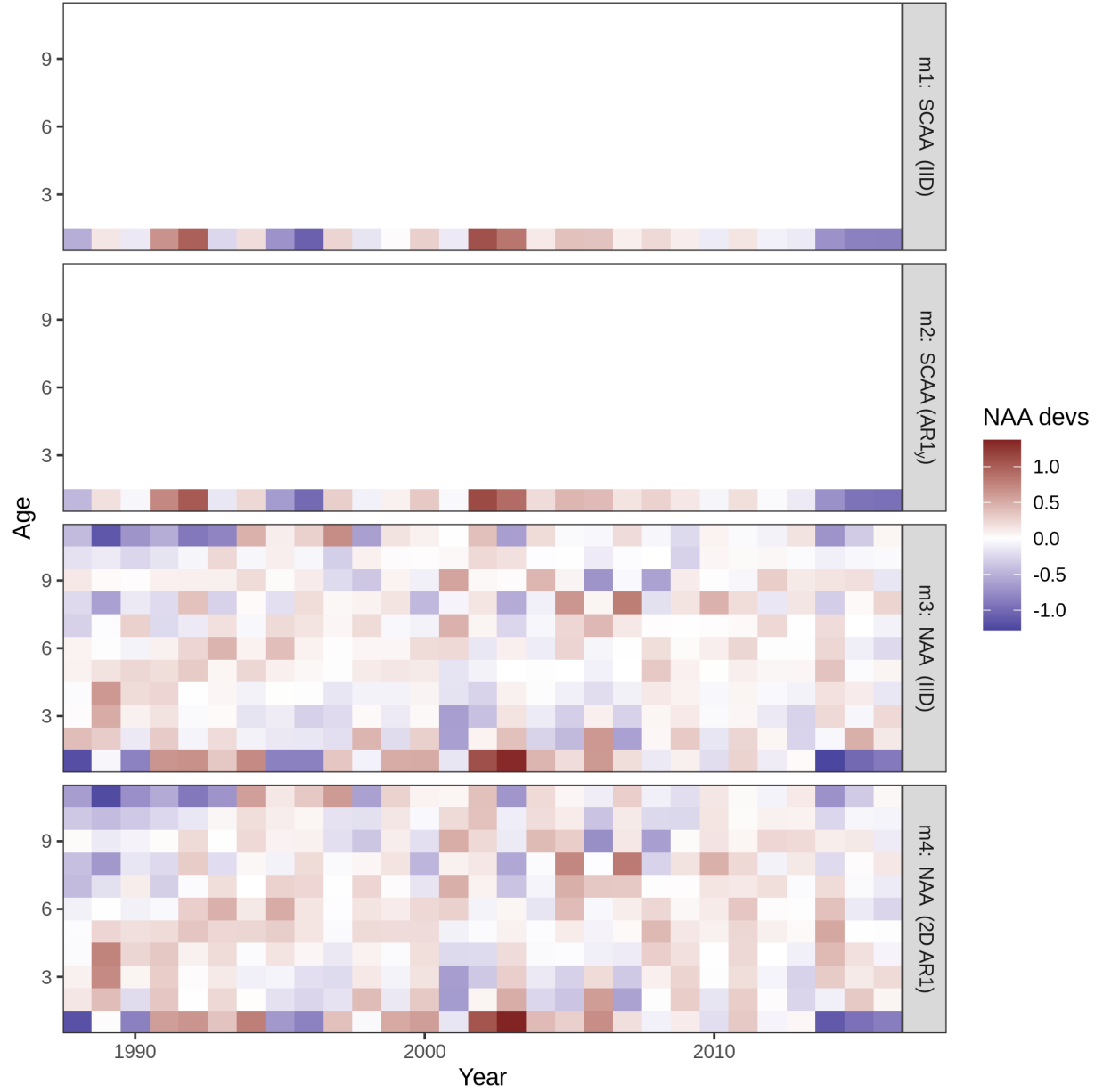


Figure 1: Survival deviations estimated for Icelandic herring using four models of numbers-at-age (NAA) random effects. m1 = only recruitment deviations are random effects (most similar to traditional statistical catch-at-age, SCAA), and deviations are independent and identically distributed (IID). m2 = as m1, but with autocorrelated recruitment deviations (AR1_y). m3 = all NAA deviations are IID random effects. m4 = as m3, but deviations are correlated by age and year (2D AR1).

