

## **Supplemental Information**

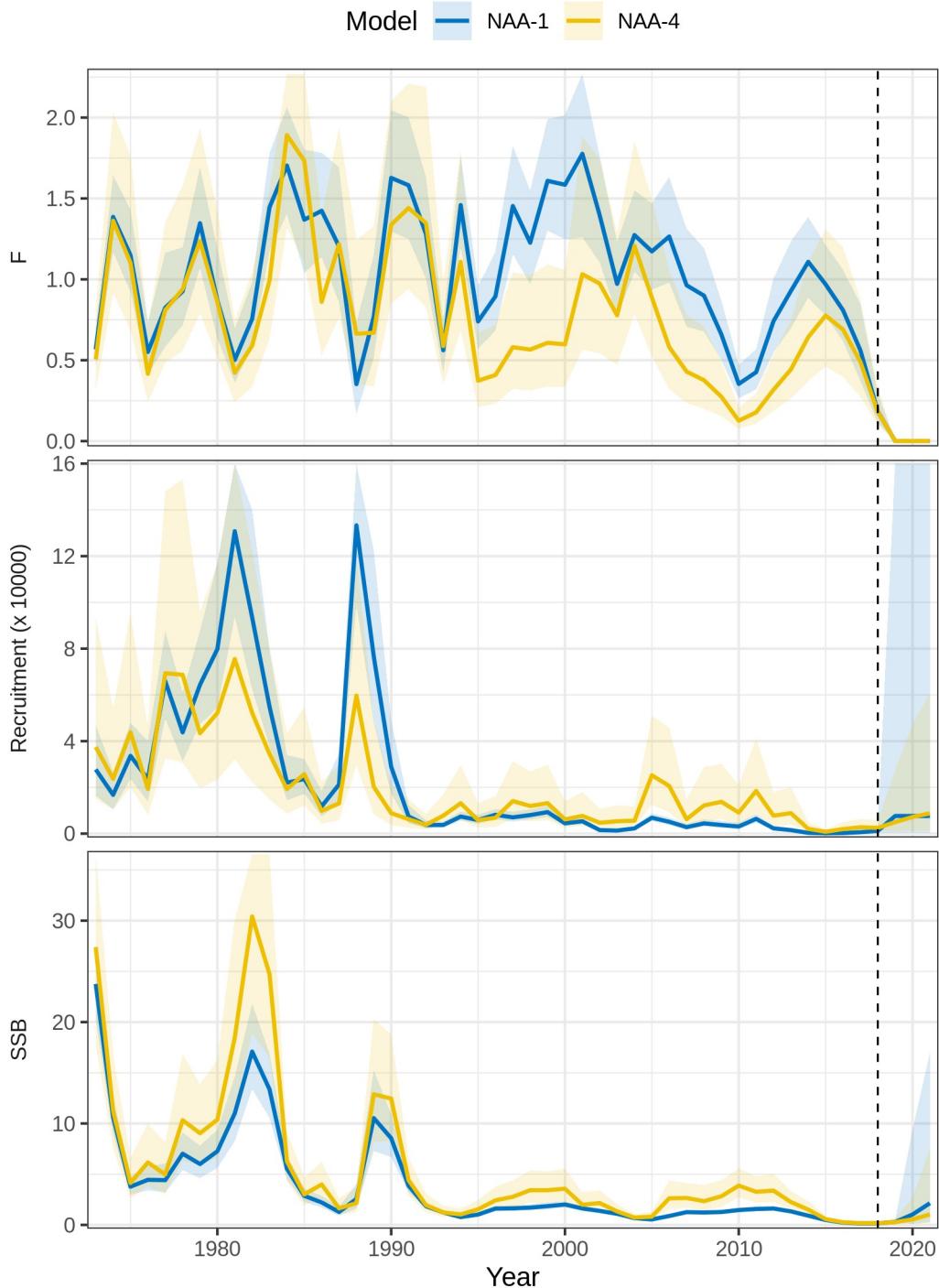
The Woods Hole Assessment Model (WHAM): a general state-space assessment framework that incorporates time- and age-varying processes via random effects and links to environmental covariates

Brian C. Stock\* and Timothy J. Miller

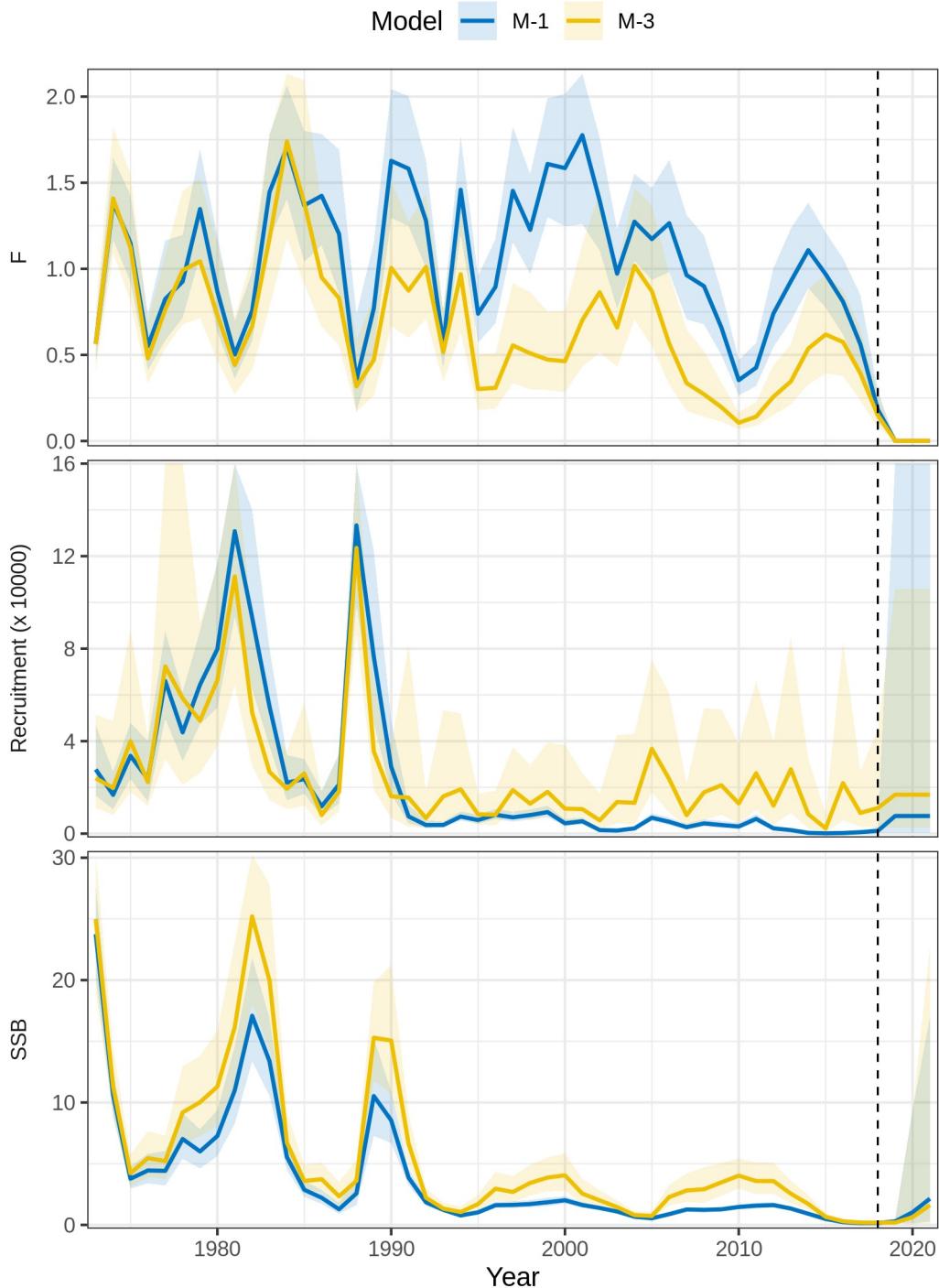
\*[brian.stock@noaa.gov](mailto:brian.stock@noaa.gov)

Model	Mohn's $\rho_R$	Mohn's $\rho_{SSB}$	Mohn's $\rho_F$
<b>SNEMA yellowtail flounder</b>			
Base	5.57	0.85	-0.32
NAA-1	6.42	1.02	-0.43
NAA-2	4.56	0.90	-0.38
NAA-3	0.81	0.17	-0.10
NAA-4	0.51	0.06	0.00
<b>Butterfish</b>			
Base	0.18	0.25	-0.16
NAA-1	0.05	0.10	-0.07
NAA-2	0.07	0.11	-0.08
NAA-3	0.05	0.05	0.02
NAA-4	0.11	0.16	-0.12
<b>North Sea cod</b>			
Base	-0.09	0.17	-0.15
NAA-1	-0.04	0.22	-0.19
NAA-2	-0.05	0.19	-0.16
NAA-3	0.02	0.10	-0.11
NAA-4	0.00	0.05	-0.05
<b>Georges Bank haddock</b>			
Base	0.94	0.36	-0.31
NAA-1	0.89	0.38	-0.33
NAA-2	0.88	0.36	-0.31
NAA-3	0.61	0.11	-0.11
NAA-4	0.31	0.06	-0.07
<b>Icelandic herring</b>			
Base	0.51	0.63	-0.36
NAA-1	-0.01	0.02	-0.03
NAA-2	0.01	0.03	-0.04
NAA-3	-0.07	-0.05	0.06
NAA-4	-0.05	-0.03	0.03

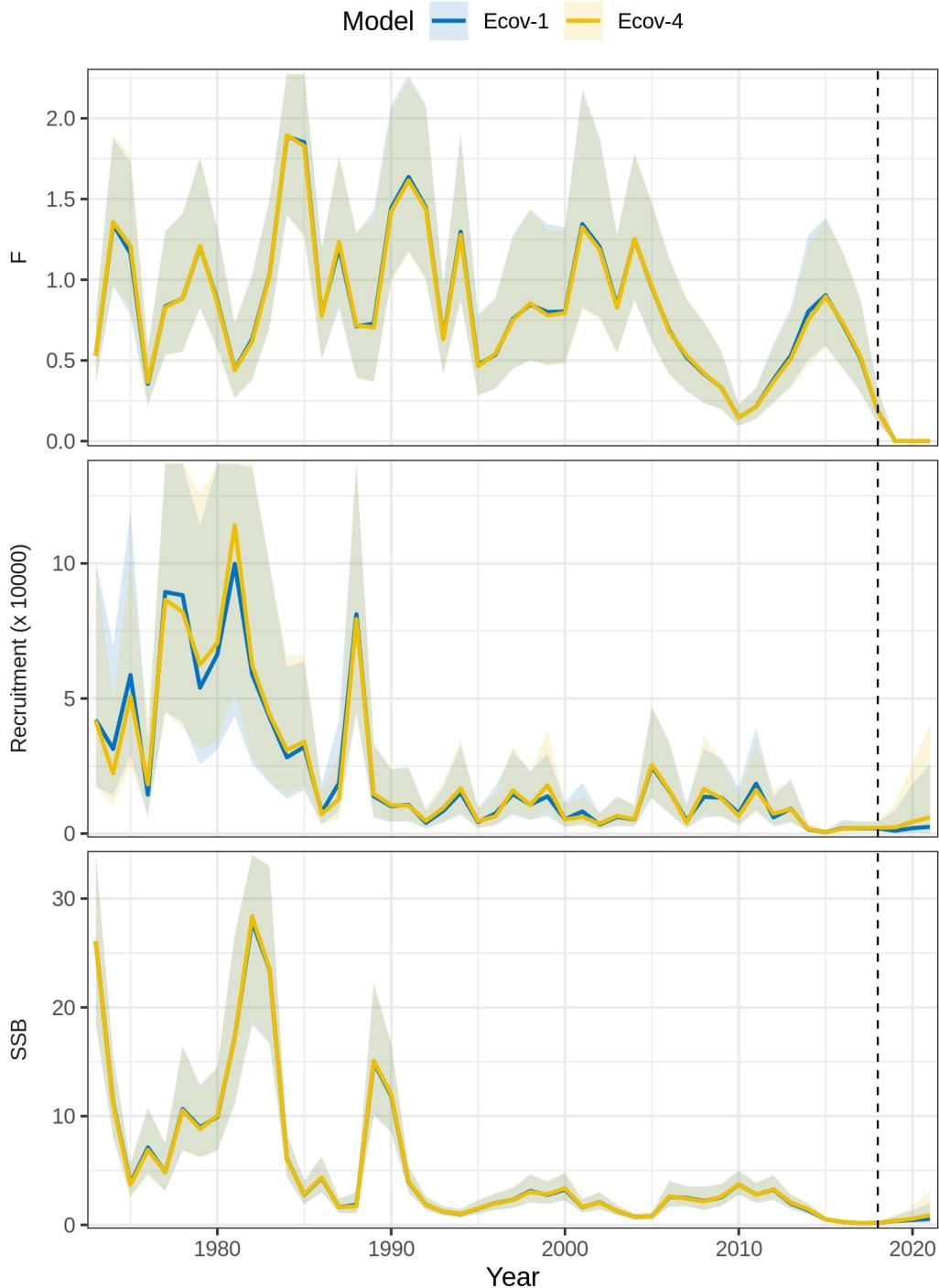
**Table S1.** Mohn's  $\rho$  values for recruitment ( $R$ ), spawning stock biomass (SSB), and fishing mortality ( $F$ ), estimated for all five stocks by the Base (recruitment deviations are fixed effects) and numbers-at-age (NAA) random effects models.



