

Background
oooo

Inputs
ooo

OM
oooooo
EM
oo

Analysis & Results
oooooooooooooooooooooooooooo

Alert
oo

Results
oooooooooooooooooooooooooooo

Conclusions
oo

Future Work
ooo



Identification and performance of stock-recruitment functions in state space assessment models

Working Paper 1

Greg Britten, Liz Brooks, Tim Miller

2024-02-07

Background

Inputs

OM EM
oooooo oo

Analysis & Results

Alert
00

Results

Conclusions

Future Work

Background

Introduction

- State Space Research Track Working Group (SSRTWG) is investigating performance of the Woods Hole Assessment Model (WHAM)
 - Simulation studies with data generated from operating models (OMs), then fit with a series of estimation models (EMs)
 - This framework allows us to evaluate how the EM fits compare to the known OM “true values”, evaluate model selection, bias, precision, etc.

Terms of Reference (TORs) Addressed

- TOR 2: Investigate the efficacy of estimating stock-recruit functions within state-space models and their utility in generating scientific advice.
 - TOR 3: Develop guidelines for including ecosystem and environmental effects in assessment models and how to treat them for generating biological reference points and scientific advice.

Background
ooo●

Inputs
ooo

OM
oooooo EM
oo

Analysis & Results
oooooooooooooooooooooooooooo

Alert
oo

Results
oooooooooooooooooooooooooooo

Conclusions
oo

Future Work
ooo

Outline

- 'Stock' parameters, fishery and index parameters
- OM factors and simulated data examples
- EM models
- Beta standardization
- Analysis & Results
- Conclusions
- Future Work

Background
oooo

Inputs
●oo

OM
oooooo EM
oo

Analysis & Results
oooooooooooooooooooo

Alert
oo

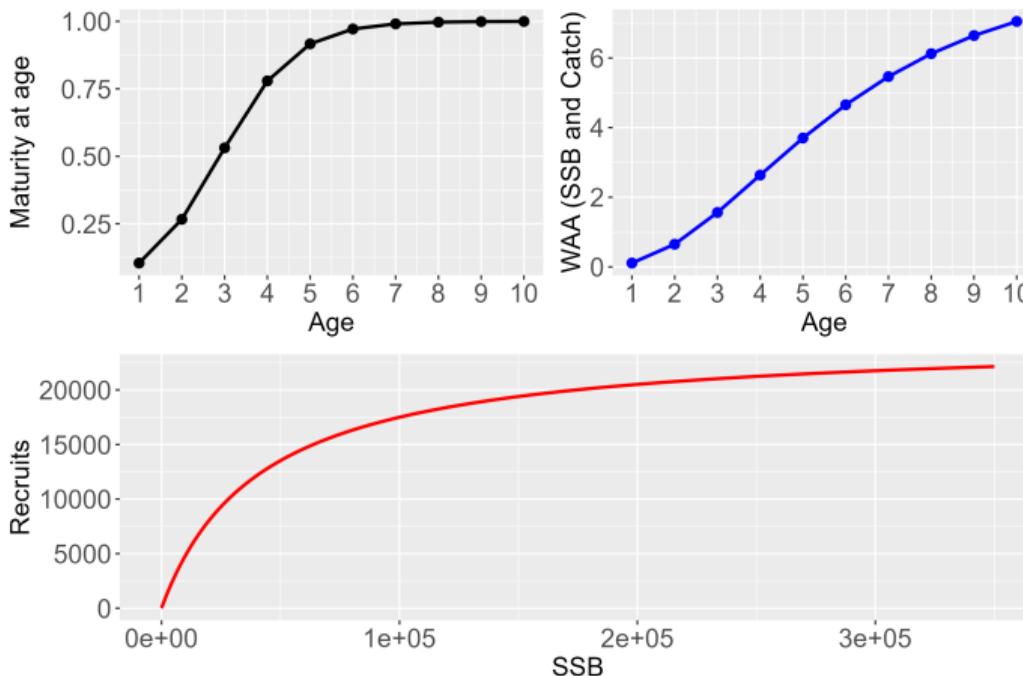
Results
oooooooooooooooooooo

Conclusions
oo

Future Work
ooo

Inputs

Stock parameters for generic gadid



- Natural mortality = 0.2 at all ages
 - Maturity, weight, natural mortality are time-invariant

Background
oooo

Inputs
ooo●

OM
oooooo

EM
oo

Analysis & Results
oooooooooooooooooooo

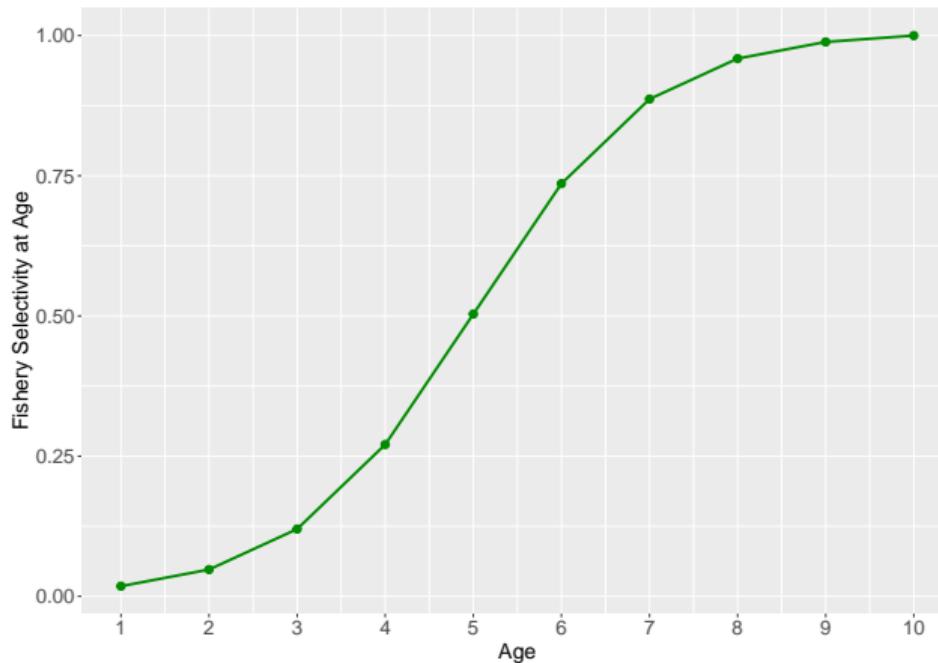
Alert
oo

Results
oooooooooooooooooooo

Conclusions
oo

Future Work
ooo

Fishery and index parameters



- two fishery independent indices were also generated, taking place at 0.25, 0.75 yr
- catchability for both indices was 0.3; selectivity was same as fishery

Background
oooo

Inputs
ooo

OM
●ooooo

EM
oo

Analysis & Results
oooooooooooooooooooooooooooo

Alert
oo

Results
oooooooooooooooooooooooooooo

Conclusions
oo

Future Work
ooo

OM

OM factors

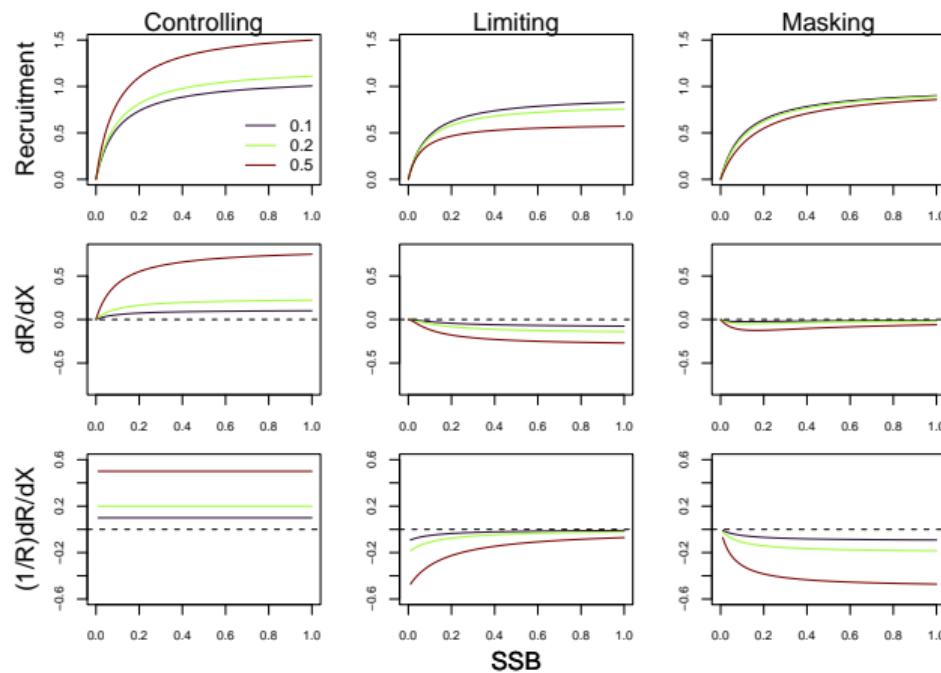
Fhistory	R_sigma	R_cor	Ecov_effect	Ecov_cor	Obs_error
H-MSY	0.1	0.2	0.1	0.2	L
MSY	1	0.8	1	0.8	H

- These levels were combined factorially with 4 stock recruit models (all Beverton-Holt) for 256 different OM
- Factors that did not vary were the observation and process error of the Ecov (both fixed at 0.1)