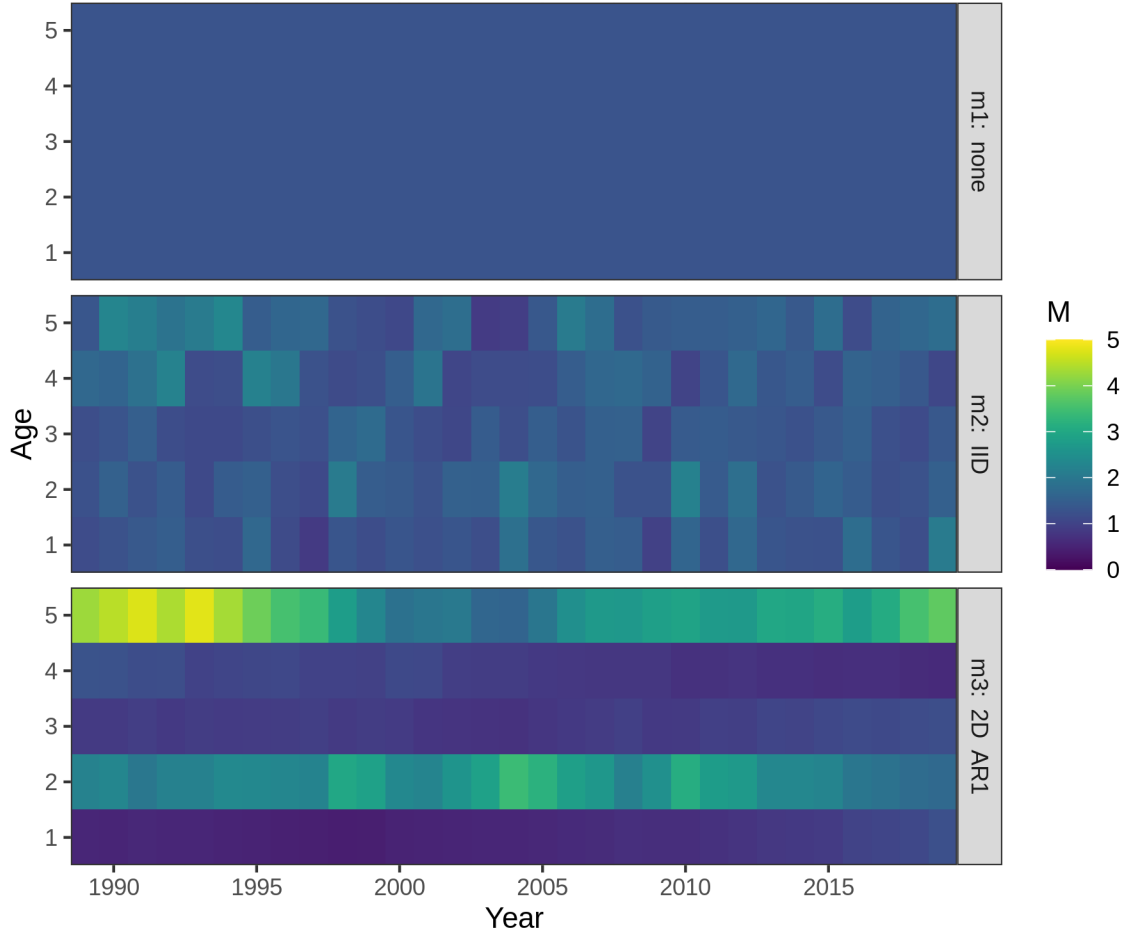


Figure 2: Natural mortality (M) estimated for butterfish using three random effects models. m1 = no random effects on M . m2 = M deviations are independent and identically distributed (IID). m3 = M deviations are correlated by age and year (2D AR1).

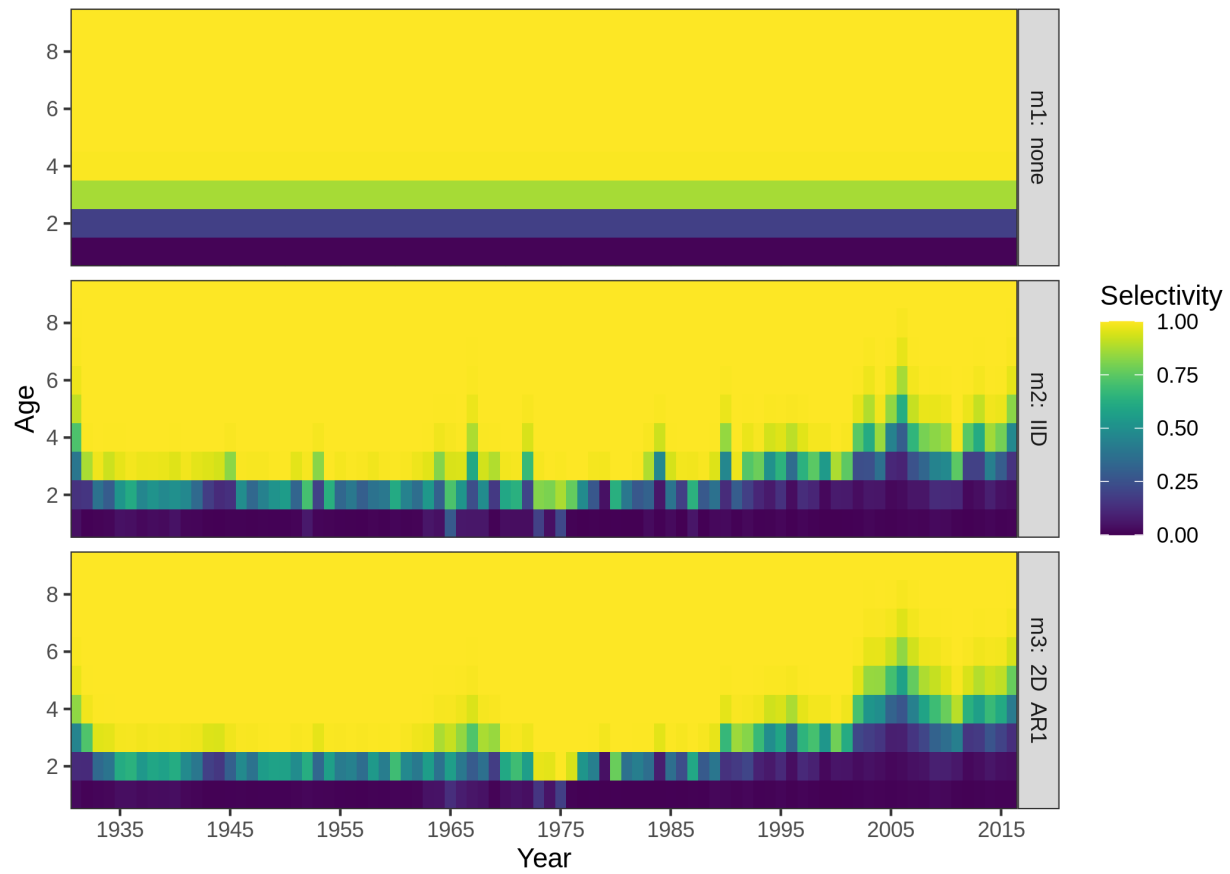


Figure 3: Selectivity estimated for Georges Bank haddock using three random effects models. m1 = no random effects (constant logistic selectivity). m2 = selectivity deviations are independent and identically distributed (IID). m3 = selectivity deviations are correlated by parameter and year (2D AR1).