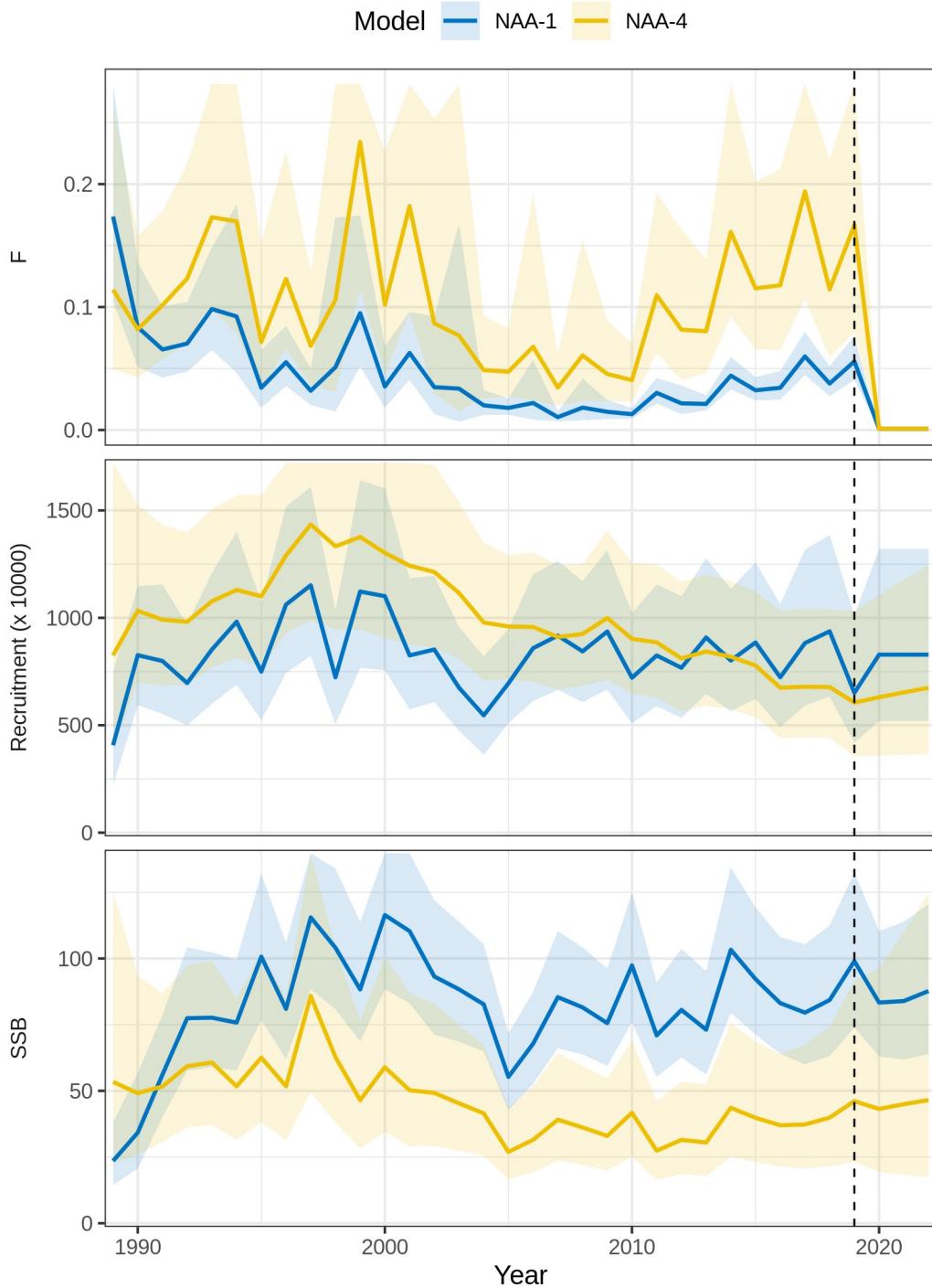
**Figure S1.** Trends in  $F$ , recruitment, and SSB estimated for SNEMA yellowtail flounder using two models of numbers-at-age (NAA) random effects. NAA-1 = only recruitment deviations are independent random effects (most similar to traditional statistical catch-at-age). NAA-4 = all NAA deviations are random effects correlated by age and year (2D AR1). NAA-4 had the lowest AIC (Fig. 1 in main text). The vertical dashed line indicates the terminal year in the assessment.



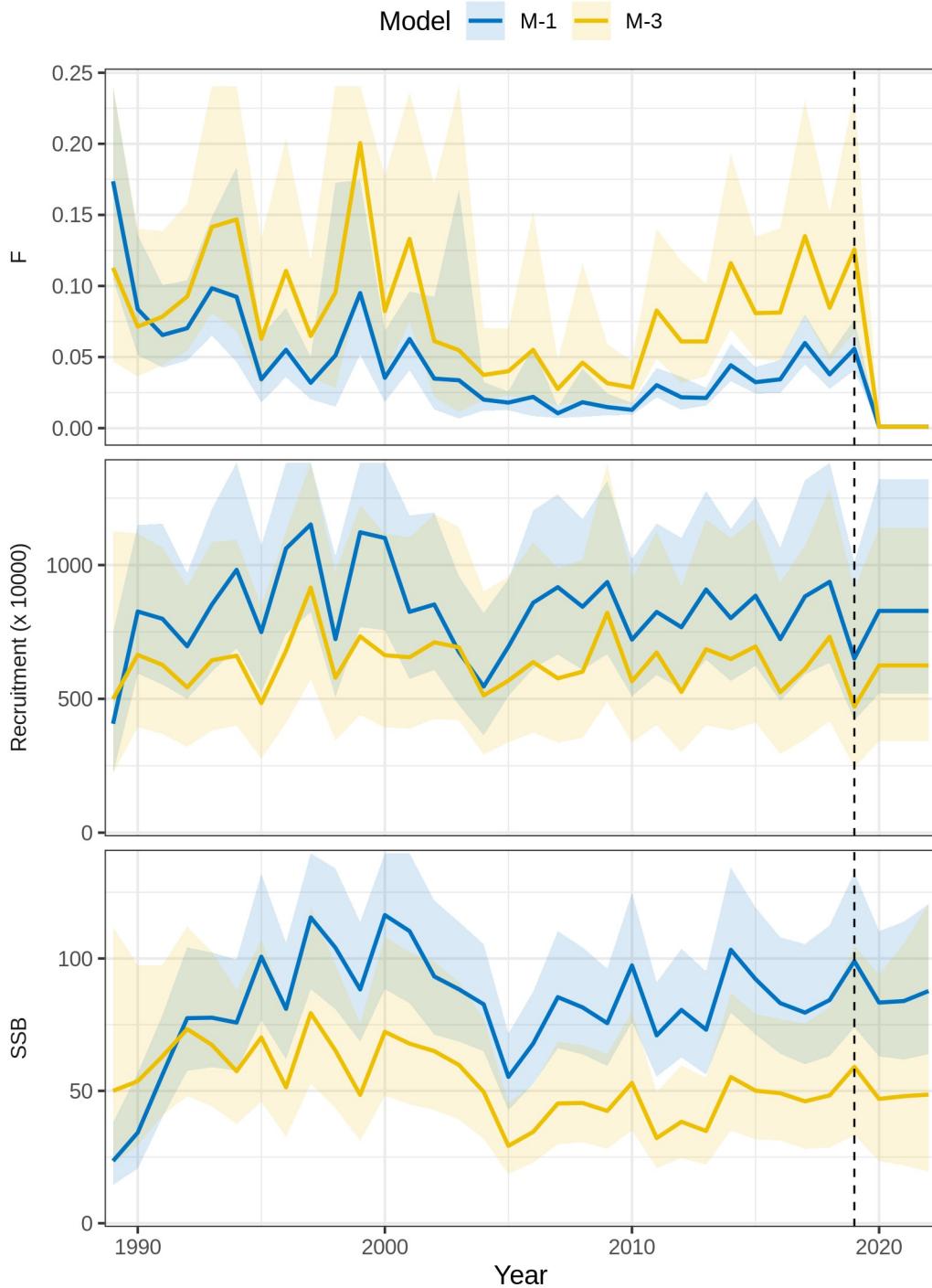
**Figure S2.** Trends in  $F$ , recruitment, and SSB estimated for SNEMA yellowtail flounder using two models of natural mortality ( $M$ ) random effects. M-1 = no random effects on  $M$ . M-3 =  $M$  deviations are random effects correlated by age and year (2D AR1). M-3 had the lowest AIC (Fig. 1 in main text). The vertical dashed line indicates the terminal year in the assessment.



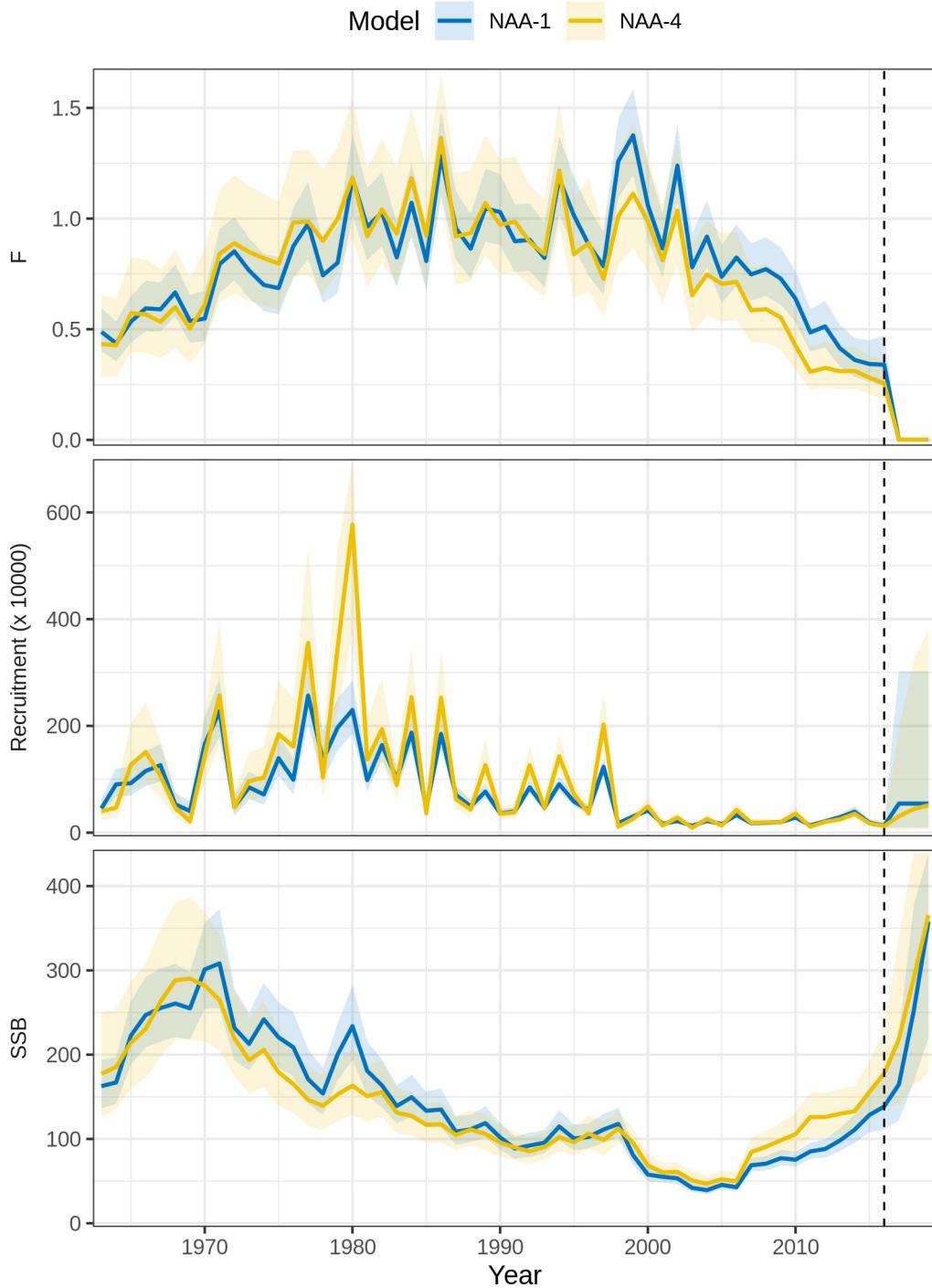
**Figure S3.** Trends in  $F$ , recruitment, and SSB estimated for SNEMA yellowtail flounder using models with and without effects of the Cold Pool Index (CPI) on recruitment. Ecov-1 = no CPI effect. Ecov-4 = CPI modeled as AR1 process with linear effect on the Beverton-Holt  $\beta$  parameter. Ecov-4 had the lowest AIC (Fig. 1 in main text). The vertical dashed line indicates the terminal year in the assessment.