OM factor for Recruitment variability (R_sig) relative to recent assessments

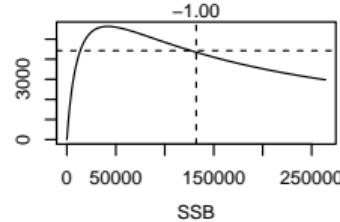
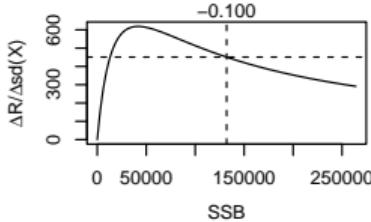
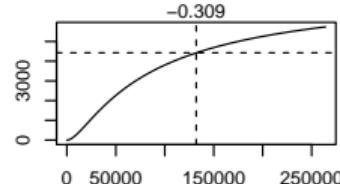
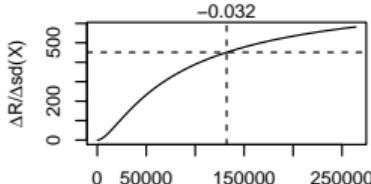
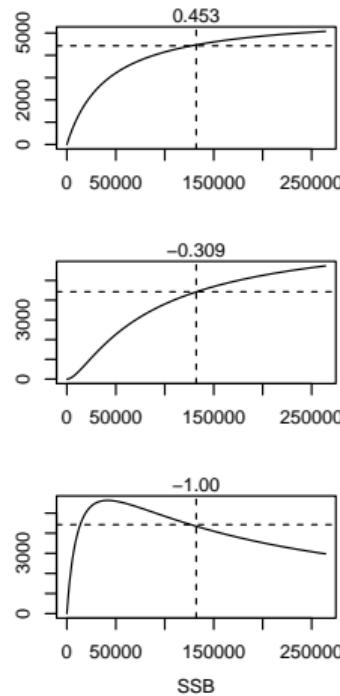
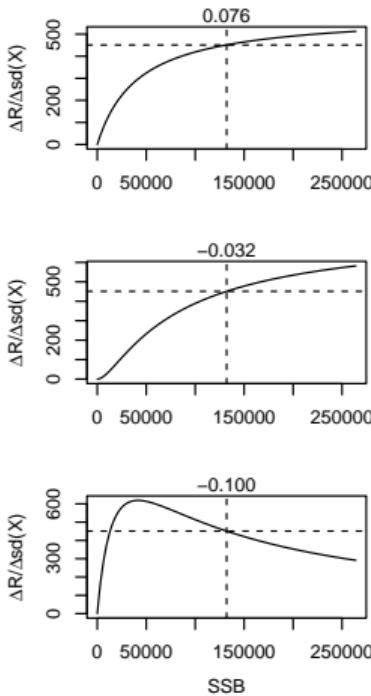
Stock	SRR	Recr_RE	Rec_Cor	sigmaR	sigmaNA	Arho_age	rho_year
GB.Haddock	mean	rec+1	2dar1	1.64	0.38	0.15	0.28
Redfish	BH	rec	iid	1.40	NA	0.00	0.00
GB.Cod	mean	rec+1	iid	1.25	0.52	0.00	0.00
GM.Haddock	mean	rec+1	2dar1	1.06	0.35	0.32	0.52
Mackerel	random.walk	rec+1	2dar1	0.92	0.49	0.50	-0.11
N.Black.Sea.Bass	mean	rec+1	2dar1	0.74	0.81	0.08	0.26
GM.Cod	mean	rec+1	2dar1	0.68	0.25	0.34	0.70
SNE.Cod	BH	rec	iid	0.60	NA	0.00	0.00
Plaice	mean	rec+1	iid	0.51	0.29	0.00	0.00
S.Black.Sea.Bass	mean	rec+1	2dar1	0.51	0.60	-0.13	0.33
GB.Winter.Flounder	mean	rec+1	2dar1	0.34	0.19	0.63	0.69
Bluefish	mean	rec+1	2dar1	0.33	0.16	-0.20	0.77
Butterfish	mean	rec+1	ar1	0.30	0.25	0.00	0.18

OM factor: Beverton-Holt functional (1 with no Ecov effect, 3 with effect)



Background
ooooInputs
oooOM
oooo●○○ EM
ooAnalysis & Results
ooooooooooooooooooooAlert
ooResults
ooooooooooooooooooooConclusions
ooFuture Work
ooo

OM factor: Beta standardization



Some Example OM Simulated Data (100 data sets for each OM)

- still need to make this

Background Inputs OM EM Analysis & Results Alert Results Conclusions Future Work

EM

EM models

- All OMs used Beverton-Holt stock recruit relationship (SRR)
- For the EMs, we also fit a mean SRR with random effects
- There were 6 total EMs fit to each simulated data set (100) for all 256 OMs
- 153,600 fitted models X 2 for β unstandardized and β standardized cases

EM	SRR	Ecov.How	Ecov	EM_mod
1	Mean	0	None	Mean_0
2	Mean	1	Controlling	Mean_1
3	BH	0	None	BH_0
4	BH	1	Controlling	BH_1
5	BH	2	Limiting	BH_2
6	BH	4	Masking	BH_4

Background
oooo

Inputs
ooo

OM
oooooo EM
oo

Analysis & Results
●oooooooooooooooooooooooooooo

Alert
oo Results
oooooooooooooooooooo

Conclusions
oo Future Work
ooo

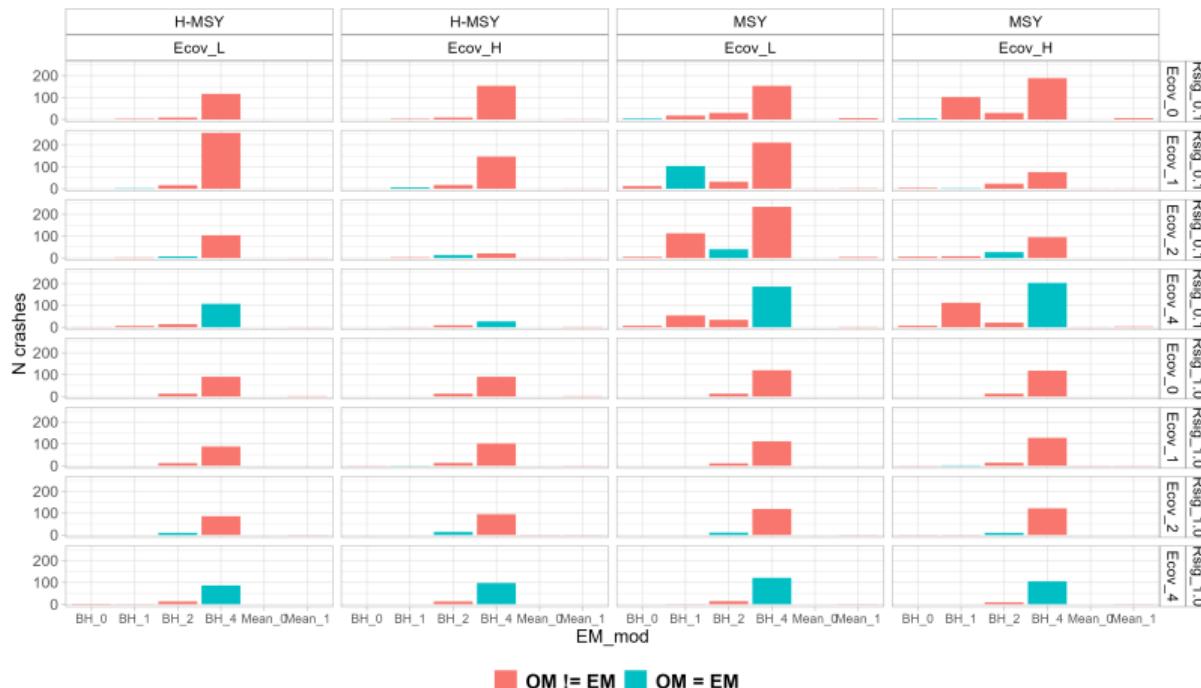
Analysis & Results

Analyses

1. Convergence of the estimating models
2. Model identifiability of an underlying stock recruitment model and/or an underlying relationship between environmental covariate
3. ΔAIC and model probability
4. Assessment error (recruitment, spawning stock biomass, and Fbar)
5. Bias of estimated parameters
6. Mohn's ρ
7. Projection performance relative to assumptions about the environmental covariate
 - Results will be shown for β unstandardized case, because β standardization had no appreciable effect
 - Most results summarized by relative error (RE): $(\text{Est}-\text{True})/\text{True}$

Crashes

- ~3.3-3.5% crashes, typically for Ecov=Masking ("BH_4") and $\sigma_R=0.1$
 - most problematic for Masking (BH_4) and exacerbated by $\sigma_R=0.1$



Background
oooo

Inputs
ooo

OM
oooooo EM
oo

Analysis & Results
oooo●oooooooooooooooooooo

Alert
oo

Results
oooooooooooooooooooo

Conclusions
oo

Future Work
ooo

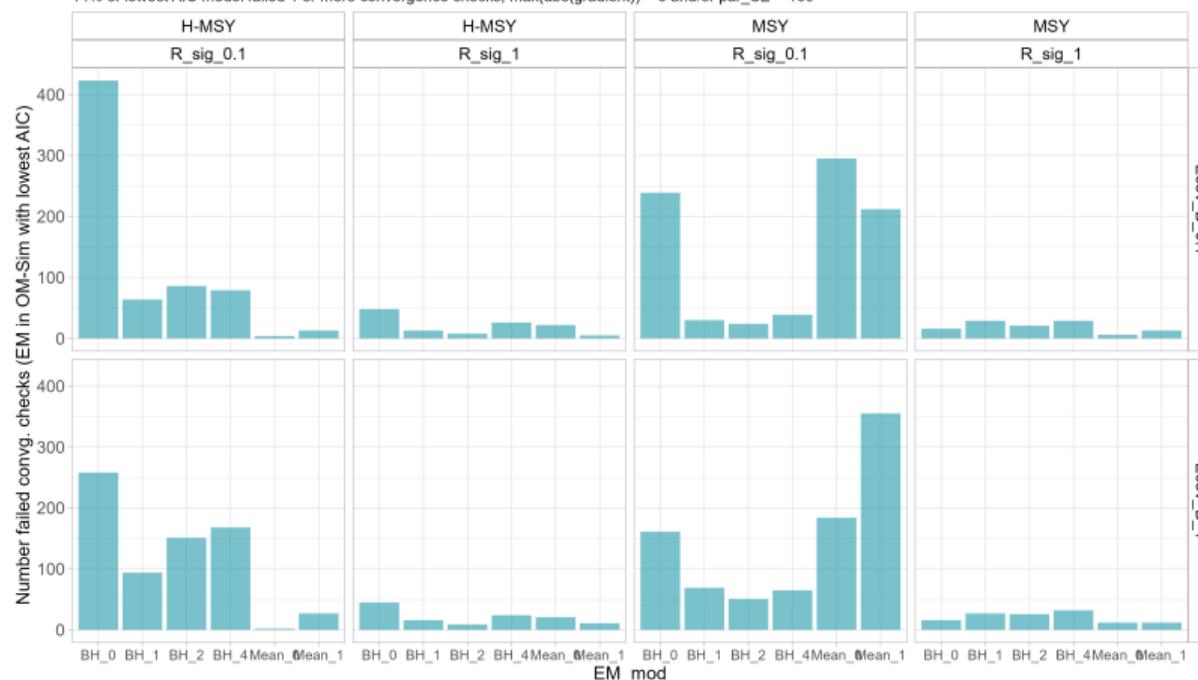
Convergence Criteria

1. the optimization routine (`stats::nlminb`) completed without error
 2. the `stats::nlminb` convergence flag = 0 indicated successful convergence
 3. the maximum absolute value of the gradient of the log-likelihood is $< 1e-6$
 4. `TMB::sdreport` provided non-NA values for all fixed effects standard errors
 5. `TMB::sdreport` provided all standard errors < 100
- Models not satisfying all of these conditions were dropped from Results summaries

Unconverged Models (from EM having 'best' EM per OM-Simulation)