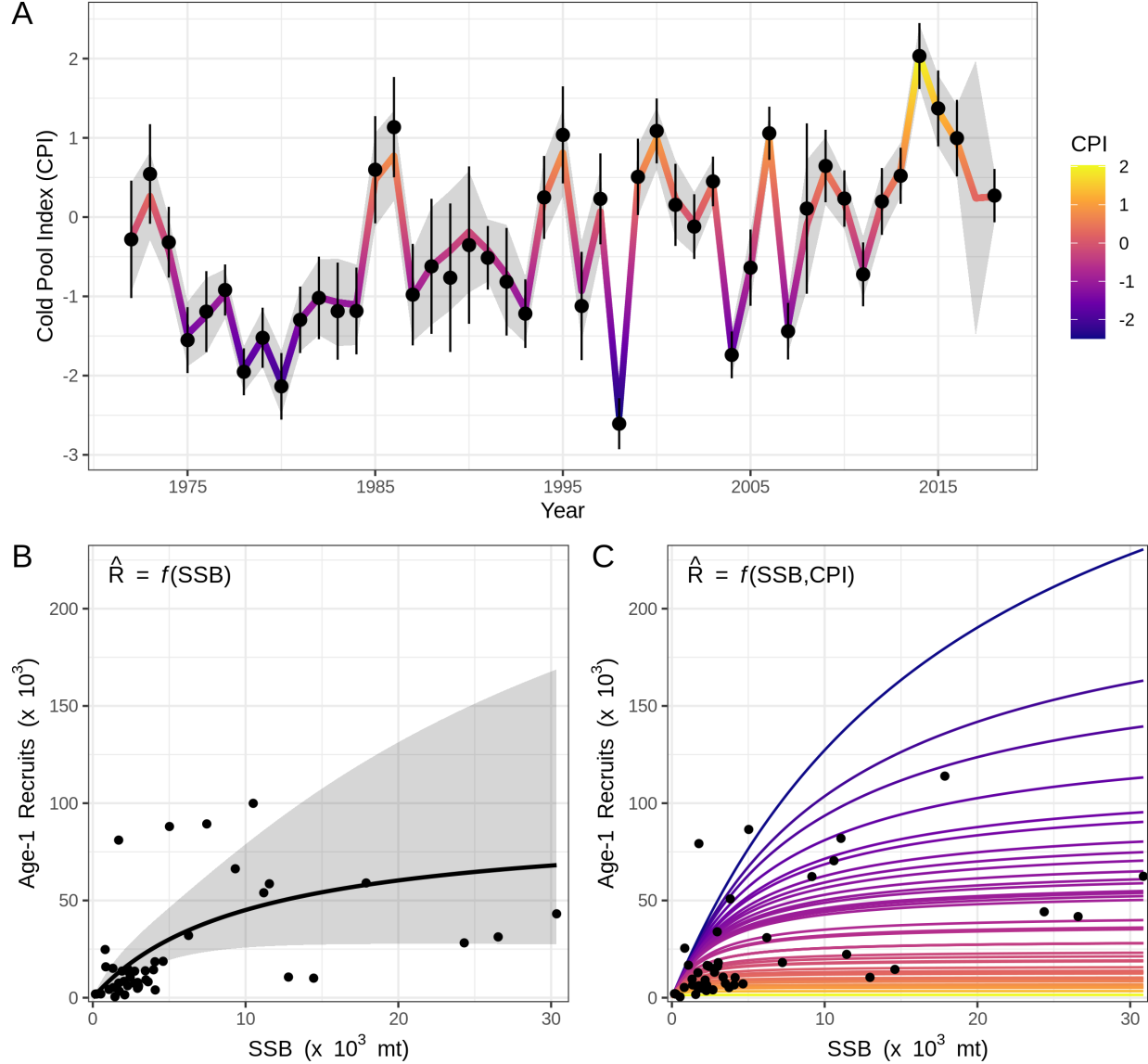


Figure 4: Beverton-Holt stock-recruit relationships fit for Southern New England-Mid Atlantic yellowtail flounder, with and without effects of the Cold Pool Index (CPI). A) CPI estimated from the model with lowest AIC (m4, AR1-linear). Points are observations with 95% CI, and the line with shading is the model-estimated CPI with 95% CI. Note the increased uncertainty surrounding the CPI estimate in 2017 (no observation). B) Estimates of spawning stock biomass (SSB), recruitment, and the stock-recruit function from the model without a CPI effect, m1. C) Estimates of SSB and recruitment from m4, with an effect of the CPI on β . Lines depict the expected stock-recruit relationship in each year t , given the CPI in year $t - 1$ (color).

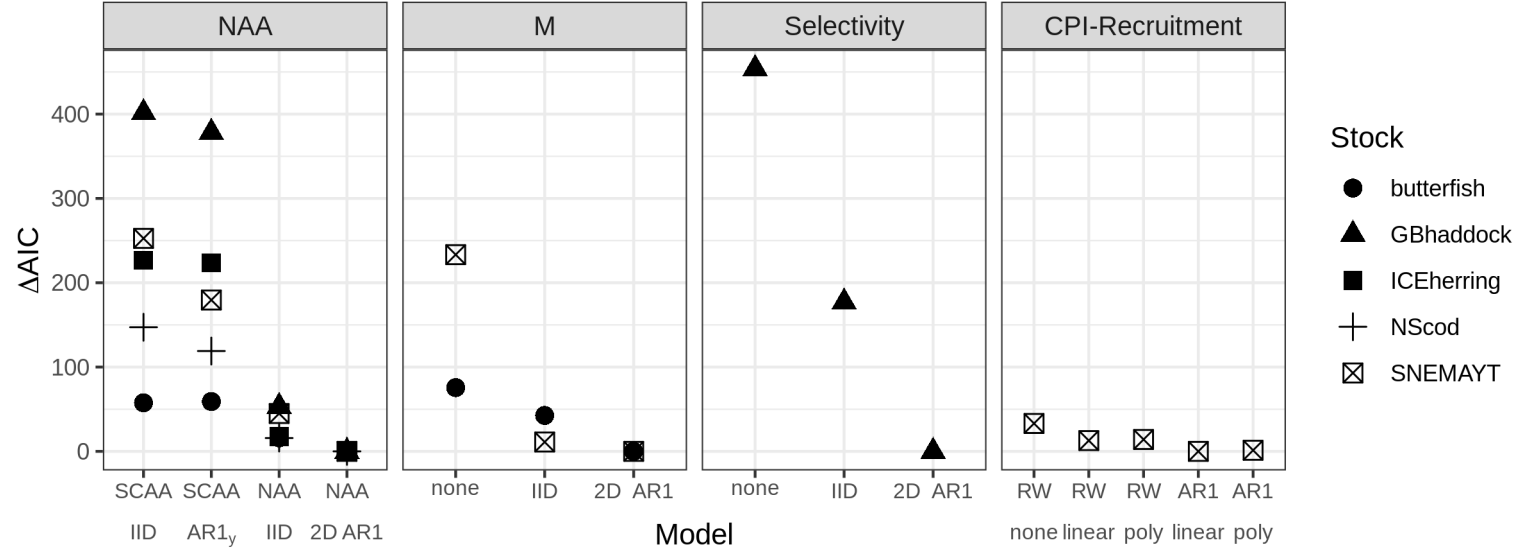


Figure 5: AIC differences by model and stock when fit to full datasets. Stock abbreviations: SNEMA yellowtail flounder (SNEMAYT), North Sea cod (NScod), Icelandic herring (ICEherring), and Georges Bank haddock (GBhaddock).