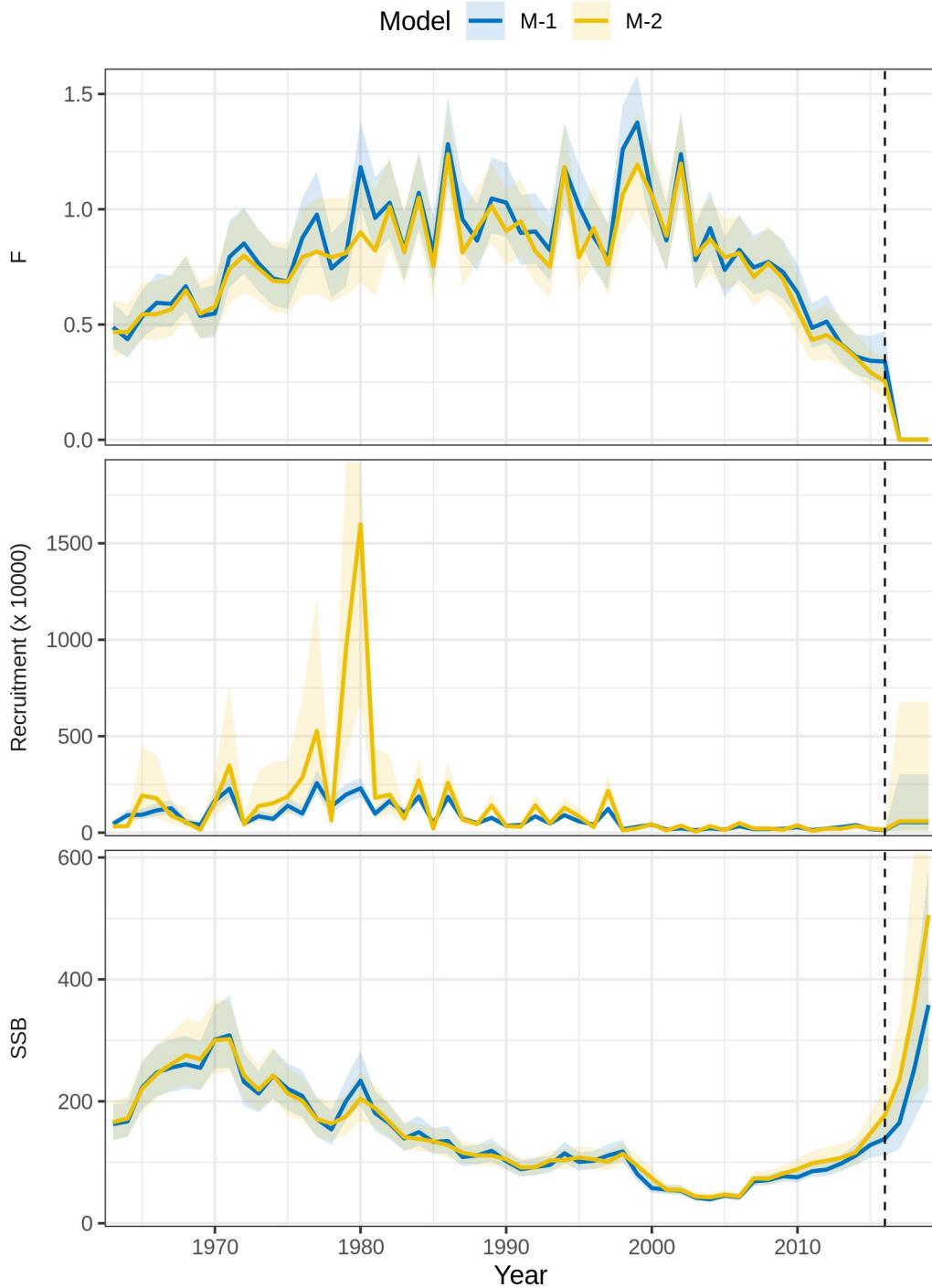
**Figure S4.** Trends in  $F$ , recruitment, and SSB estimated for butterfish using two models of numbers-at-age (NAA) random effects. NAA-1 = only recruitment deviations are independent random effects (most similar to traditional statistical catch-at-age). NAA-4 = all NAA deviations are random effects correlated by age and year (2D AR1). NAA-4 had the lowest AIC (Fig. 1 in main text). The vertical dashed line indicates the terminal year in the assessment.



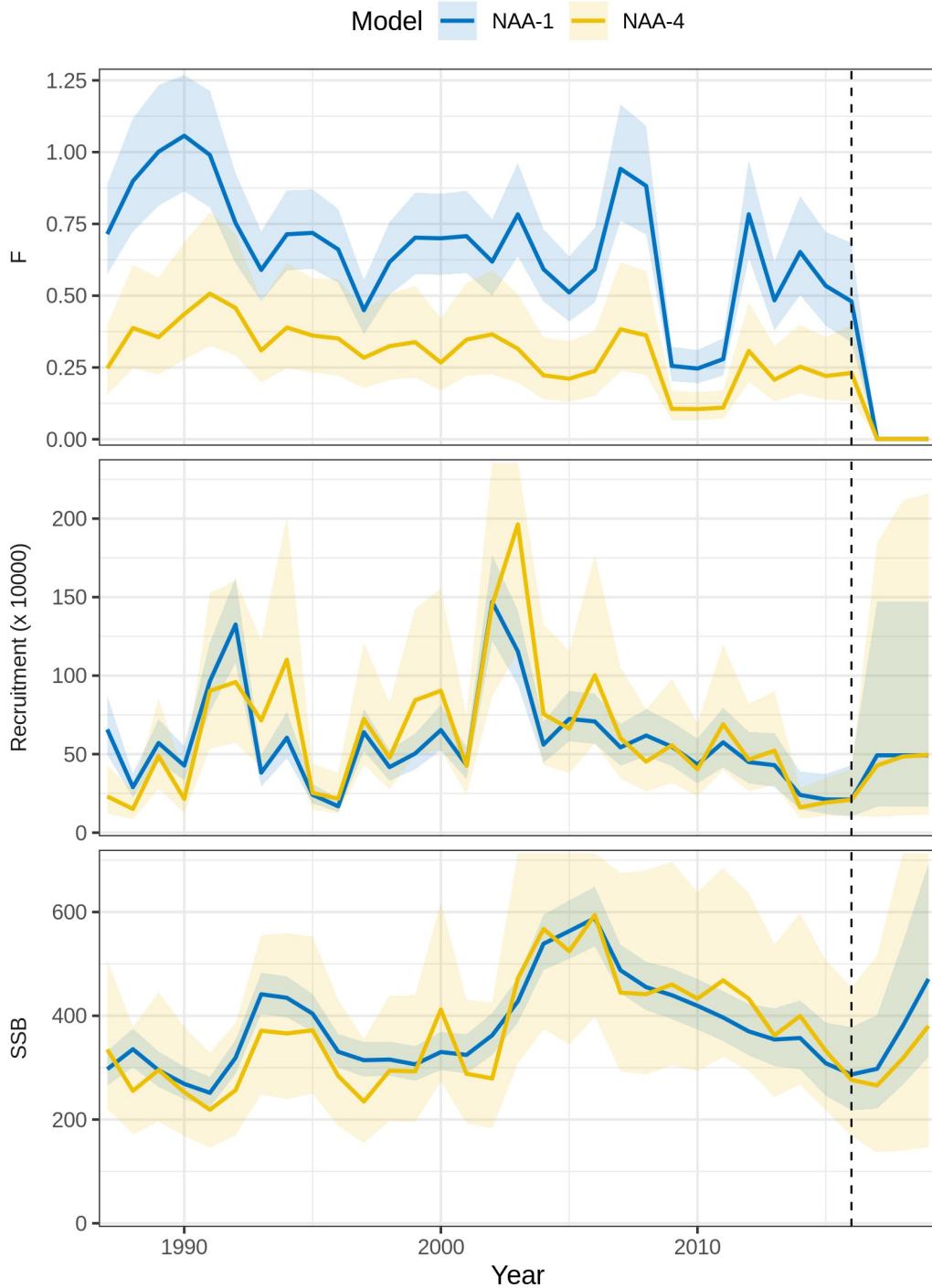
**Figure S5.** Trends in  $F$ , recruitment, and SSB estimated for butterfish using two models of natural mortality ( $M$ ) random effects. M-1 = no random effects on  $M$ . M-3 =  $M$  deviations are random effects correlated by age and year (2D AR1). M-3 had the lowest AIC (Fig. 1 in main text). The vertical dashed line indicates the terminal year in the assessment.



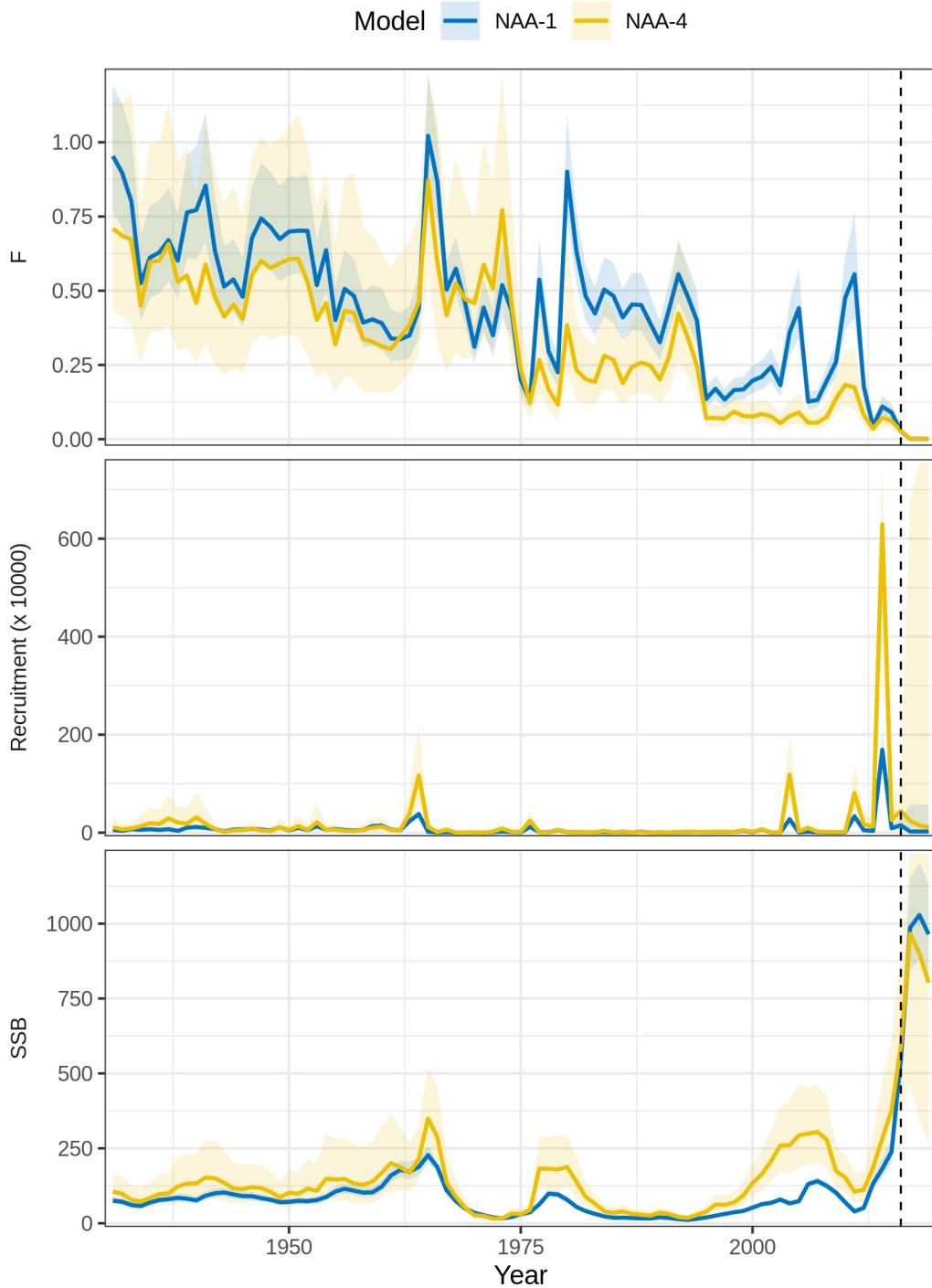
**Figure S6.** Trends in  $F$ , recruitment, and SSB estimated for North Sea cod using two models of numbers-at-age (NAA) random effects. NAA-1 = only recruitment deviations are independent random effects (most similar to traditional statistical catch-at-age). NAA-4 = all NAA deviations are random effects correlated by age and year (2D AR1). NAA-4 had the lowest AIC (Fig. 1 in main text). The vertical dashed line indicates the terminal year in the assessment.