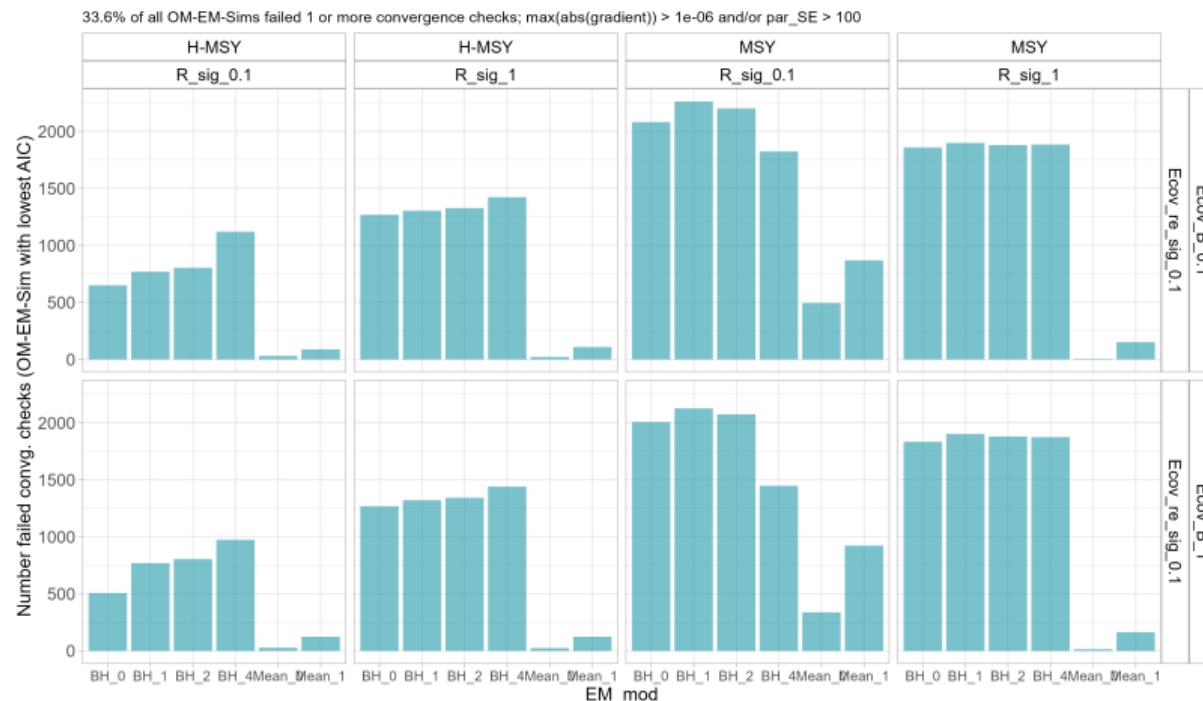
- most problematic for $\sigma_R=0.1$; BH with no Ecov and high fishing contrast (H-MSY) and mean SRR with no contrast (MSY)

14% of lowest AIC model failed 1 or more convergence checks: max(abs(gradient)) > 6 and/or par.SE > 1000.



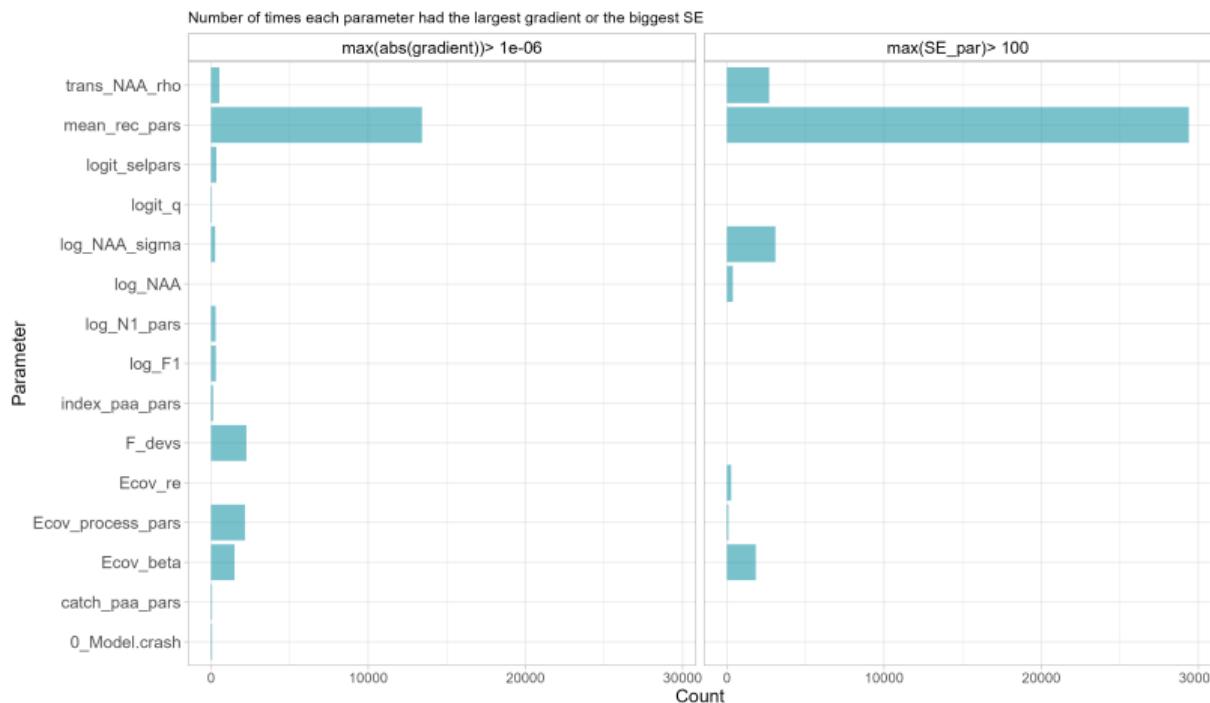
Unconverged Models (considering all EMs per OM-Simulation)

- most problematic for Beverton-Holt SRR, worse with no contrast in fishing history (MSY)



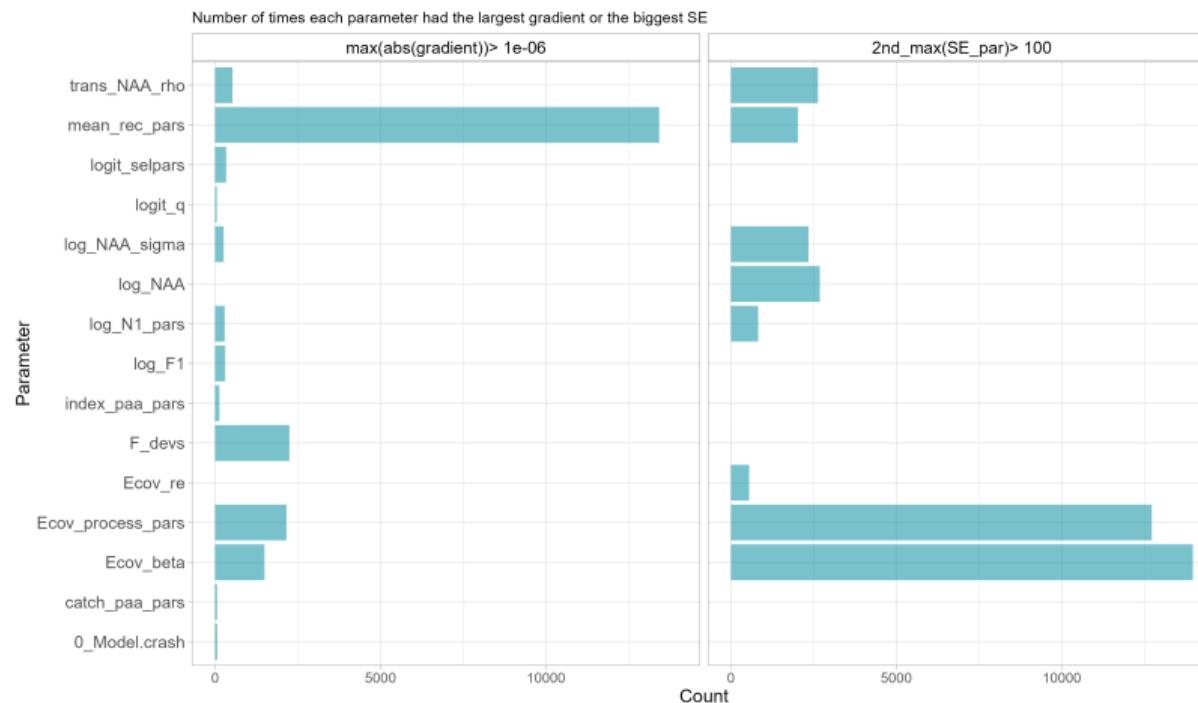
Unconverged Model - The Offending Parameters

- the SRR recruitment parameters led to most convergence failures



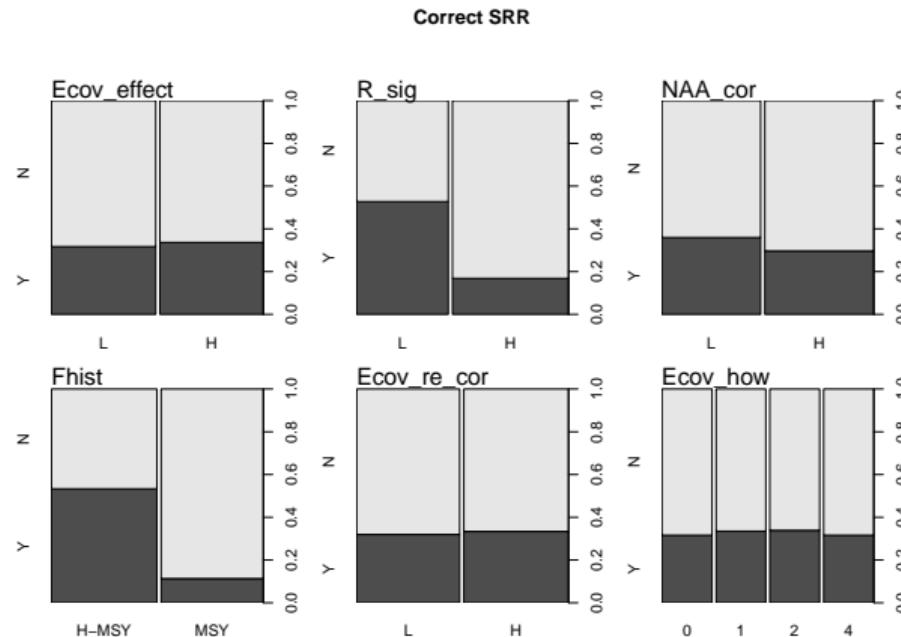
Unconverged Model - Second Most Offending Parameters

- the Ecov parameters led to the second most convergence failures



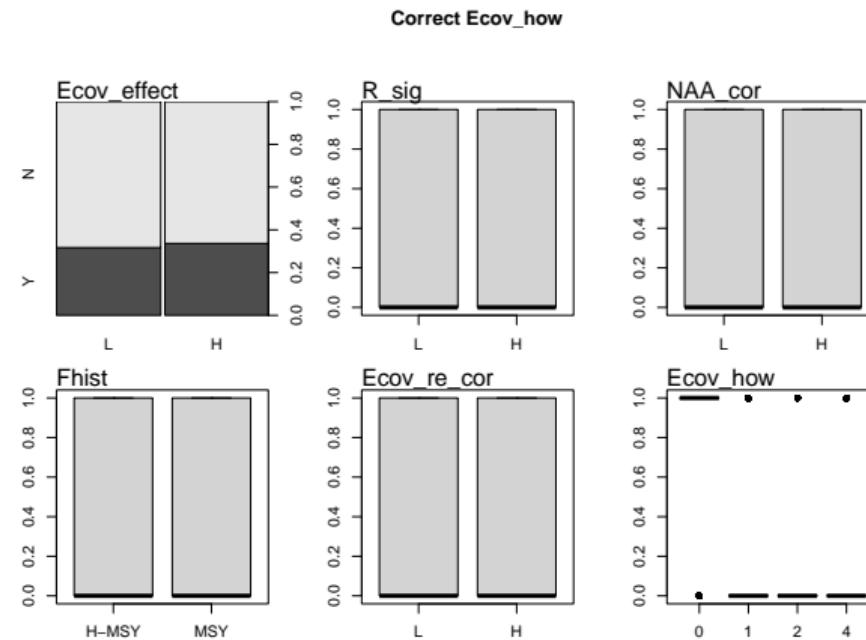
Model Identifiability - Marginal barplots for SRR