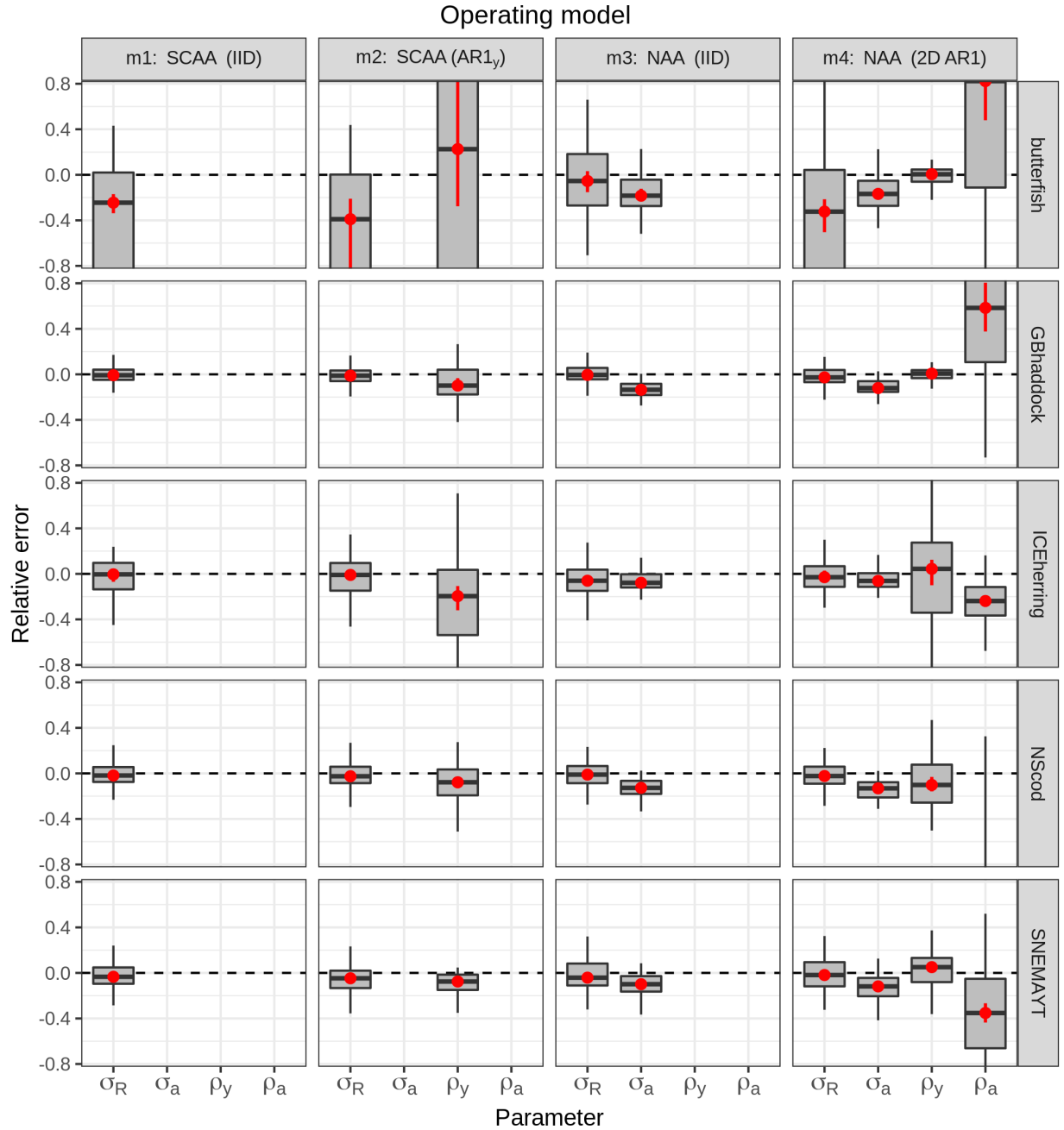


Figure 6: Relative error of parameters constraining numbers-at-age (NAA) random effects. Four models were used to simulate 100 datasets keeping fixed effect parameters constant, and then re-fit to each simulated dataset. m1 = only recruitment deviations are random effects (most similar to traditional statistical catch-at-age, SCAA), and deviations are independent and identically distributed (IID). m2 = as m1, but with autocorrelated recruitment deviations ($AR1_y$). m3 = all NAA deviations are IID random effects. m4 = as m3, but deviations are correlated by age and year (2D $AR1$). Relative error was calculated as $\frac{\hat{\theta}_i}{\theta} - 1$, where $\hat{\theta}_i$ was the estimate in simulation i for parameter θ , and θ was the true value (estimate from original dataset). Red points and lines show median relative error with 95% CI.