**Figure S7.** Trends in  $F$ , recruitment, and SSB estimated for North Sea cod using two models of natural mortality ( $M$ ) random effects. M-1 = no random effects on  $M$ . M-2 =  $M$  deviations are independent random effects. M-2 had the lowest AIC (Fig. 1 in main text) and M-3 did not converge. The vertical dashed line indicates the terminal year in the assessment.



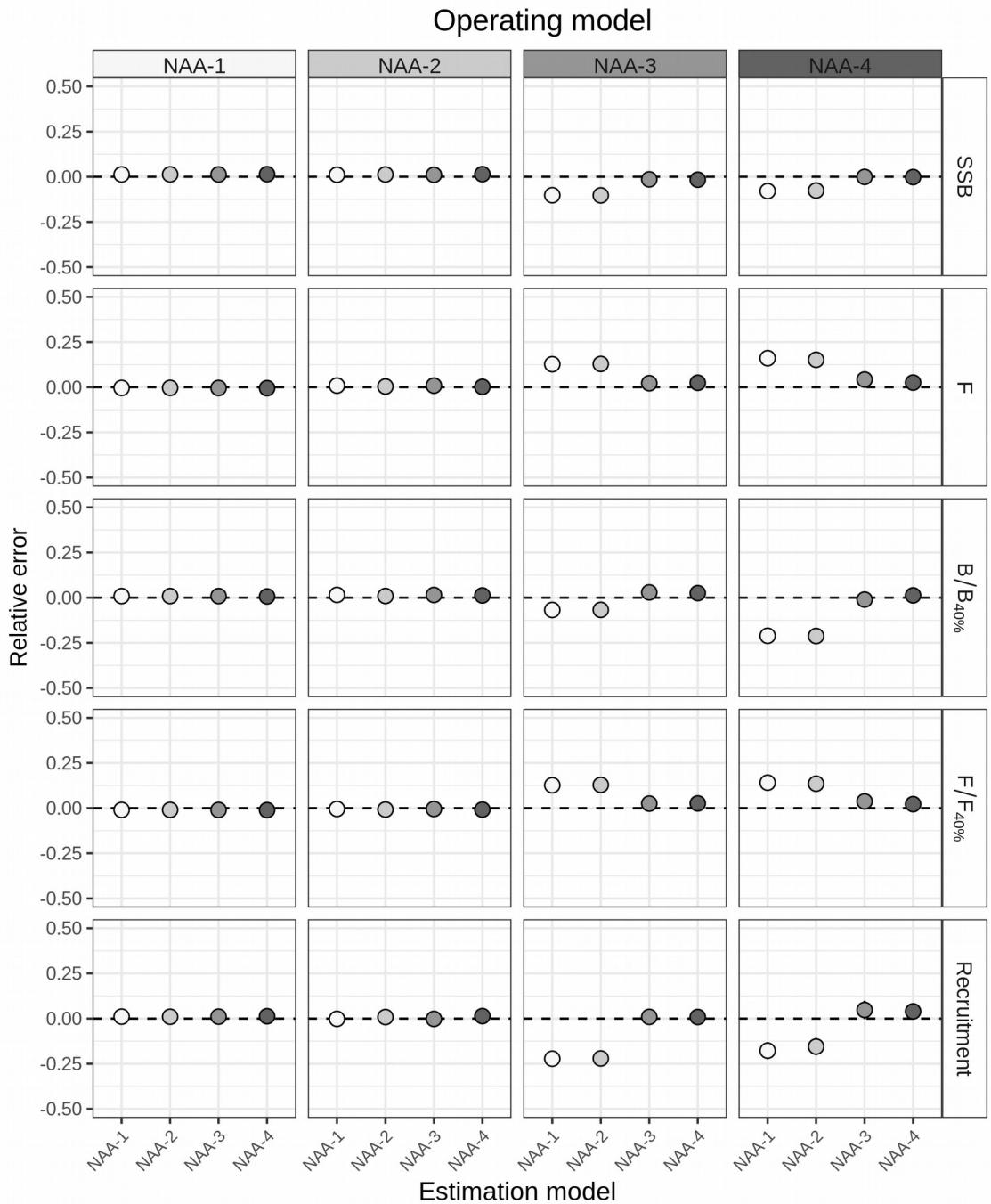
**Figure S8.** Trends in  $F$ , recruitment, and SSB estimated for Icelandic herring using two models of numbers-at-age (NAA) random effects. NAA-1 = only recruitment deviations are independent random effects (most similar to traditional statistical catch-at-age). NAA-4 = all NAA deviations are random effects correlated by age and year (2D AR1). NAA-4 had the lowest AIC (Fig. 1 in main text). The vertical dashed line indicates the terminal year in the assessment.



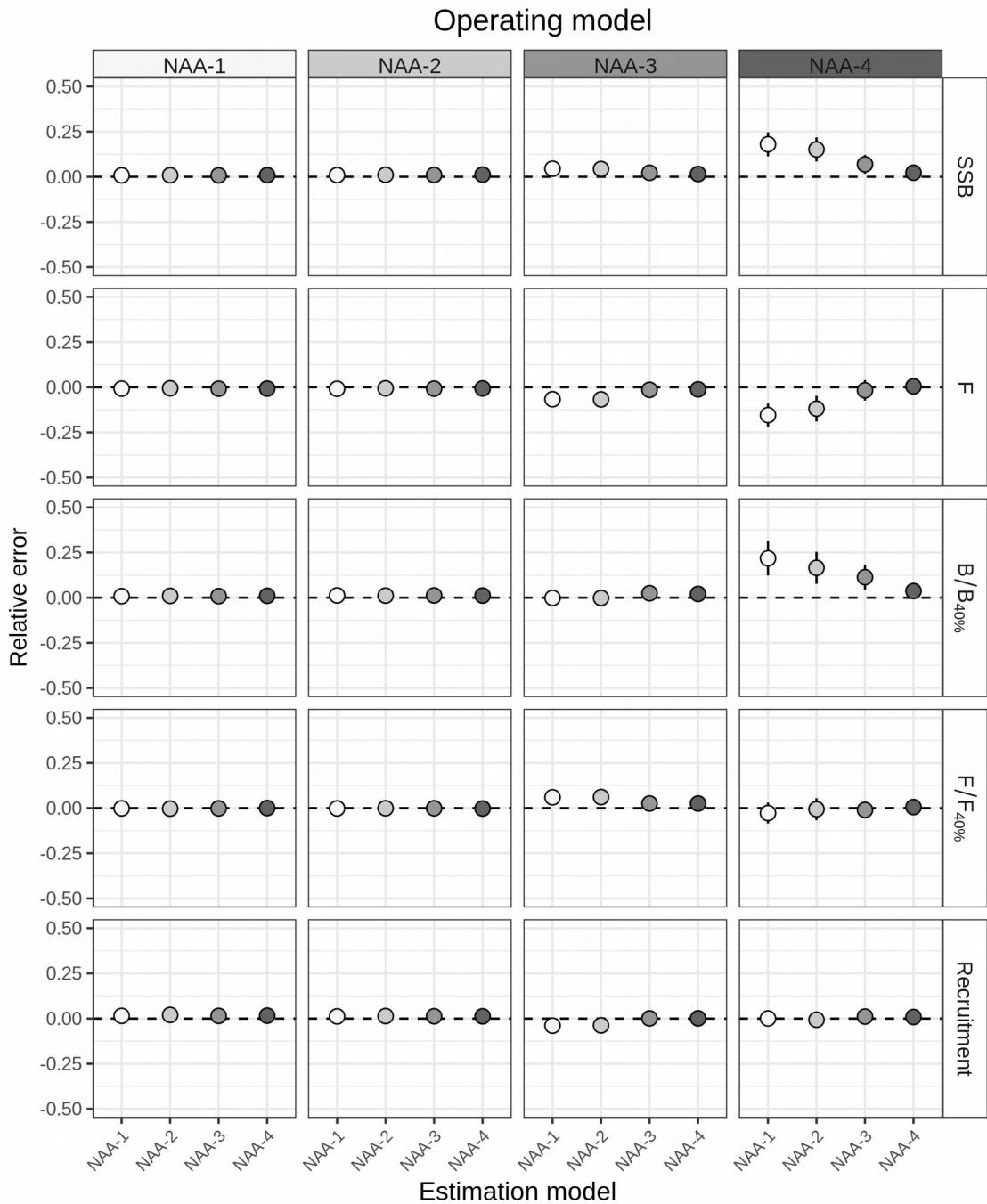
**Figure S9.** Trends in  $F$ , recruitment, and SSB estimated for Georges Bank haddock using two models of numbers-at-age (NAA) random effects. NAA-1 = only recruitment deviations are independent random effects (most similar to traditional statistical catch-at-age). NAA-4 = all NAA deviations are random effects correlated by age and year (2D AR1). NAA-4 had the lowest AIC (Fig. 1 in main text). The vertical dashed line indicates the terminal year in the assessment.



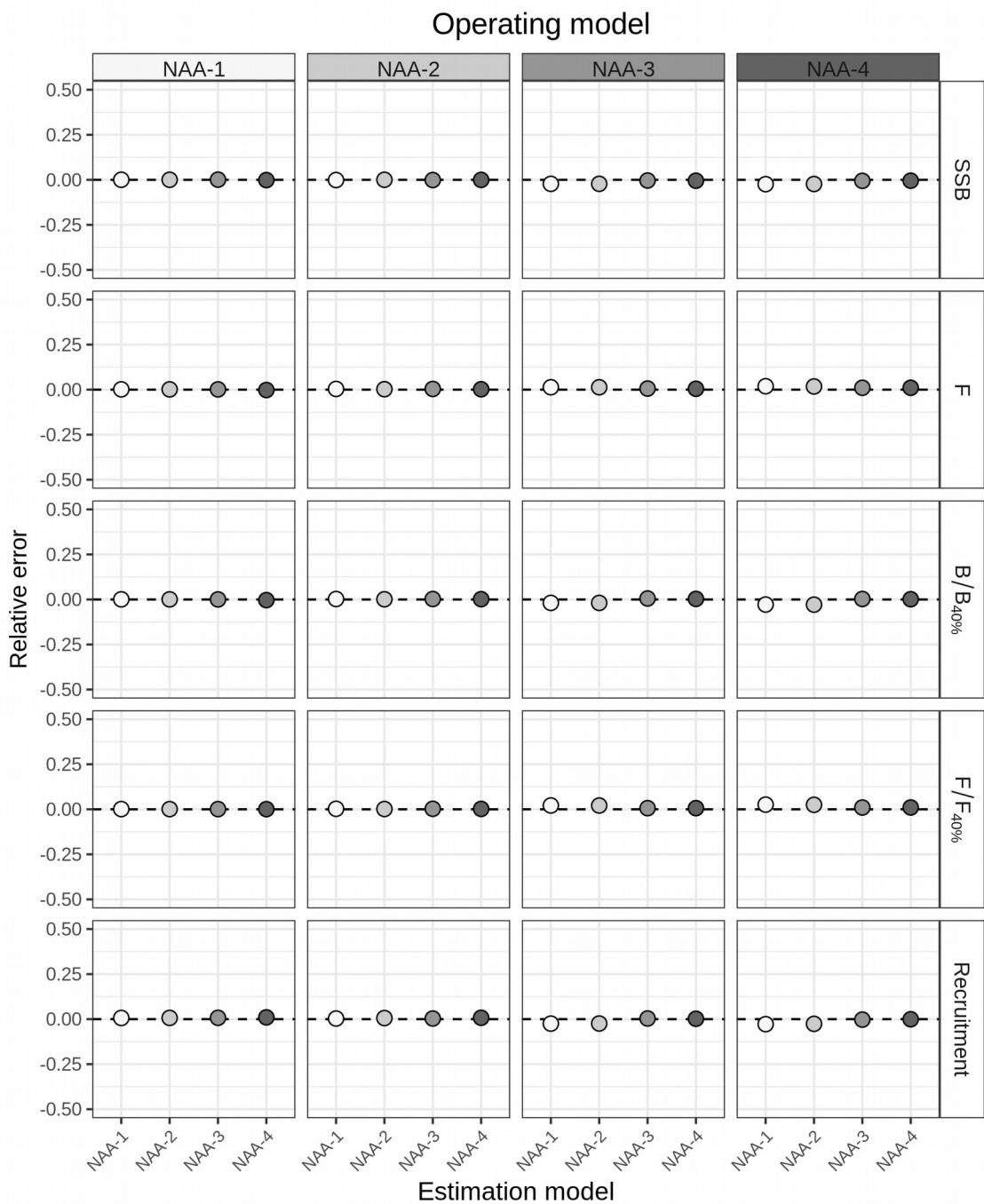
**Figure S10.** Trends in  $F$ , recruitment, and SSB estimated for Georges Bank haddock using two models of selectivity random effects. Sel-1 = time-constant selectivity. Sel-3 = random effect deviations in the logistic selectivity parameters correlated by parameter and year (2D AR1). Sel-3 had the lowest AIC (Fig. 1 in main text). The vertical dashed line indicates the terminal year in the assessment.