- Marginal univariate model selection results for whether an SRR relationship was correctly identified (indicated by dark shaded regions)



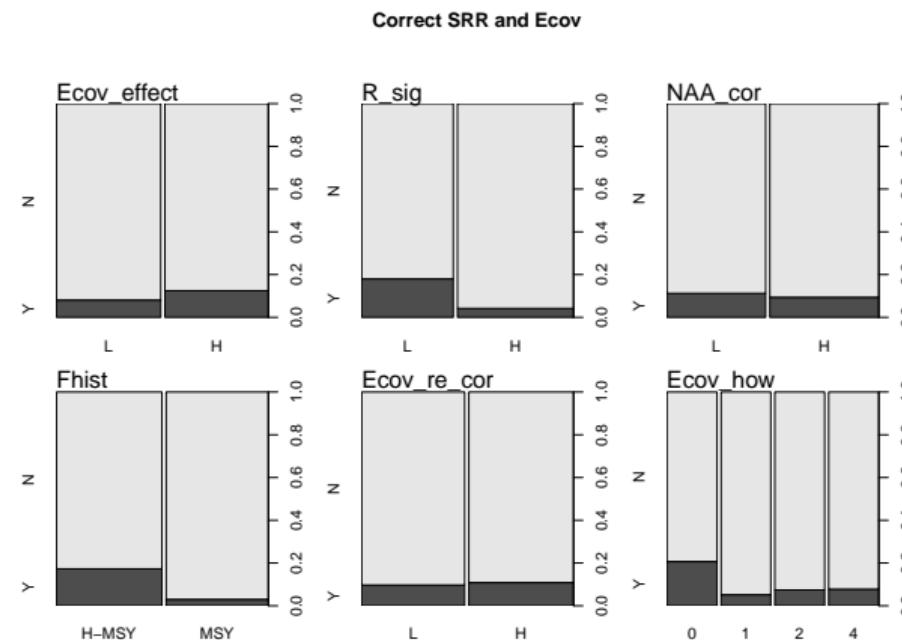
Model Identifiability - Marginal barplots for Ecov_how

- Marginal univariate model selection results for whether an Ecov was correctly identified (indicated by dark shaded regions)



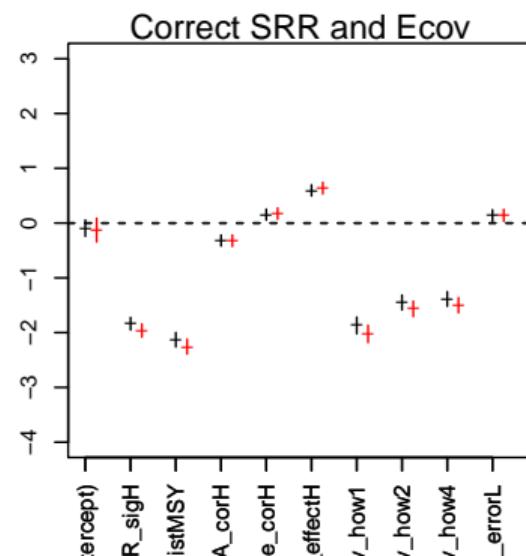
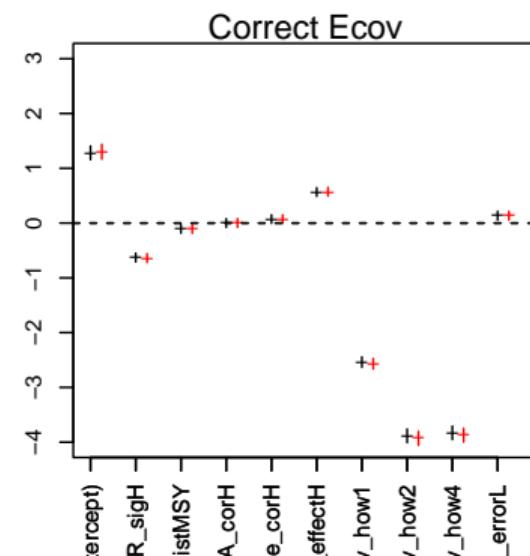
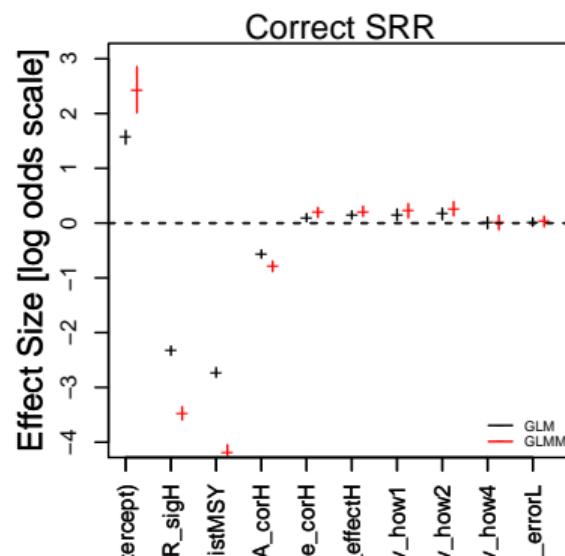
Model Identifiability - Marginal barplots for both SRR & Ecov.How

- Marginal univariate model selection results for whether the SRR AND Ecov were both correctly identified (indicated by dark shaded regions)



Model Identifiability - Effect size

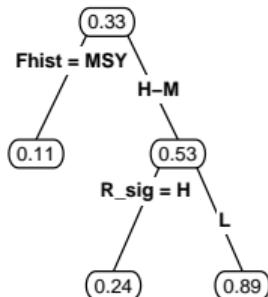
- Effect sizes for proportion correctly identified using a binomial generalized linear model and a binomial generalized linear mixed model with simulation random number seed as a random intercept



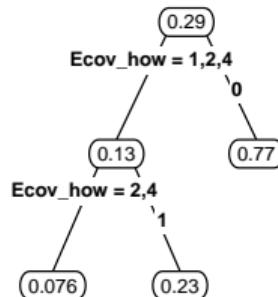
Model Identifiability - Classification tree for identifying EM = OM

- Classification tree analysis of model selection results, with default settings in R package 'rpart'. The complexity parameter is set at $cp = 0.01$, which represents the minimum classification improvement required for any split.

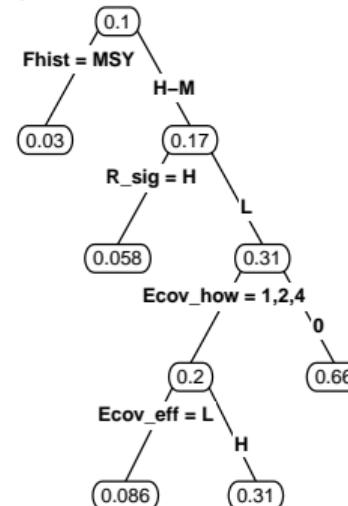
a) correct SRR



b) correct Ecov.How



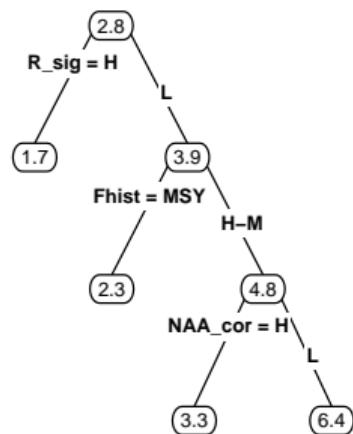
c) correct SRR and Ecov.How



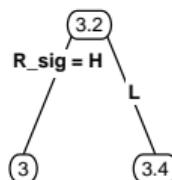
ΔAIC and model probability - Classification tree

- Classification tree analysis of for the difference in AIC among EMs, the rank of EMs, and the model probability of EMs (calculated from Akaike weights).