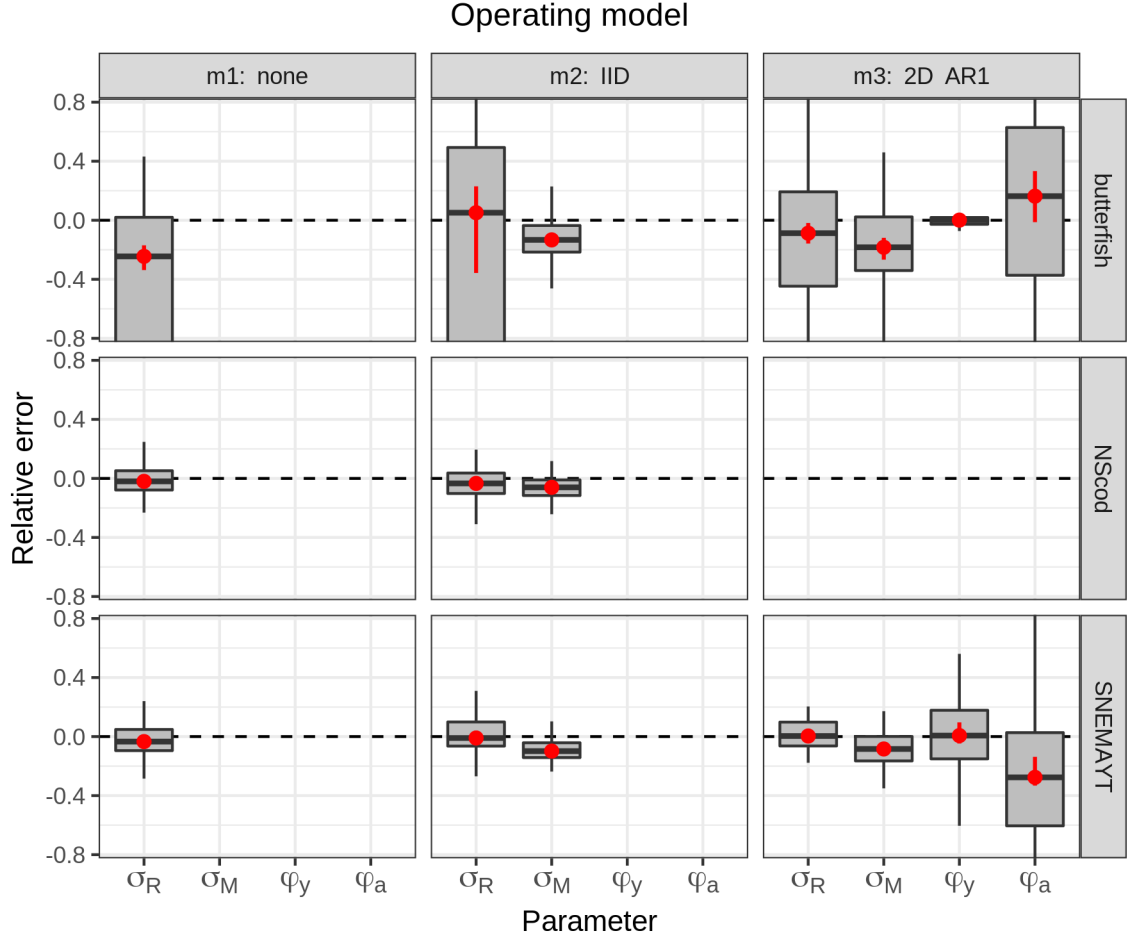


Figure 7: Relative error of parameters constraining natural mortality (M) random effects. Three models were used to simulate 100 datasets keeping fixed effect parameters constant, and then re-fit to each simulated dataset. m1 = no random effects on M . m2 = M deviations were independent and identically distributed (IID). m3 = M deviations were correlated by age and year (2D AR1). Relative error was calculated as $\frac{\hat{\theta}_i}{\theta} - 1$, where $\hat{\theta}_i$ was the estimate in simulation i for parameter θ , and θ was the true value (estimate from original dataset). Red points and lines show median relative error with 95% CI. Stock abbreviations: SNEMAYT yellowtail flounder (SNEMAYT) and North Sea cod (NScod, m3 did not converge).

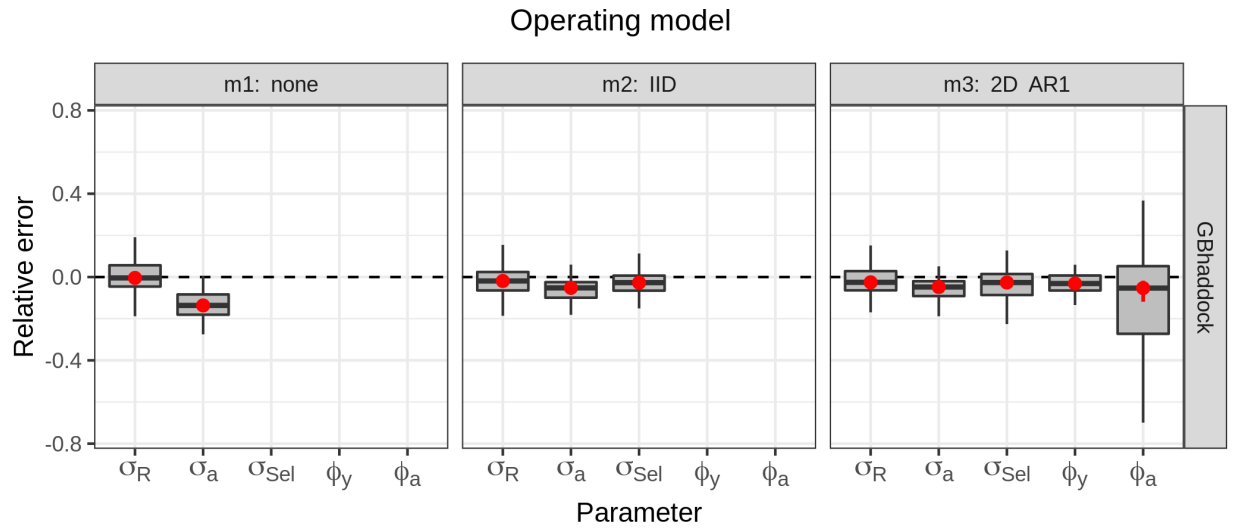


Figure 8: Relative error of parameters constraining selectivity random effects for Georges Bank haddock (GBhaddock). Three models were used to simulate 100 datasets keeping fixed effect parameters constant, and then re-fit to each simulated dataset. m1 = no random effects (constant selectivity). m2 = selectivity deviations were independent and identically distributed (IID). m3 = selectivity deviations were correlated by parameter and year (2D AR1). Relative error was calculated as $\frac{\hat{\theta}_i}{\theta} - 1$, where $\hat{\theta}_i$ was the estimate in simulation i for parameter θ , and θ was the true value (estimate from original dataset). Red points and lines show median relative error with 95% CI.