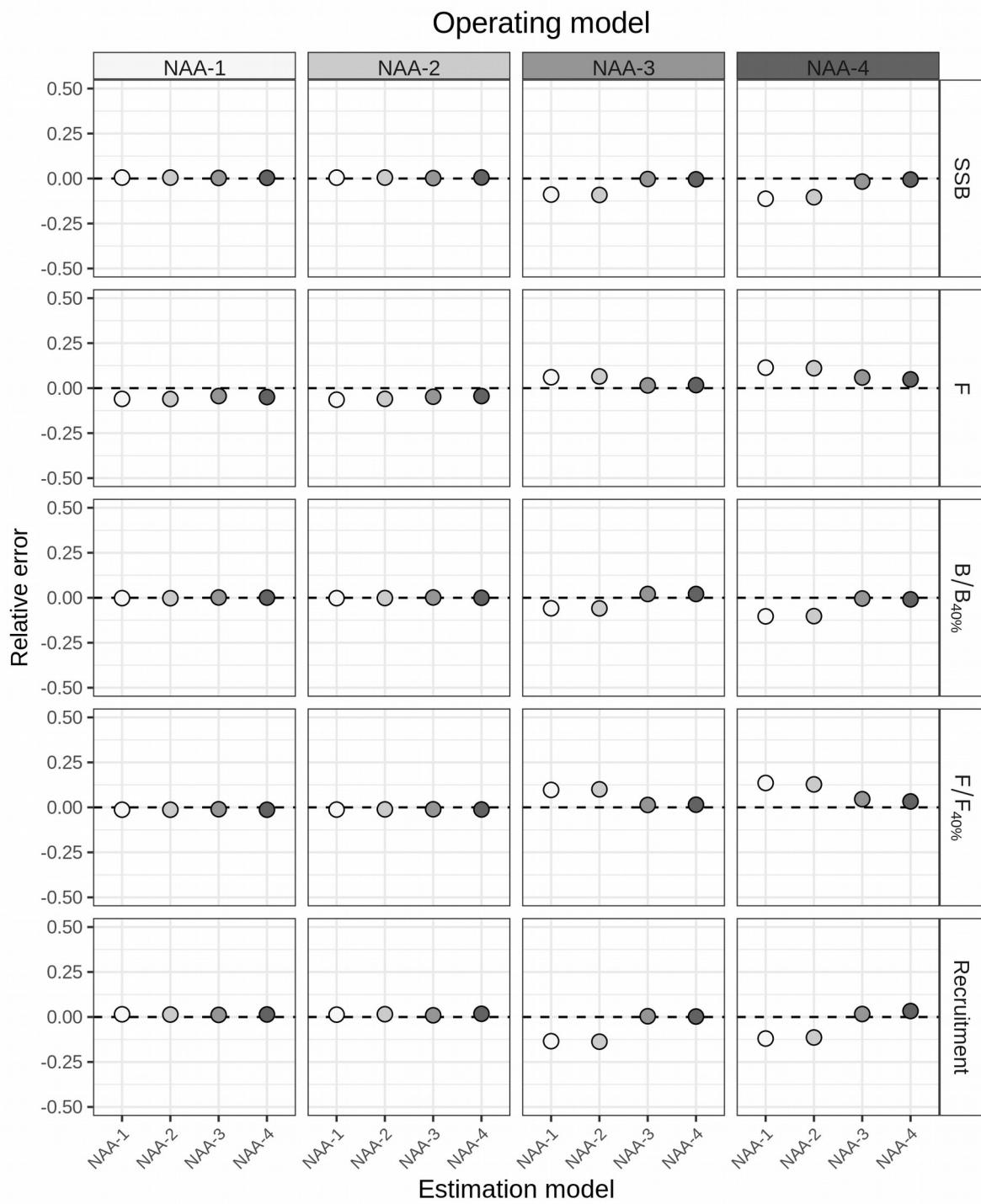
**Figure S11.** Relative error of key quantities estimated for SNEMA yellowtail flounder using four models of numbers-at-age (NAA) random effects. NAA-1 = only recruitment deviations are independent random effects (most similar to Base). NAA-2 = as NAA-1, but with autocorrelated recruitment deviations (AR1). NAA-3 = all NAA deviations are independent random effects. NAA-4 = as NAA-3, but deviations are correlated by age and year (2D AR1). Points without lines indicate that 95% CI are smaller than the points.



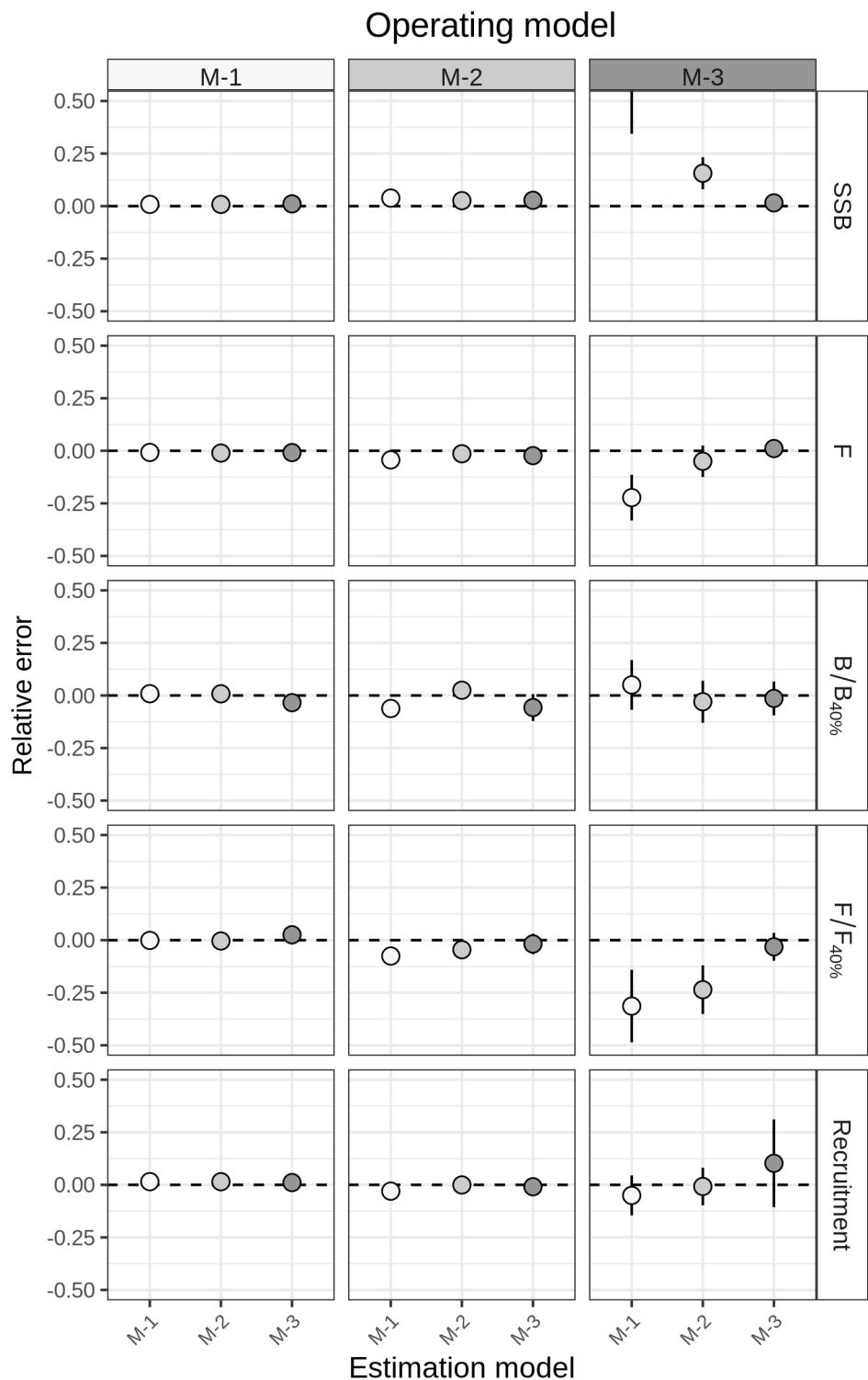
**Figure S12.** Relative error of key quantities estimated for butterfish using four models of numbers-at-age (NAA) random effects. NAA-1 = only recruitment deviations are independent random effects (most similar to Base). NAA-2 = as NAA-1, but with autocorrelated recruitment deviations (AR1). NAA-3 = all NAA deviations are independent random effects. NAA-4 = as NAA-3, but deviations are correlated by age and year (2D AR1). Points without lines indicate that 95% CI are smaller than the points.



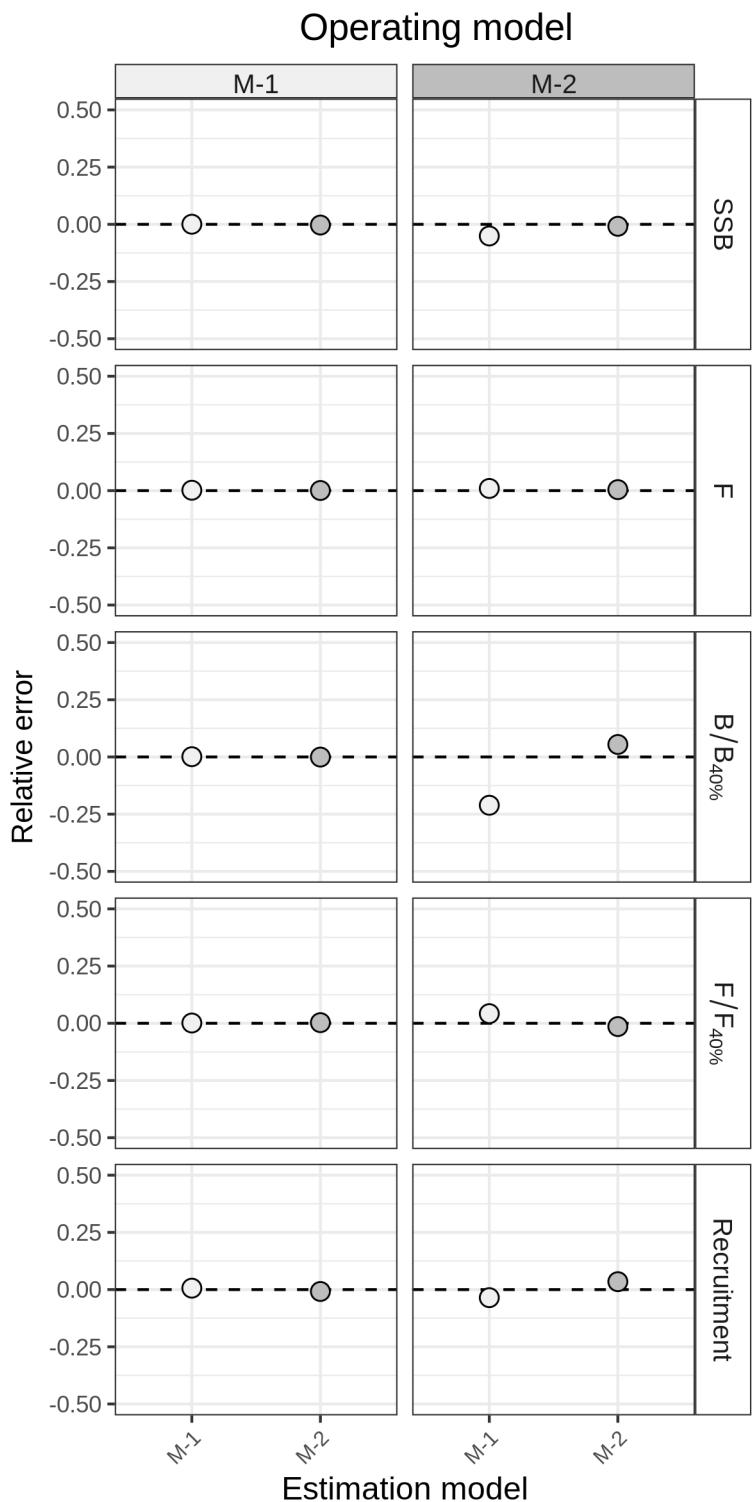
**Figure S13.** Relative error of key quantities estimated for North Sea cod using four models of numbers-at-age (NAA) random effects. NAA-1 = only recruitment deviations are independent random effects (most similar to Base). NAA-2 = as NAA-1, but with autocorrelated recruitment deviations (AR1). NAA-3 = all NAA deviations are independent random effects. NAA-4 = as NAA-3, but deviations are correlated by age and year (2D AR1). Points without lines indicate that 95% CI are smaller than the points.