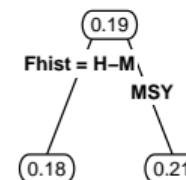
a) dAIC



b) AIC rank

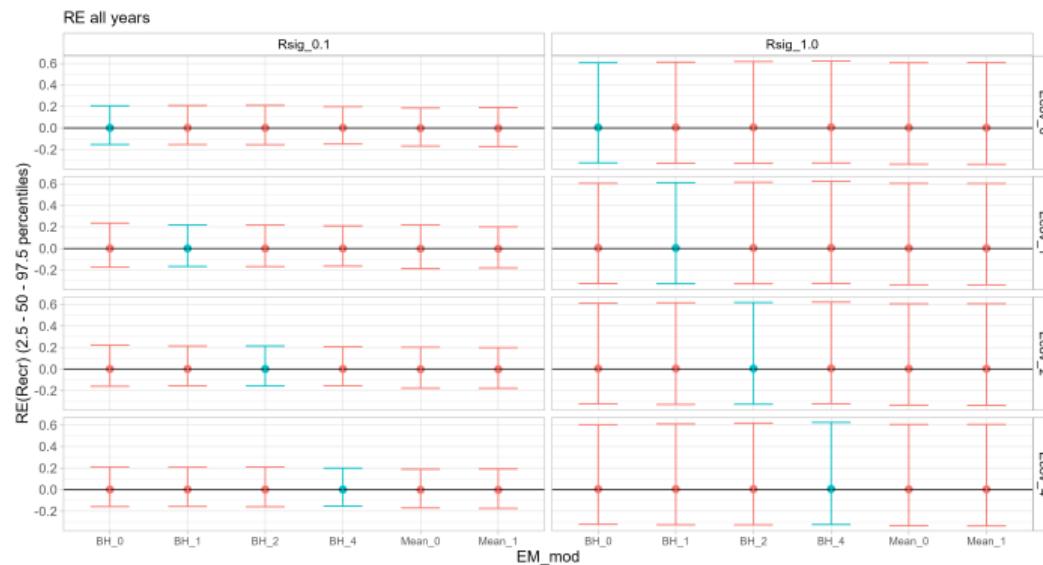


c) Model probability



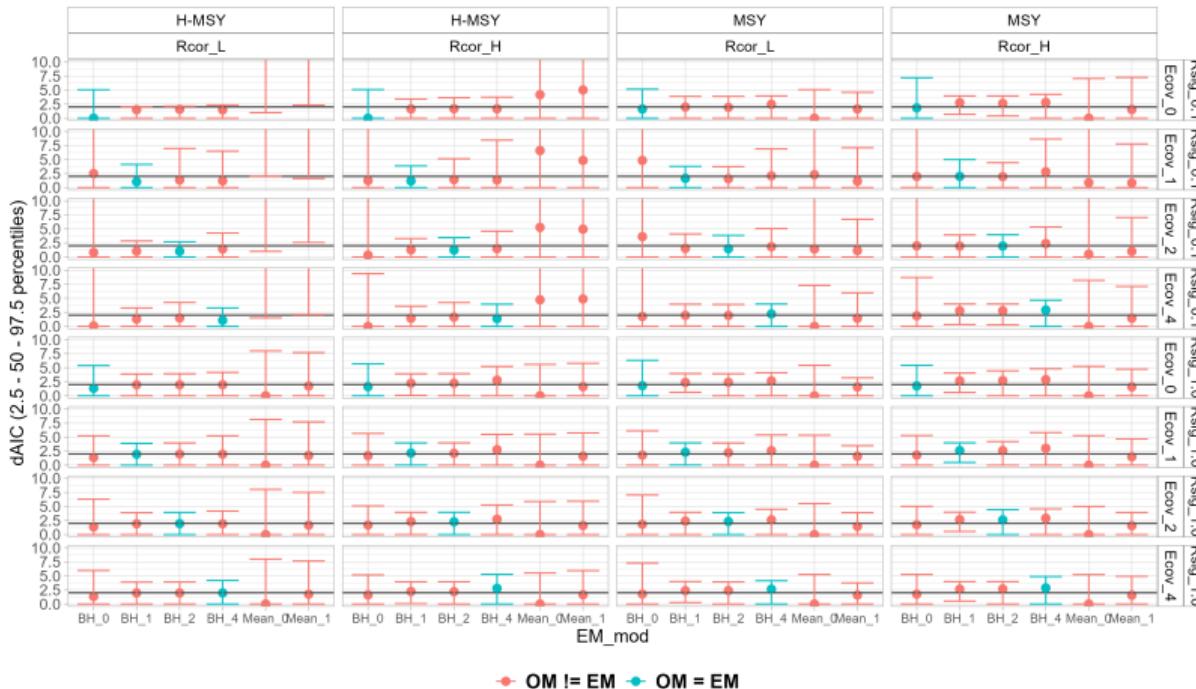
Results layout (example)

- Plots show median relative error (filled circle) with 2.5th and 97.5th percentiles
- OM factors as facets (top and right side)
- EM models on the x-axis
- Blue indicates EM=OM (diagonal of first 4 columns w/in facet blocks), incorrect EMs in red



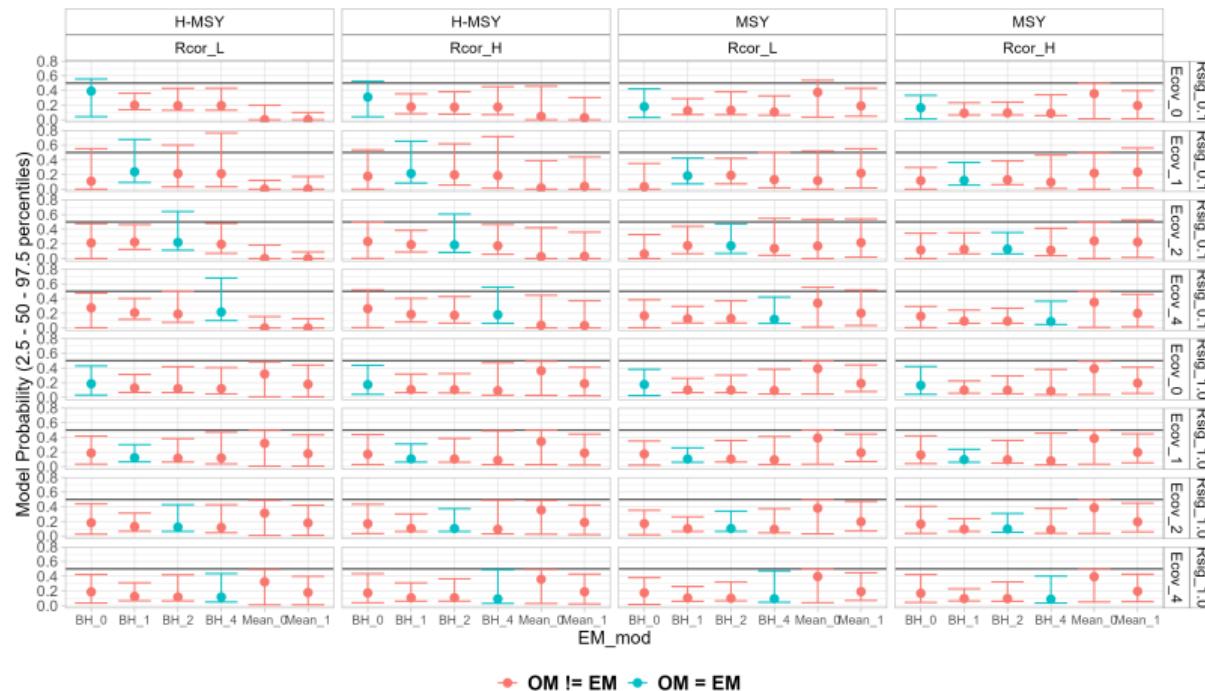
ΔAIC - Median and 95% probability interval

- Smaller ΔAIC when σ_R is larger



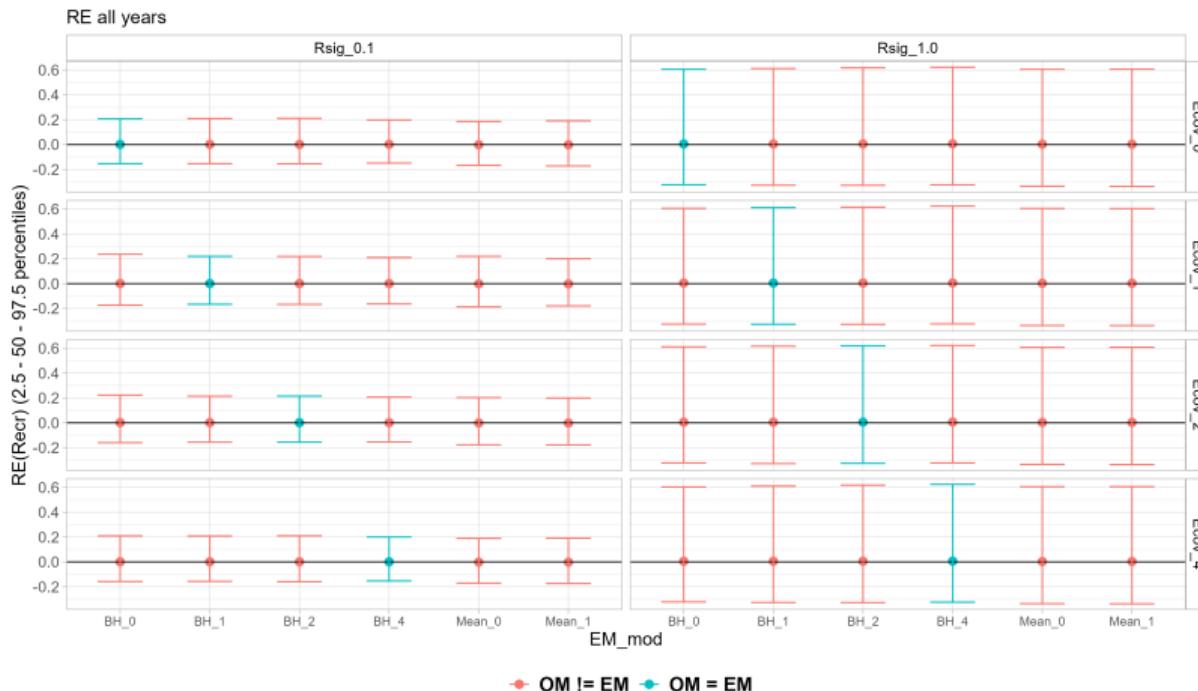
Model probability - Median and 95% probability interval

- True OMs receive more probability with $\sigma_R=0.1$ and high fishing contrast; otherwise, mean SRR received more probability



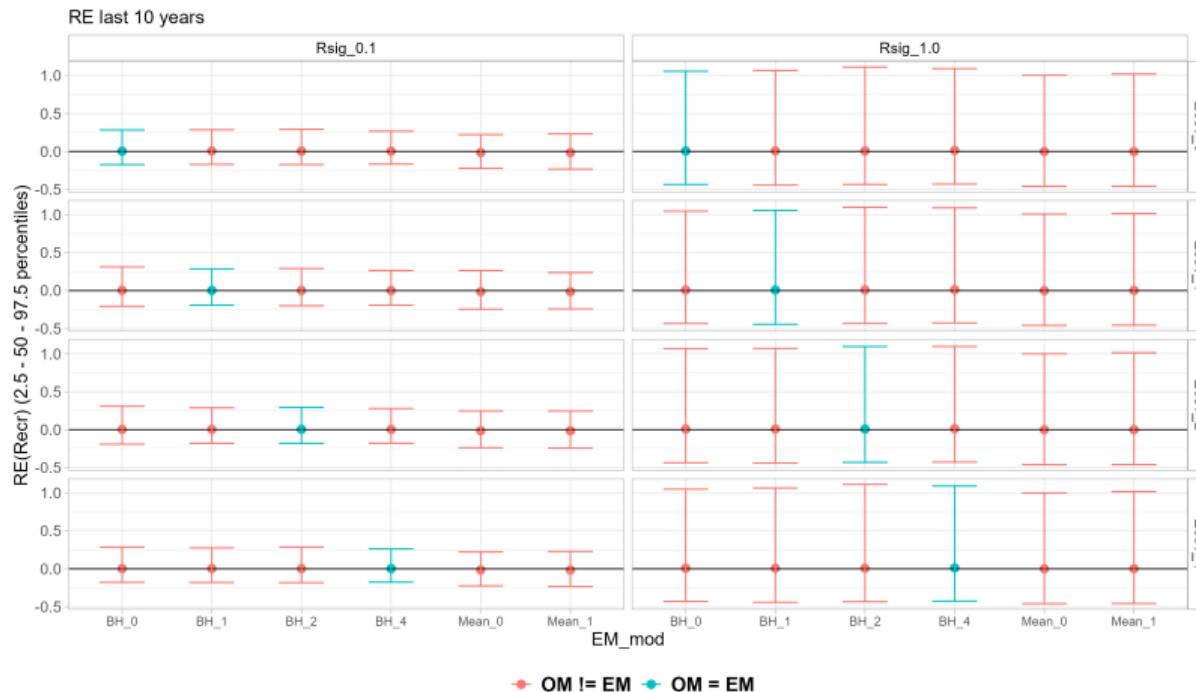
Assessment error - Recruitment (all years)