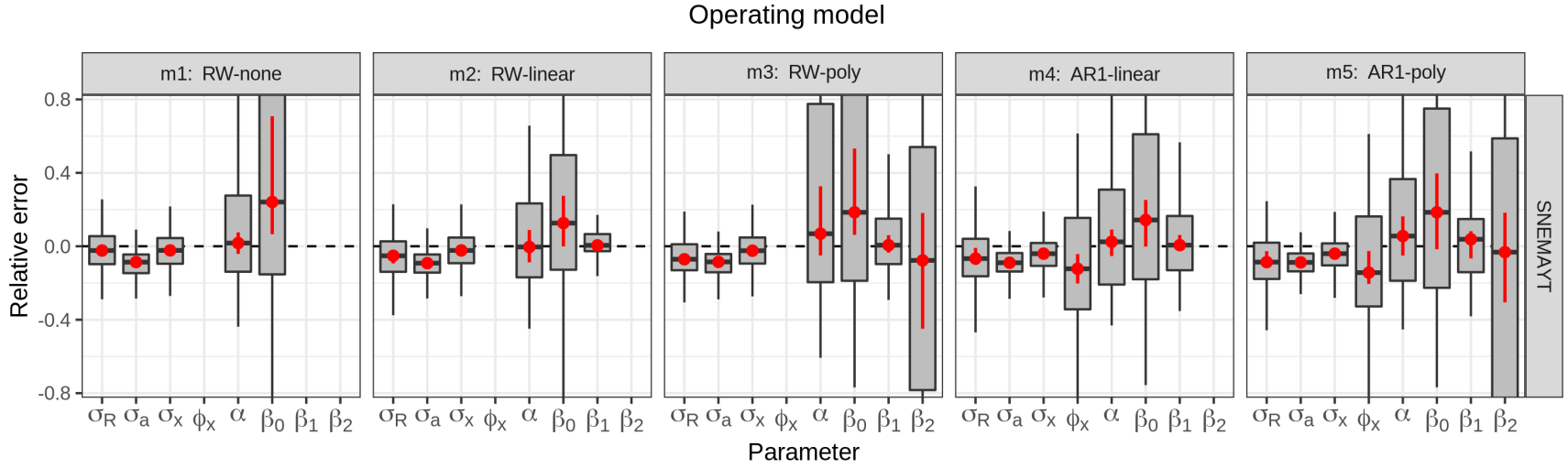


Figure 9: Relative error of parameters constraining variation in recruitment for Southern New England-Mid Atlantic yellowtail flounder (SNEMAYT). Five models were used to simulate 100 datasets keeping fixed effect parameters constant, and then re-fit to each simulated dataset. All models estimated recruitment using the Beverton-Holt function and included CPI effects on β : $\hat{R}_{t+1} = \frac{\alpha S_t}{1 + e^{\beta_0 + \beta_1 x_t + \beta_2 x_t^2 S_t}}$. m1 = Cold Pool Index (CPI) modeled as a random walk (RW) with no effect on recruitment ($\beta_1 = \beta_2 = 0$). m2 = CPI as RW, linear effect on β . m3 = CPI as RW, 2nd order polynomial effect on β . m4 = CPI as AR1, linear effect. m5 = CPI as AR1, polynomial effect. Relative error was calculated as $\frac{\hat{\theta}_i}{\theta} - 1$, where $\hat{\theta}_i$ was the estimate in simulation i for parameter θ , and θ was the true value (estimate from original dataset). Red points and lines show median relative error with 95% CI.