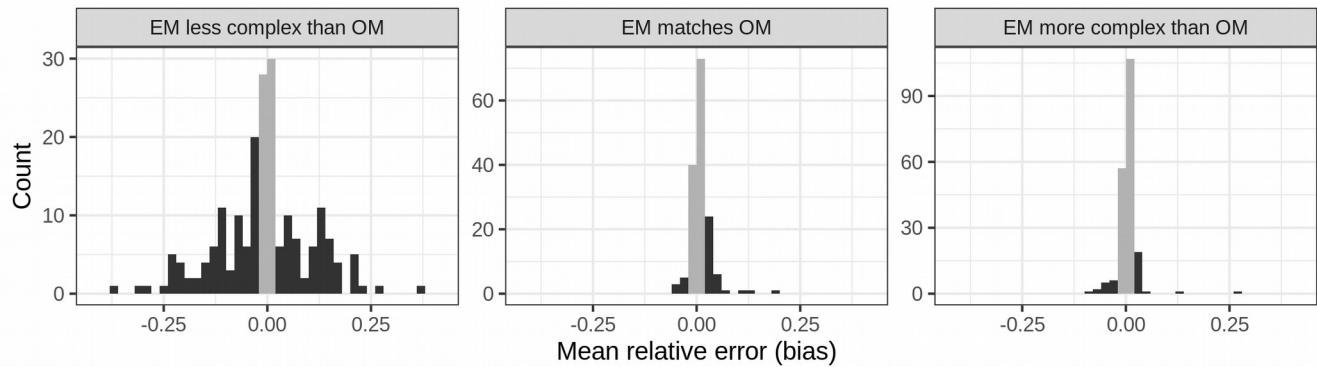
**Figure S14.** Relative error of key quantities estimated for Icelandic herring using four models of numbers-at-age (NAA) random effects. NAA-1 = only recruitment deviations are independent random effects (most similar to Base). NAA-2 = as NAA-1, but with autocorrelated recruitment deviations (AR1). NAA-3 = all NAA deviations are independent random effects. NAA-4 = as NAA-3, but deviations are correlated by age and year (2D AR1). Points without lines indicate that 95% CI are smaller than the points.



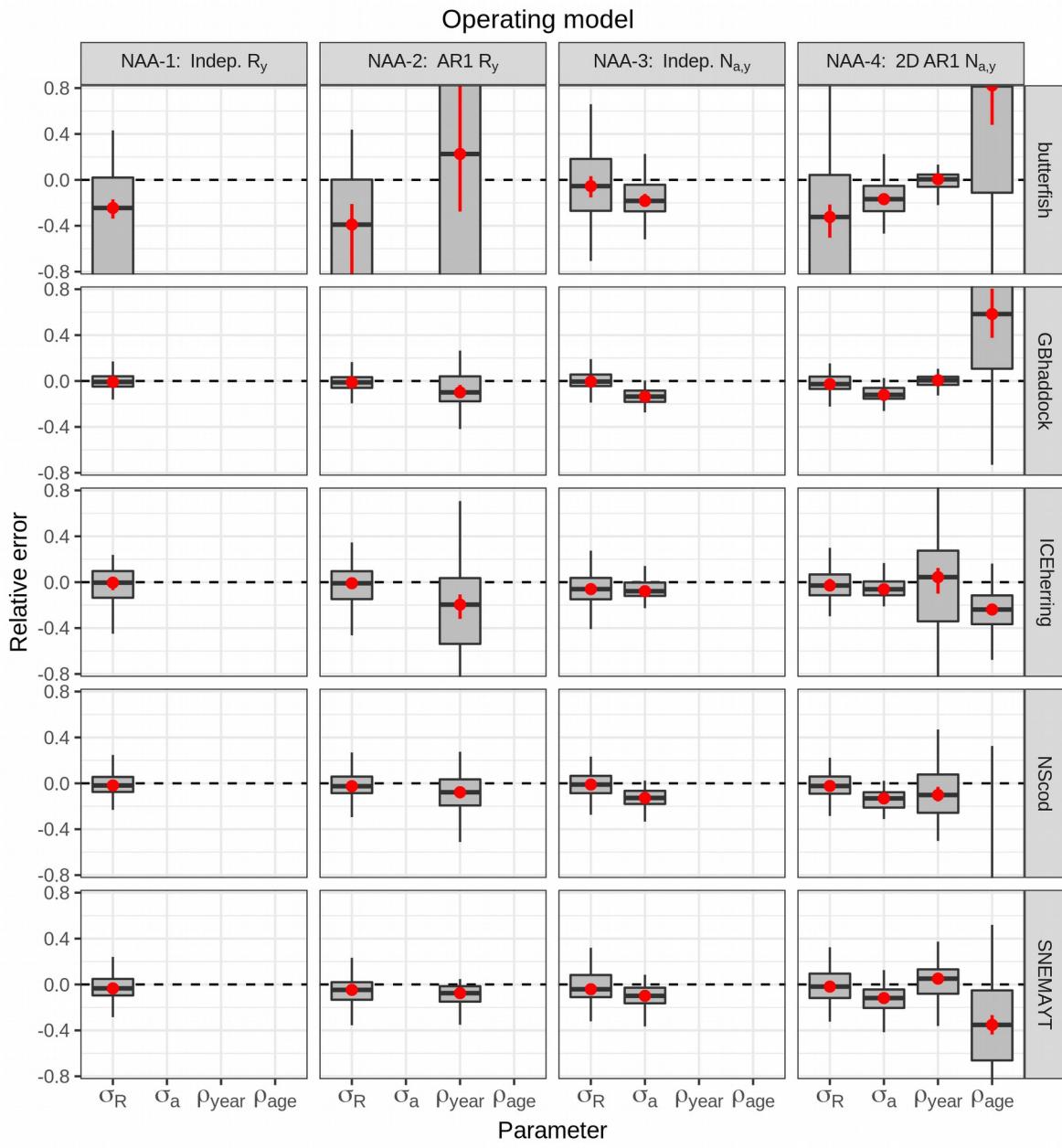
**Figure S15.** Relative error of key quantities estimated for butterfish using three models of natural mortality ( $M$ ) random effects. M-1 = no random effects on  $M$ . M-2 = independent  $M$  deviations. M-3 =  $M$  deviations are correlated by age and year (2D AR1). Points without lines indicate that 95% CI are smaller than the points.



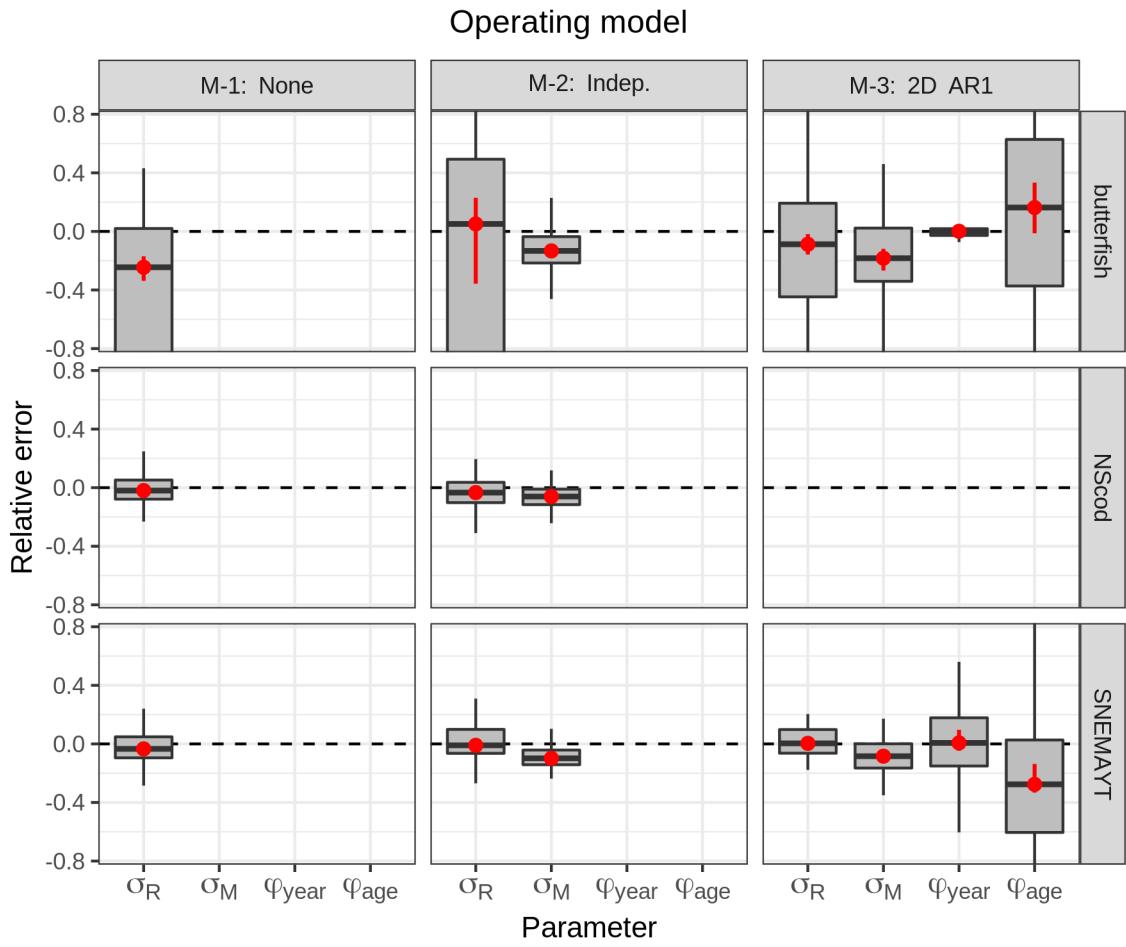
**Figure S16.** Relative error of key quantities estimated for North Sea cod using two models of natural mortality ( $M$ ) random effects. M-1 = no random effects on  $M$ . M-2 = independent  $M$  deviations. M-3, with  $M$  deviations correlated by age and year (2D AR1), did not converge. Points without lines indicate that 95% CI are smaller than the points.



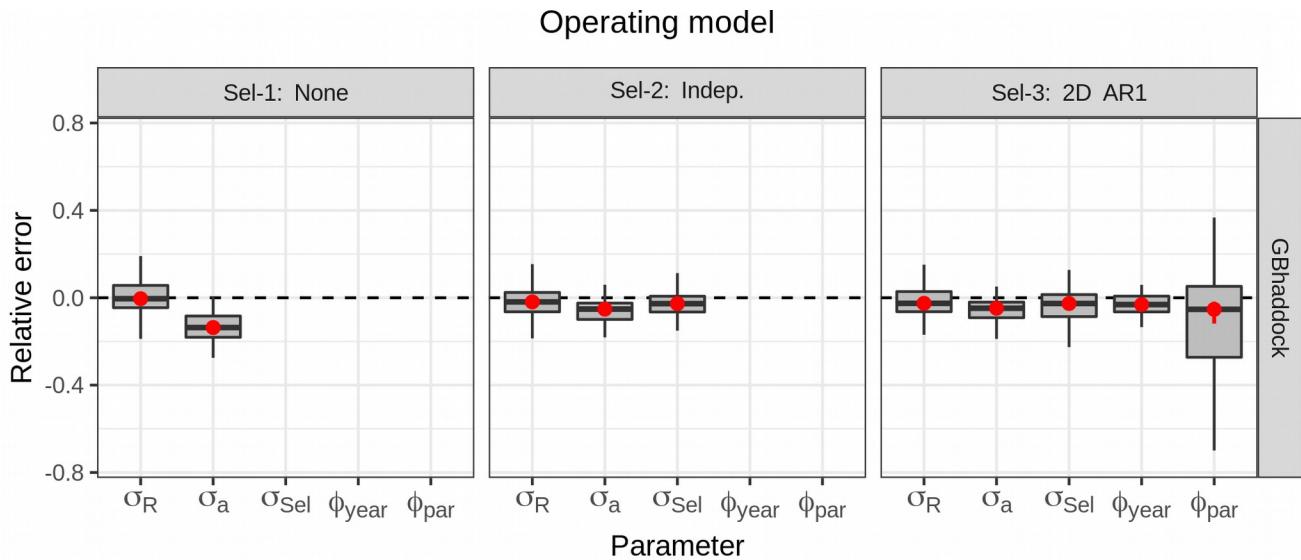
**Figure S17.** Mean relative error (bias) aggregated across all quantities ( $\text{SSB}$ ,  $\text{SSB}/\text{SSB}_{40\%}$ ,  $F$ ,  $F/F_{40\%}$ , and recruitment), stocks, and processes (numbers-at-age, natural mortality, and selectivity) in cross-test simulations when the estimating model (EM) was less complex, matched, or was more complex than the operating model (OM). Bias was frequently notable ( $> 2\%$ , dark shading) when the estimating model was less complex, i.e. treated fewer parameters as random effects, than the operating model, but not vice-versa.