- Median unbiased, σ_R determined quantile range



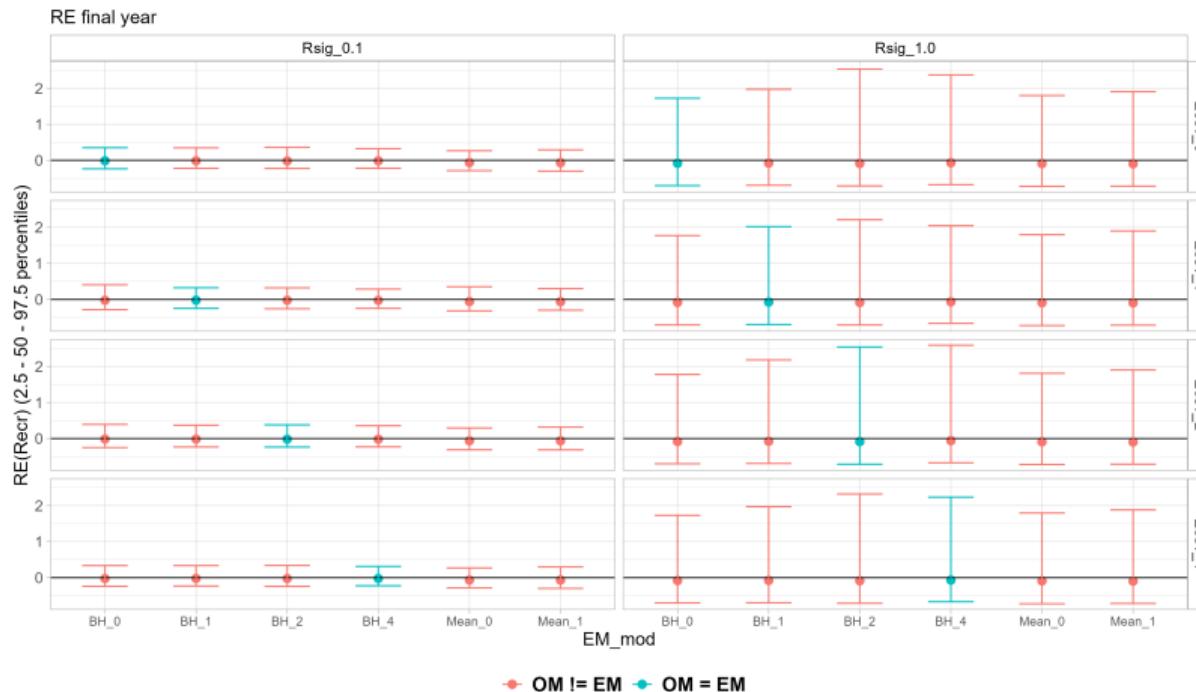
Assessment error - Recruitment (last 10 years)

- Median unbiased, σ_R determined quantile range



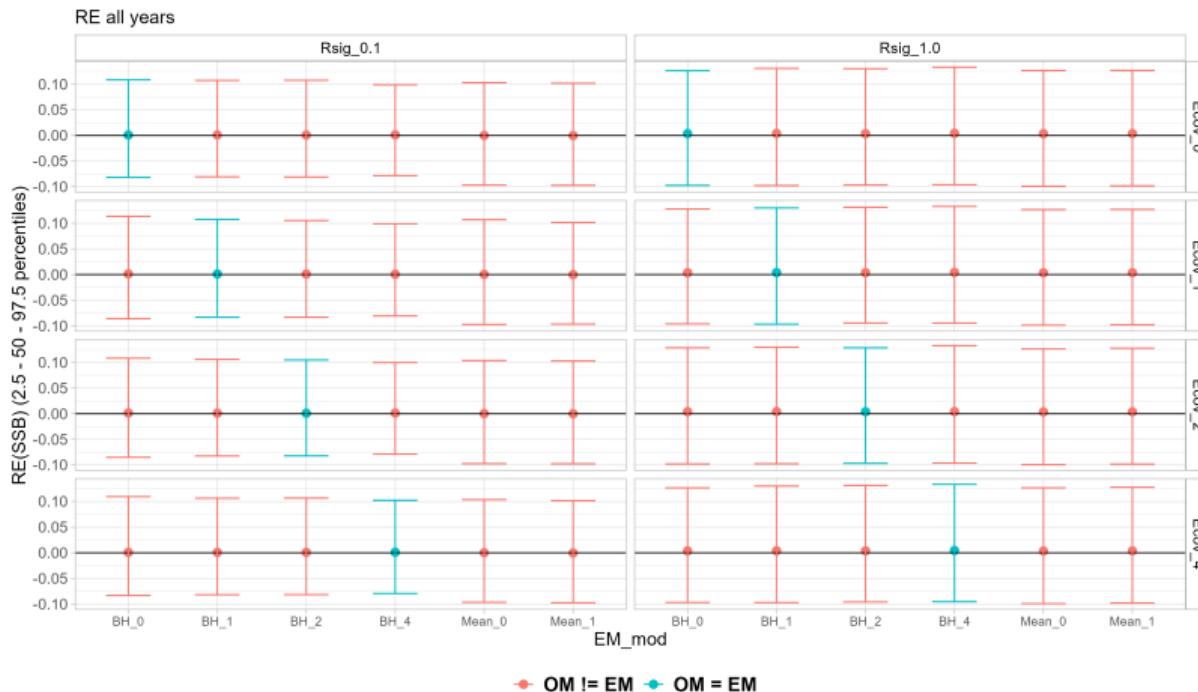
Assessment error - Recruitment (terminal year)

- Median unbiased, σ_R determined quantile range



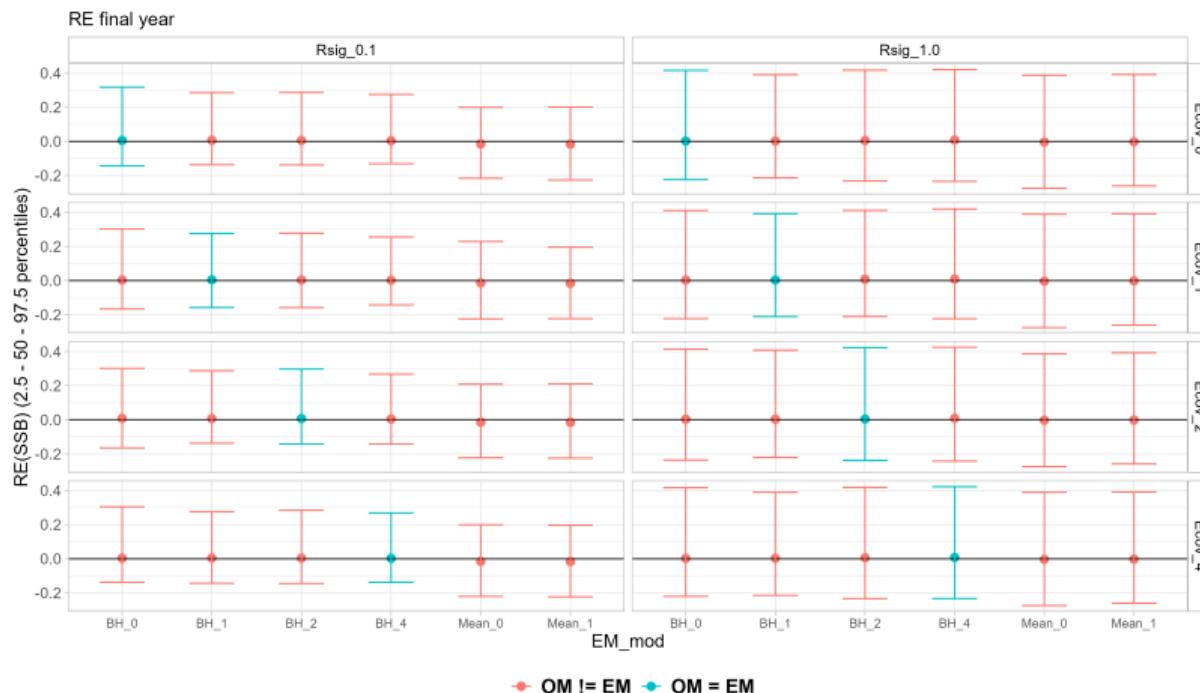
Assessment error - Spawning Stock Biomass (all years)

- Median unbiased, σ_R determined quantile range



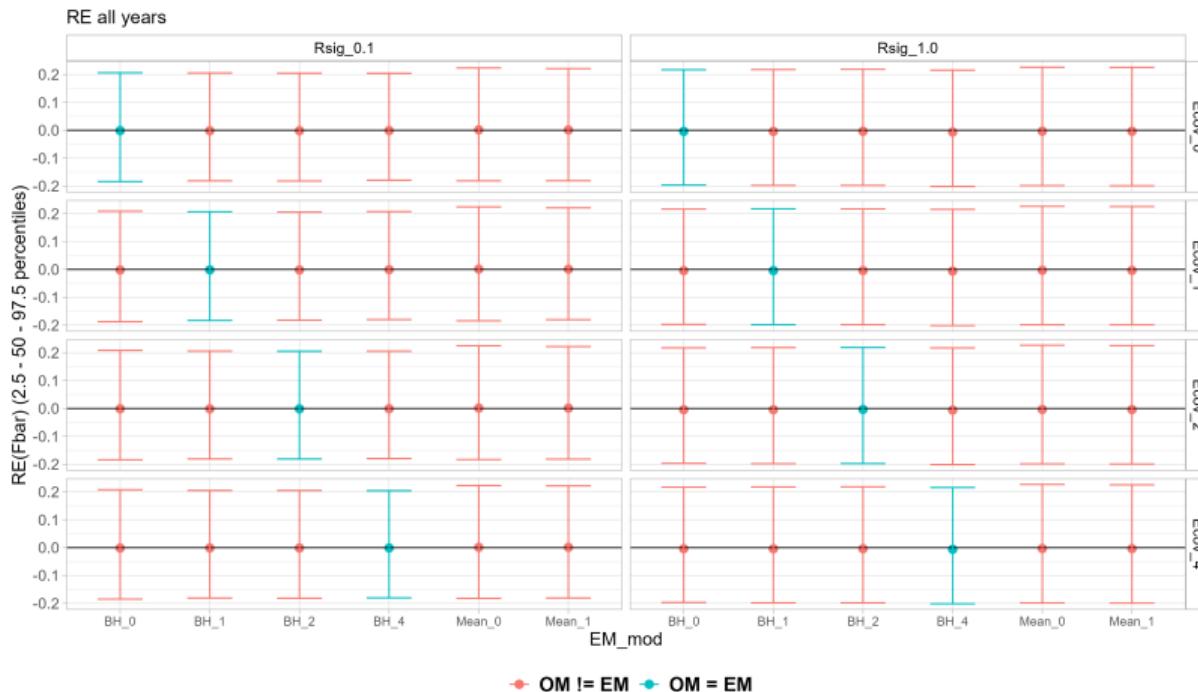
Assessment error - Spawning Stock Biomass (terminal year)