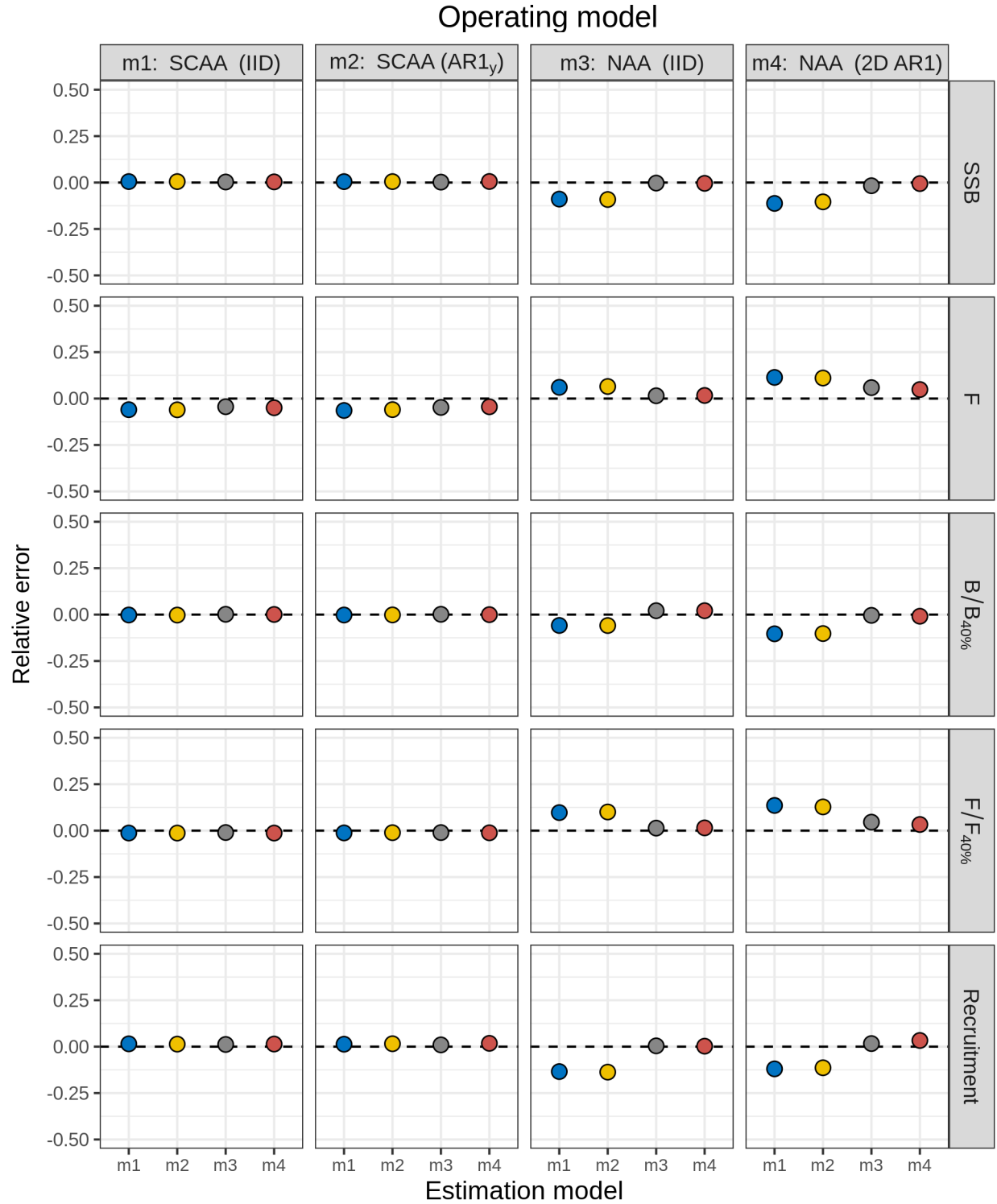


Figure 10: Relative error of key quantities estimated for Icelandic herring using four models of numbers-at-age (NAA) random effects. m1 = only recruitment deviations are random effects (most similar to traditional statistical catch-at-age, SCAA), and deviations are independent and identically distributed (IID). m2 = as m1, but with autocorrelated recruitment deviations (AR1_y). m3 = all NAA deviations are IID random effects. m4 = as m3, but deviations are correlated by age and year (2D AR1).

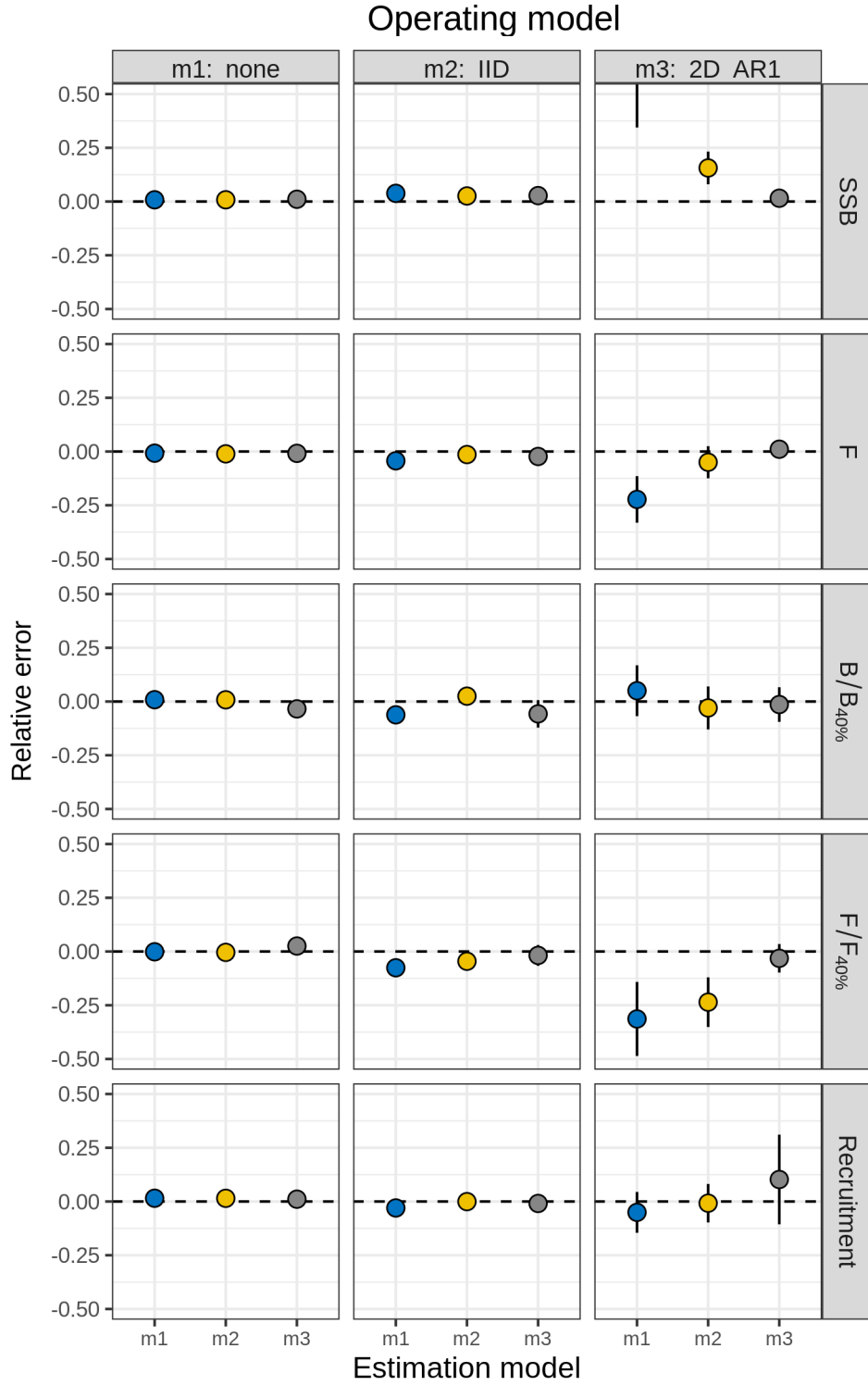


Figure 11: Relative error of key quantities estimated for butterfish using three models of natural mortality (M) random effects. m1 = no random effects on M . m2 = M deviations are independent and identically distributed (IID). m3 = M deviations are correlated by age and year (2D AR1).