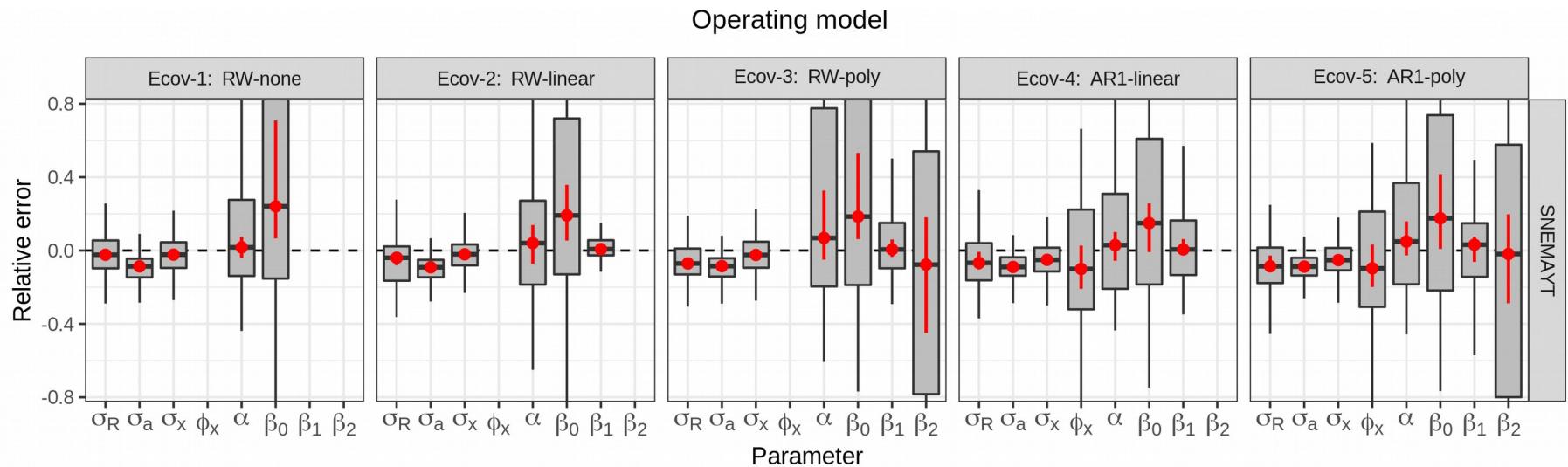
**Figure S18.** Relative error of parameters constraining numbers-at-age (NAA) random effects. Four models were used to simulate 100 datasets keeping fixed effect parameters constant, and then re-fit to each simulated dataset. NAA-1 = only recruitment deviations are independent random effects. NAA-2 = as NAA-1, but with autocorrelated (AR1) recruitment deviations. NAA-3 = all NAA deviations are independent random effects. NAA-4 = as NAA-3, but deviations are correlated by age and year (2D AR1). Relative error was calculated as  $\frac{\hat{\theta}_i}{\theta_i} - 1$ , where  $\hat{\theta}_i$  was the estimate in simulation  $i$  for parameter  $\theta$ , and  $\theta_i$  was the true value (estimate from original dataset). Red points and lines show median relative error with 95% CI.



**Figure S19.** 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. M-1 = no random effects on  $M$ . M-2 = independent  $M$  deviations. M-3 =  $M$  deviations are correlated by age and year (2D AR1). Relative error was calculated as  $\frac{\hat{\theta}_i}{\theta_i} - 1$ , where  $\hat{\theta}_i$  was the estimate in simulation  $i$  for parameter  $\theta$ , and  $\theta_i$  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, M-3 did not converge).

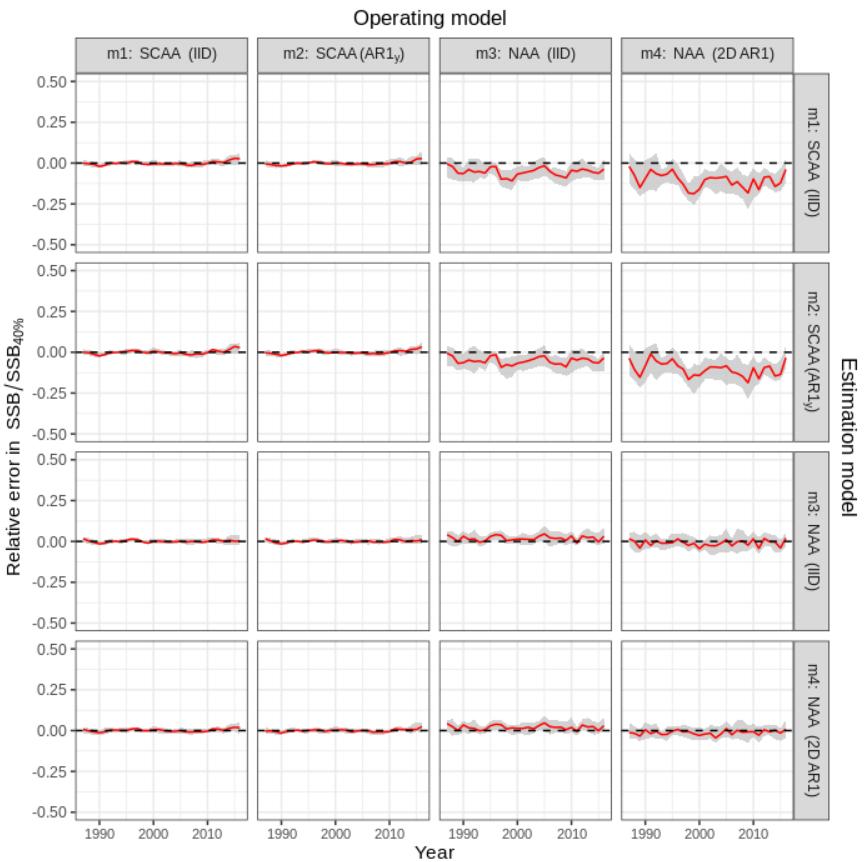
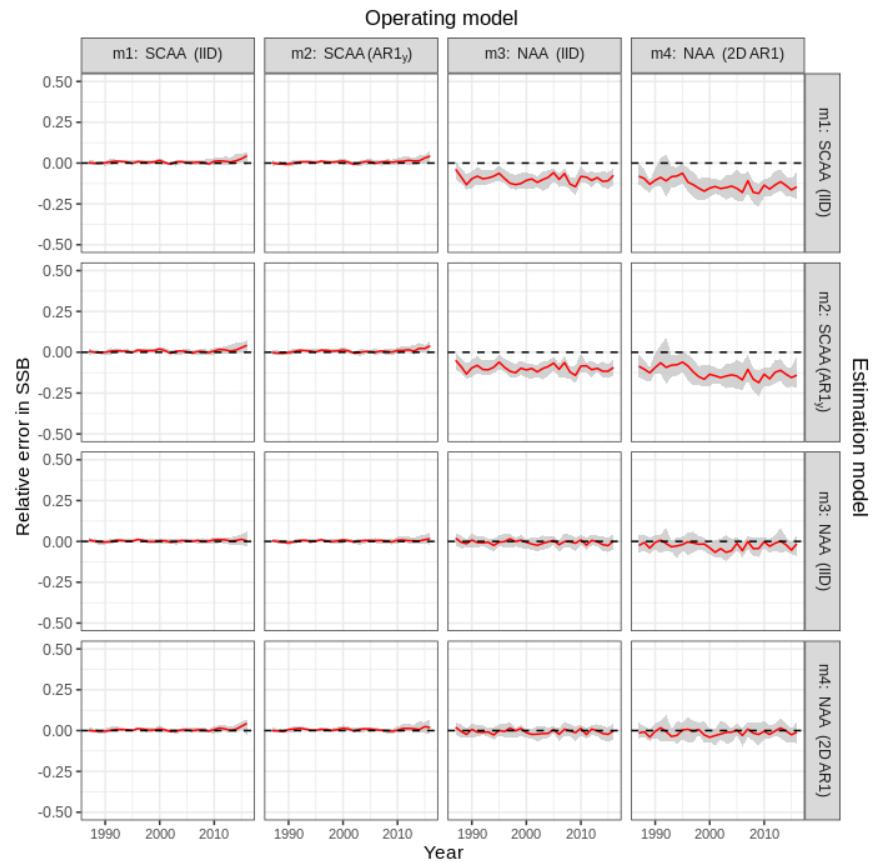


**Figure S20.** Relative error of parameters constraining selectivity random effects for Georges Bank haddock (GBhaddock). Three models were used to simulate 100 datasets keeping fixed effect parameters constant, and then re-fit to each simulated dataset. Sel-1 = no random effects (constant logistic selectivity). Sel-2 = independent selectivity deviations. Sel-3 = selectivity deviations are correlated by parameter and year (2D AR1). Relative error was calculated as  $\frac{\hat{\theta}_i}{\theta_i} - 1$ , where  $\hat{\theta}_i$  was the estimate in simulation  $i$  for parameter  $\theta$ , and  $\theta_i$  was the true value (estimate from original dataset). Red points and lines show median relative error with 95% CI.

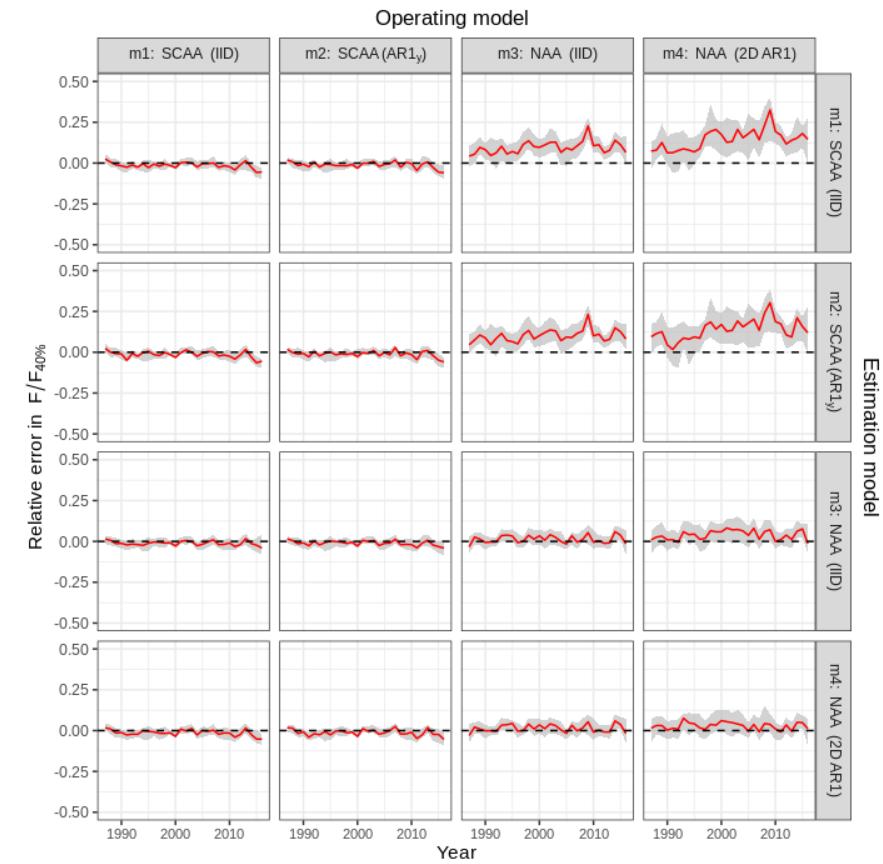
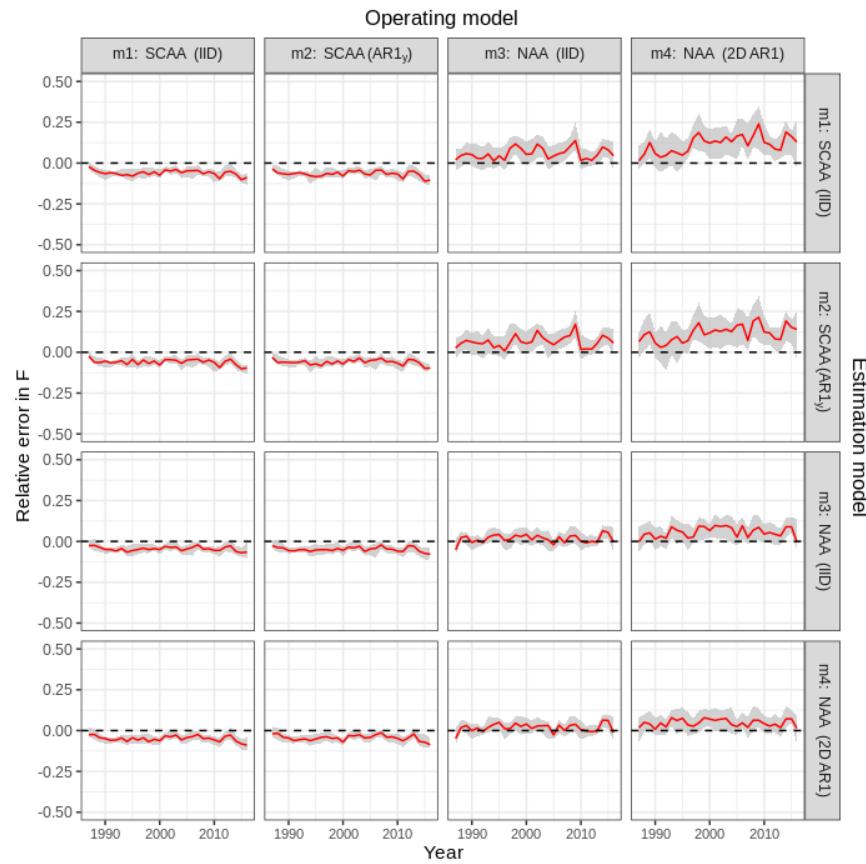