- Median unbiased, σ_R determined quantile range



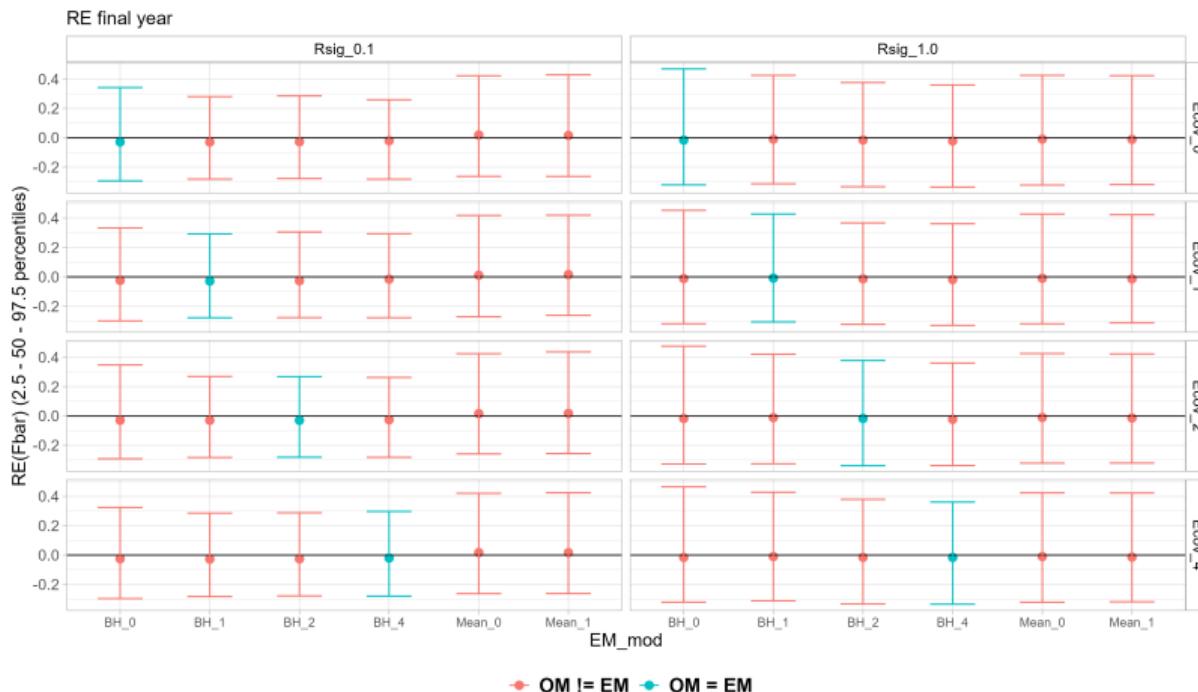
Assessment error - Fbar (all years)

- Median unbiased, σ_R determined quantile range



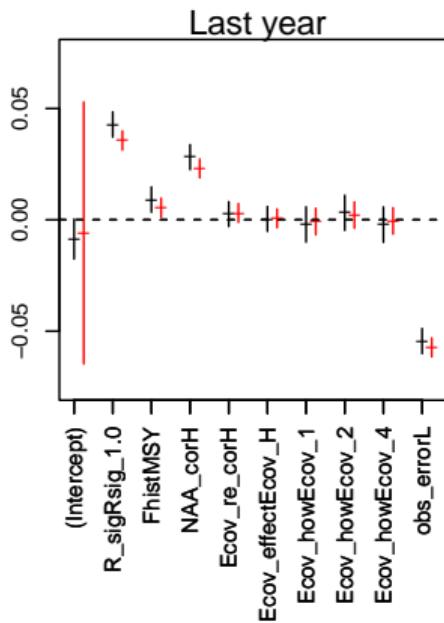
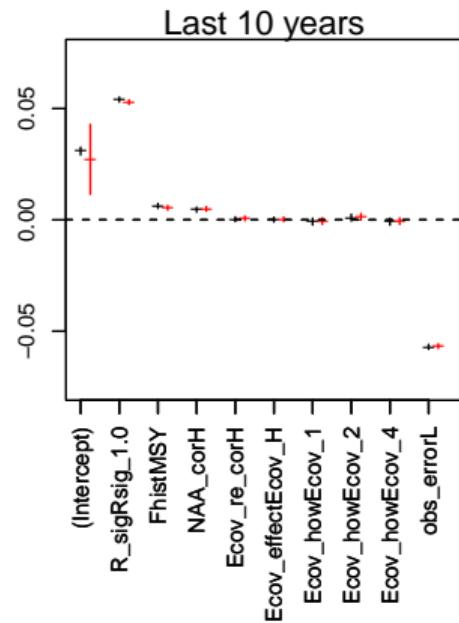
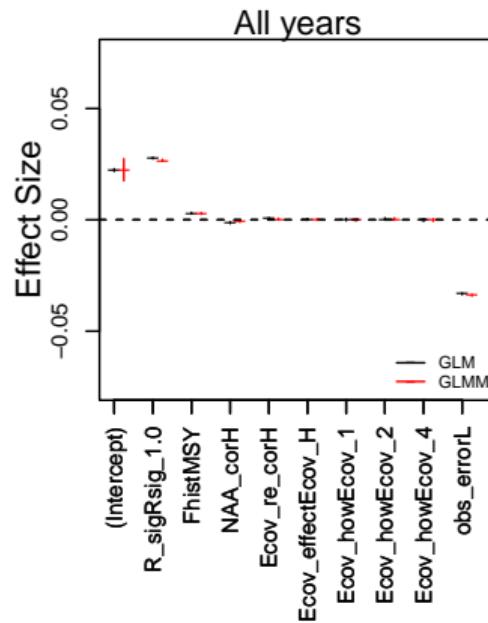
Assessment error - Fbar (terminal year)

- Median unbiased, σ_R determined quantile range



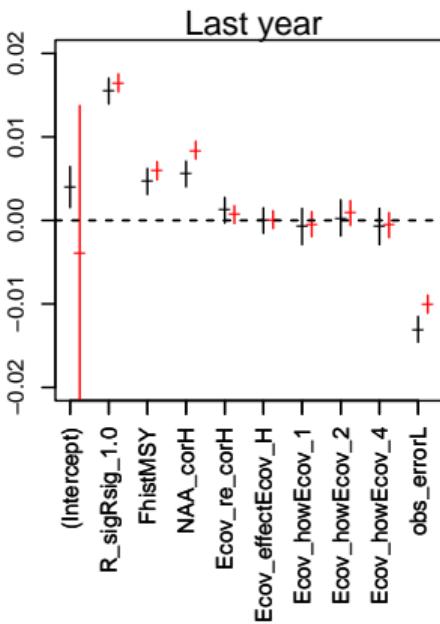
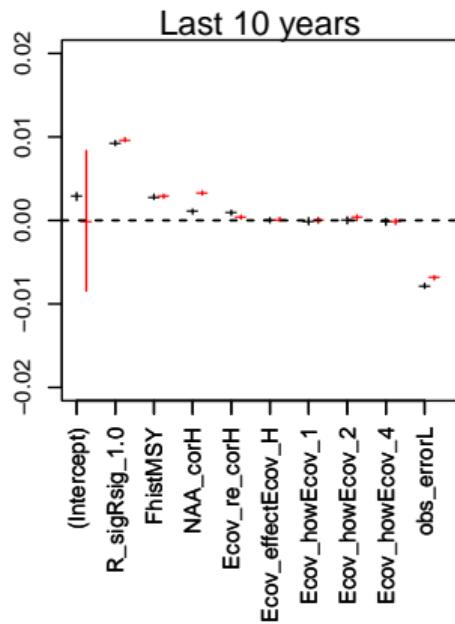
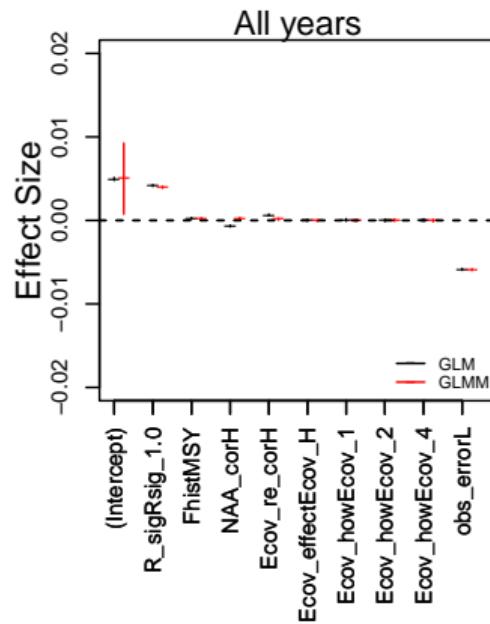
Background
ooooInputs
oooOM
oooooo
EM
ooAnalysis & Results
ooooooooooooooooooooAlert
ooResults
ooooooooooooooooooooConclusions
ooFuture Work
ooo

Effect size for Relative error in Recruitment



Background
ooooInputs
oooOM
oooooo
EM
ooAnalysis & Results
oooooooooooo
ooooooooooooAlert
ooResults
oooooooooooo
ooooooooooooConclusions
ooFuture Work
ooo

Effect size for Relative error in SSB



Effect size for Relative error in Fbar

Background
oooo

Inputs
ooo

OM
oooooo EM
oo

Analysis & Results
oooooooooooooooooooo

Alert
●○

Results
oooooooooooooooooooo

Conclusions
oo

Future Work
ooo

Alert

Background
oooo

Inputs
ooo

OM
oooooo EM
oo

Analysis & Results
oooooooooooooooooooo

Alert
o●

Results
oooooooooooooooooooo

Conclusions
oo

Future Work
ooo

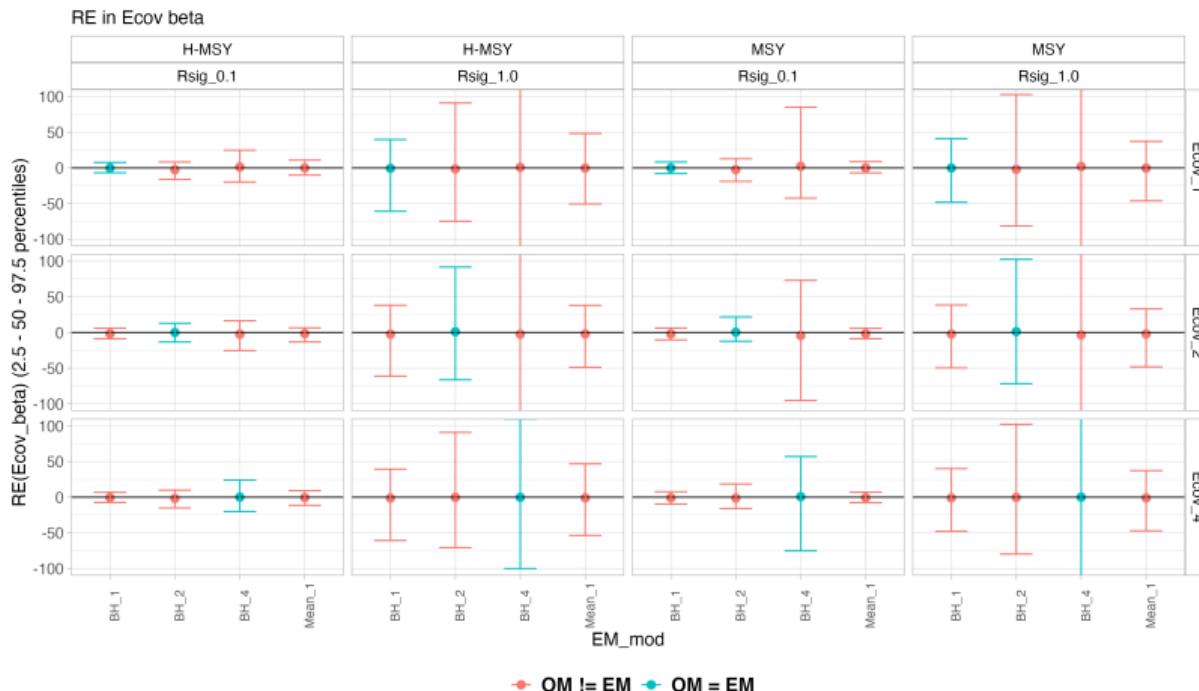
Is everybody still awake?