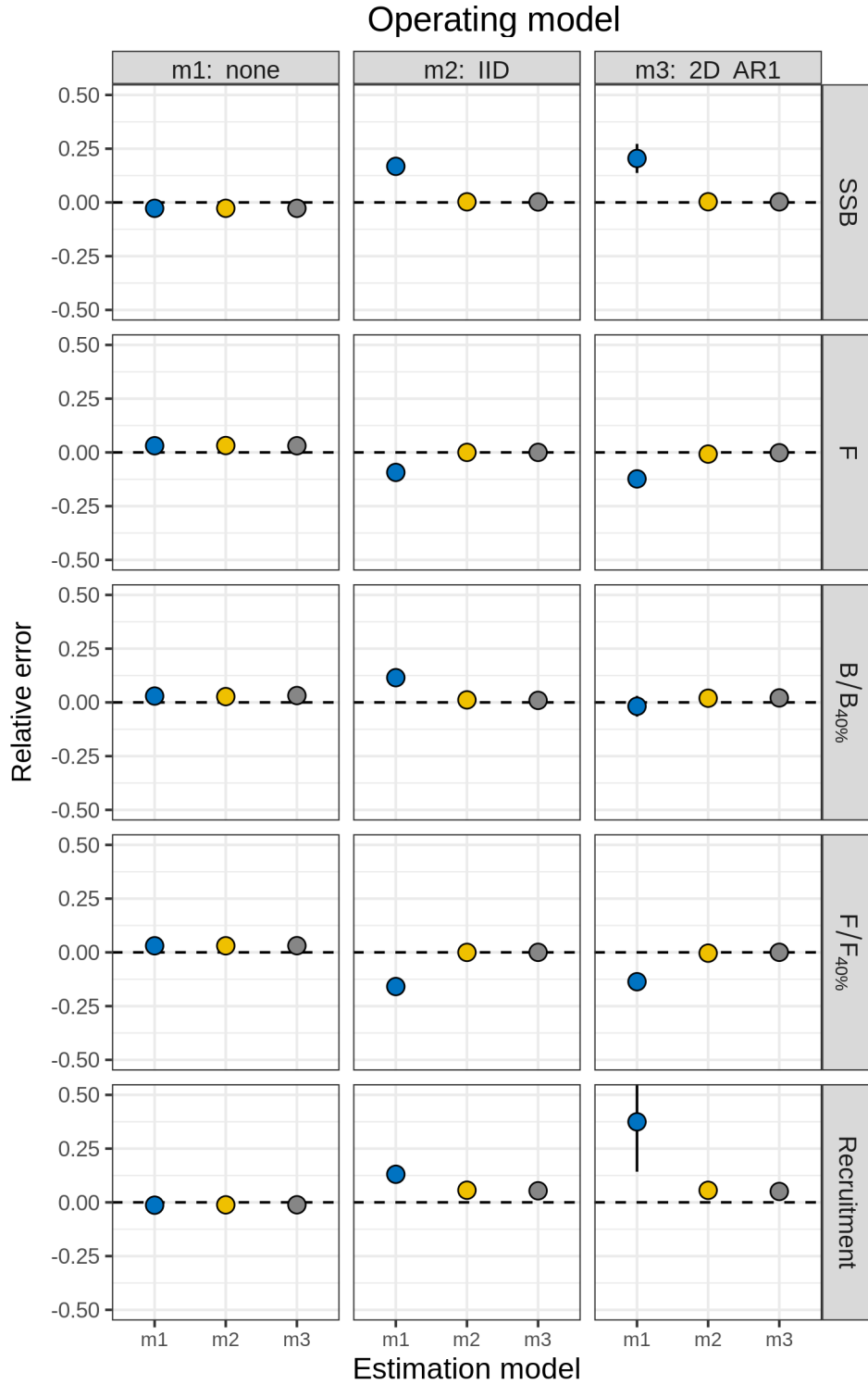


Figure 12: Relative error of key quantities estimated for Georges Bank haddock using three models of selectivity random effects. m1 = no random effects (constant logistic selectivity). m2 = selectivity deviations are independent and identically distributed (IID). m3 = selectivity deviations are correlated by parameter and year (2D AR1).

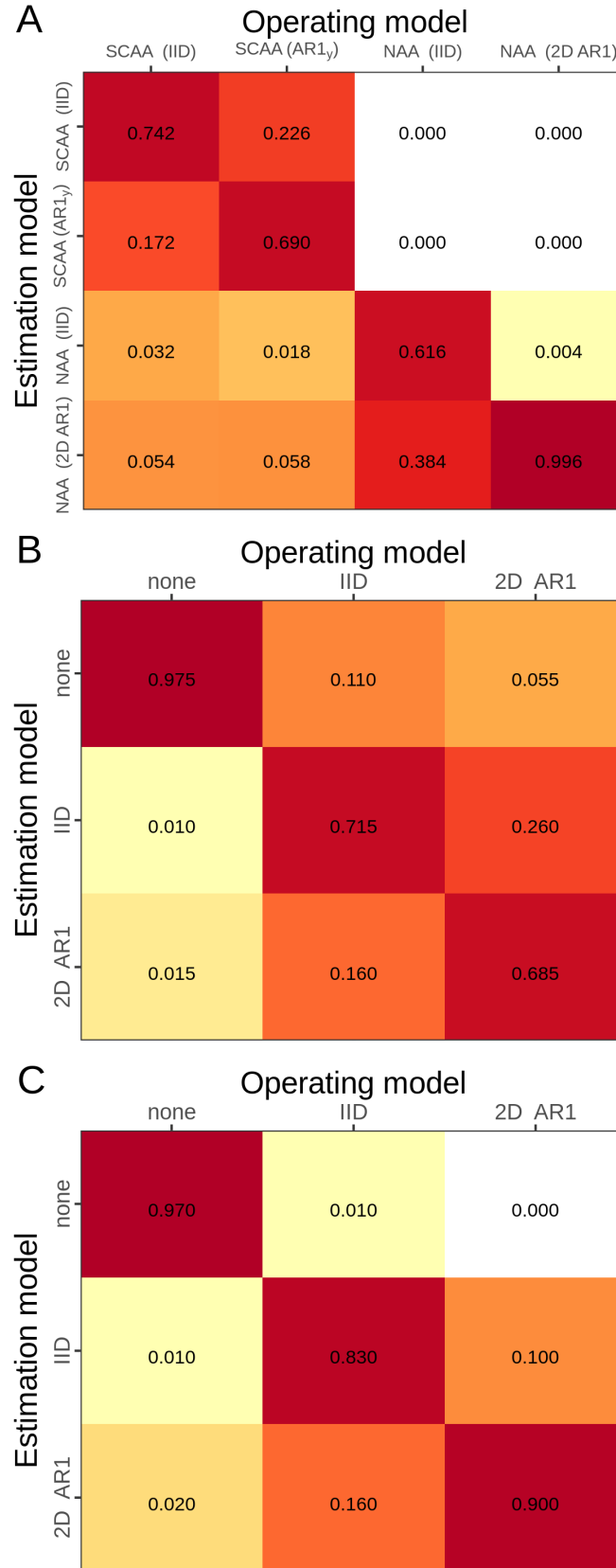


Figure 13: Proportion of simulations in which each model had the lowest AIC. A) Numbers-at-age (NAA), aggregated across all five stocks. B) Natural mortality (M), aggregated over two stocks (SNEMAYT and NScod). C) Selectivity (GBhaddock). Not all estimation models converged for each simulation, even when the operating model matched.