**Figure S21.** Relative error of parameters constraining variation in recruitment for Southern New England-Mid Atlantic yellowtail flounder (SNEMAYT). Five models were used to simulate 100 datasets keeping fixed effect parameters constant, and then re-fit to each simulated dataset. All models estimated recruitment using the Beverton-Holt function, with effect of the Cold Pool Index (CPI) as  $\beta$  terms:

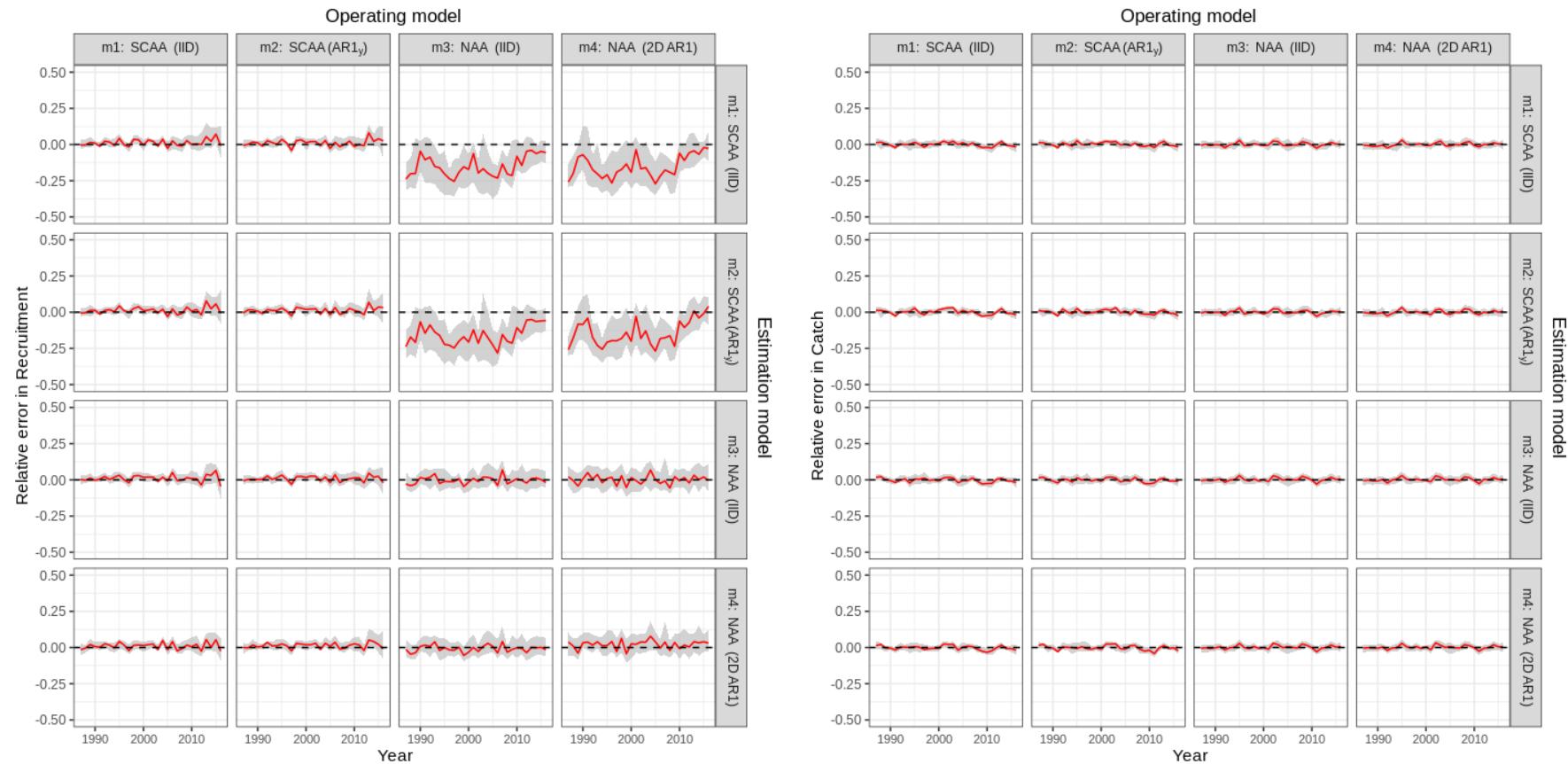
$\hat{R}_{t+1} = \frac{\alpha S_t}{1 + e^{\beta_0 + \beta_1 x_t + \beta_2 x_t^2} S_t}$ . Ecov-1 = CPI modeled as a random walk (RW) with no effect on recruitment ( $\beta_1 = \beta_2 = 0$ ). Ecov-2 = CPI as RW, linear effect on  $\beta$ . Ecov-3 = CPI as RW, 2nd order polynomial effect on  $\beta$ . Ecov-4 = CPI as AR1, linear effect. Ecov-5 = CPI as AR1, polynomial effect. Relative error was calculated as  $\frac{\hat{\theta}_i}{\theta_i} - 1$ , where  $\hat{\theta}_i$  was the estimate in simulation  $i$  for parameter  $\theta$ , and  $\theta_i$  was the true value (estimate from original dataset). Red points and lines show median relative error with 95% CI.



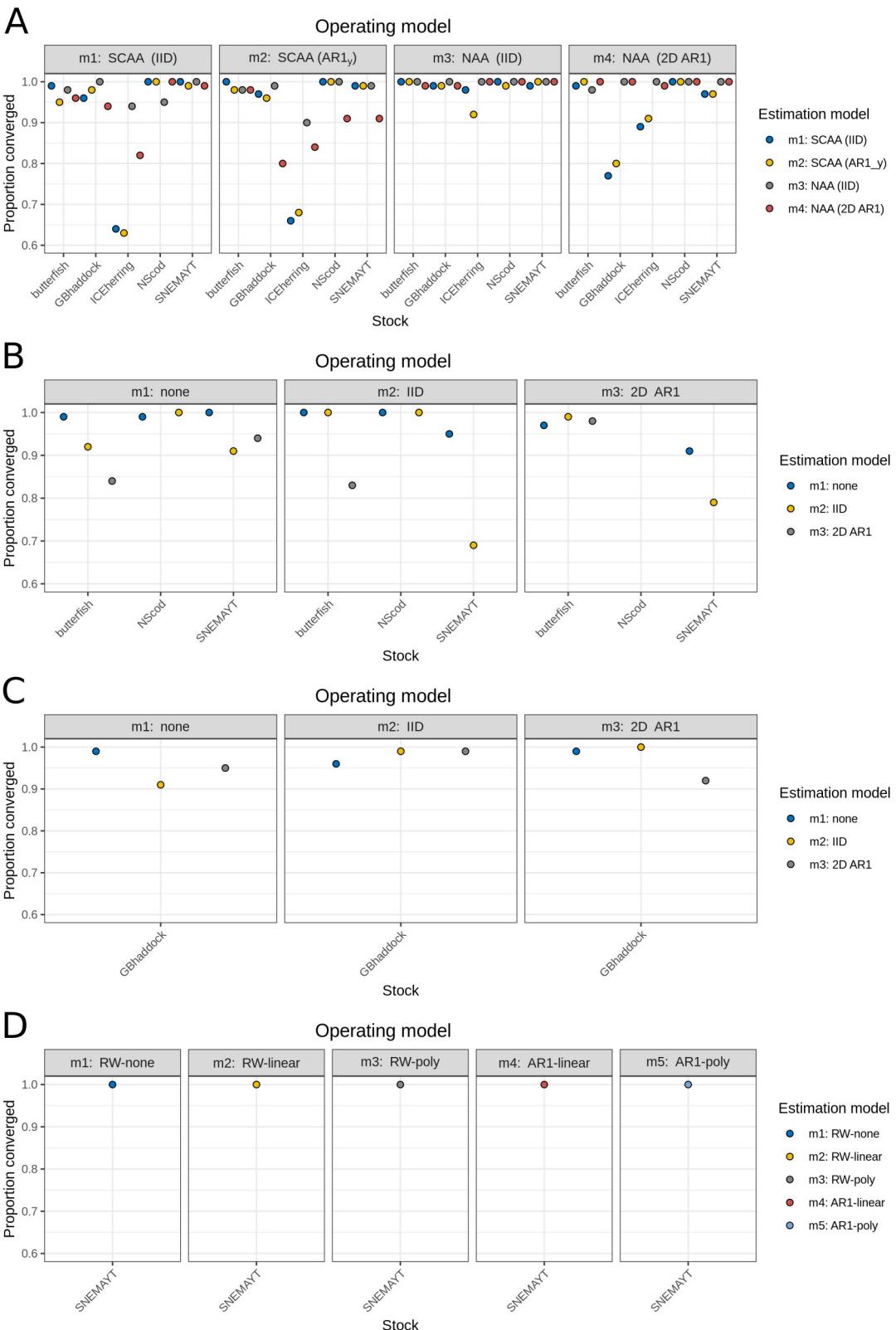
**Figure S22.** Relative error in SSB (left) and SSB / SSB<sub>40%</sub> (right) in simulation cross-tests for the four numbers at age (NAA) models fit to Icelandic herring data. Relative error was calculated as  $\frac{\hat{\theta}_i}{\theta_i} - 1$ , where  $\hat{\theta}_i$  was the estimate in simulation  $i$  for parameter  $\theta$ , and  $\theta_i$  was the true value (estimate from original dataset). Solid red lines and grey shading show the median relative error with 95% CI, calculated within year and across simulations.



**Figure S23.** Relative error in  $F$  (left) and  $F / F_{40\%}$  (right) in simulation cross-tests for the four numbers at age (NAA) models fit to Icelandic herring data. Relative error was calculated as  $\frac{\hat{\theta}_i}{\theta_i} - 1$ , where  $\hat{\theta}_i$  was the estimate in simulation  $i$  for parameter  $\theta$ , and  $\theta_i$  was the true value (estimate from original dataset). Solid red lines and grey shading show the median relative error with 95% CI, calculated within year and across simulations.



**Figure S24.** Relative error in recruitment (left) and catch (right) in simulation cross-tests for the four numbers at age (NAA) models fit to Icelandic herring data. Relative error was calculated as  $\frac{\hat{\theta}_i}{\theta_i} - 1$ , where  $\hat{\theta}_i$  was the estimate in simulation  $i$  for parameter  $\theta$ , and  $\theta_i$  was the true value (estimate from original dataset). Solid red lines and grey shading show the median relative error with 95% CI, calculated within year and across simulations.



**Figure S25.** Proportion of cross-test simulations in which models converged for a) NAA, b)  $M$ , c) selectivity, and d) Ecov random effects. Models with positive definite Hessian matrix were considered to be converged.