Results

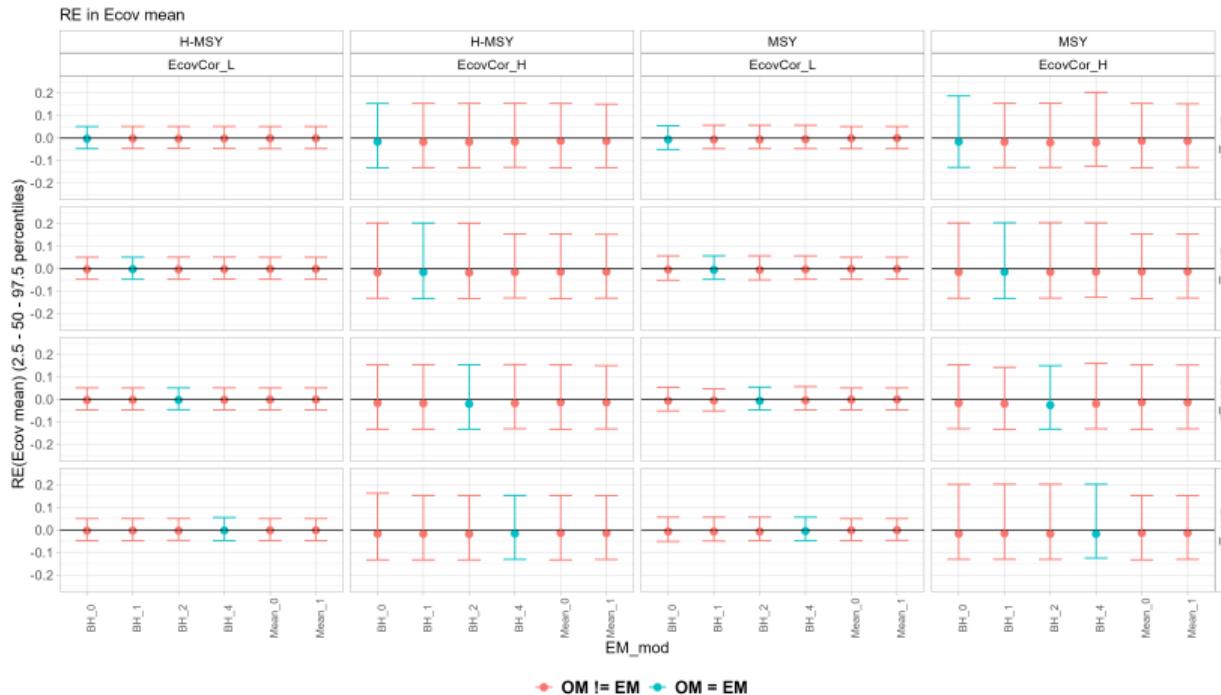
Bias of estimated parameters - Ecov β

- Limiting and Masking had worst performance (BH_3 and BH_4)



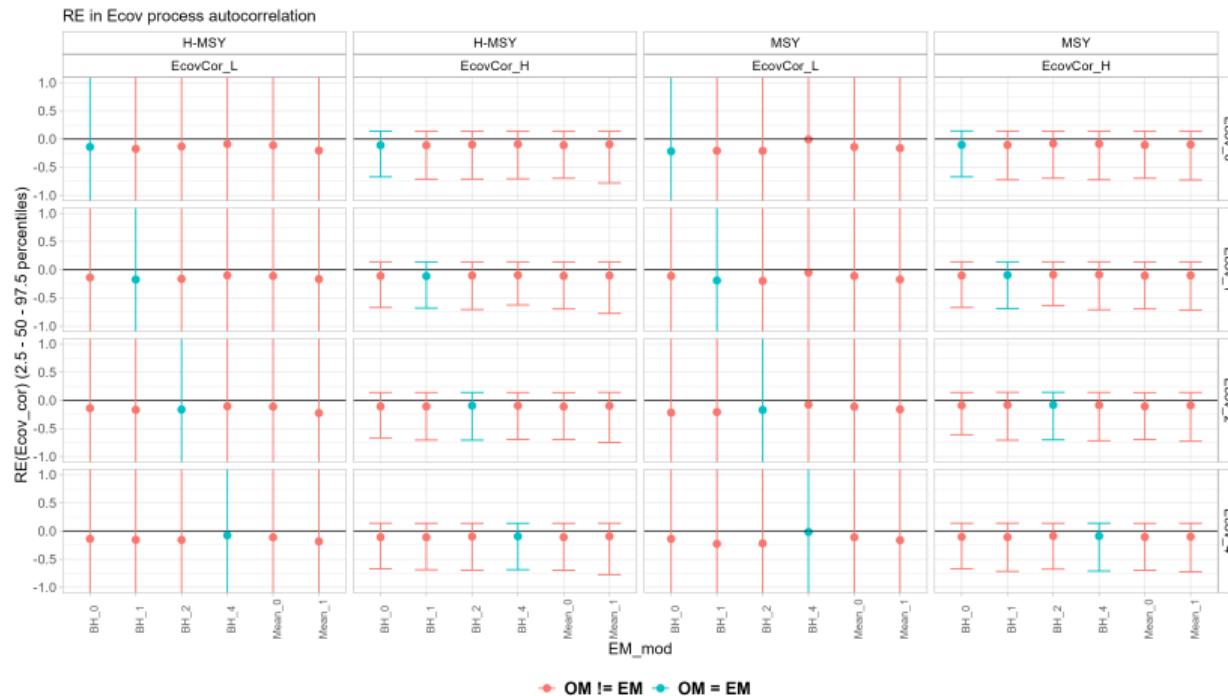
Bias of estimated parameters - Ecov mean

- Larger quantile range when Ecov correlation is high (0.8)



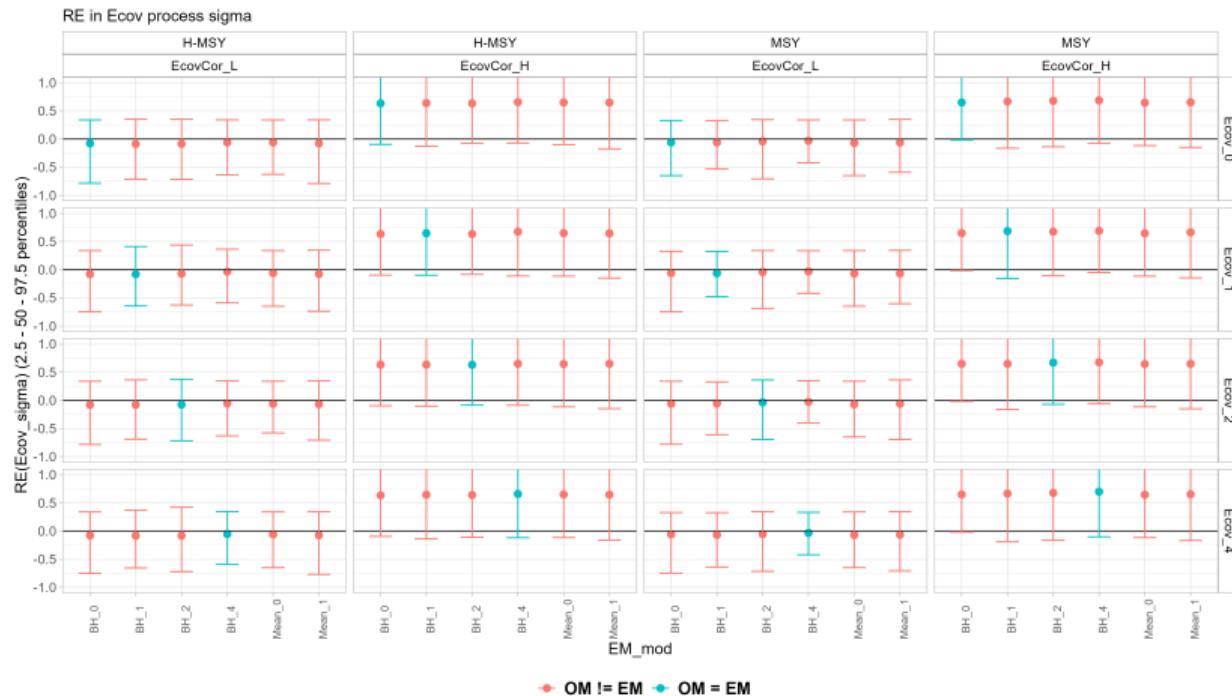
Bias of estimated parameters - ρ_{Ecov}

- Well estimated when Ecov correlation was high (0.8)



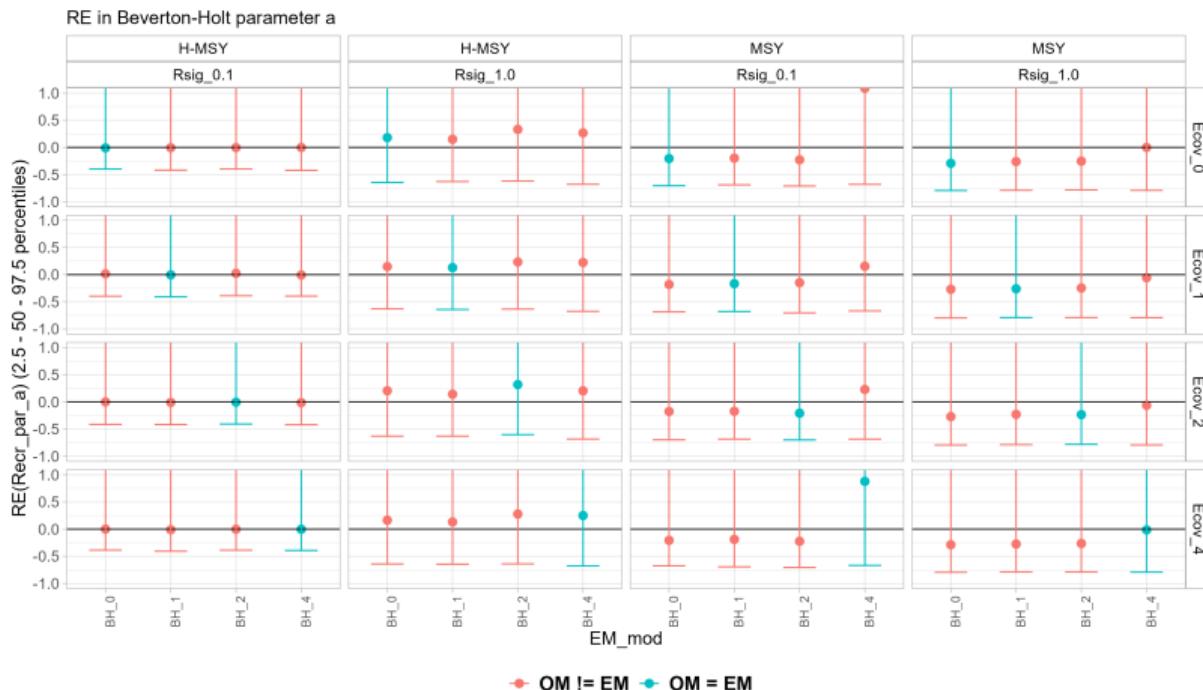
Bias of estimated parameters - σ_{Ecov}

- Positively biased when Ecov correlation is high (0.8)



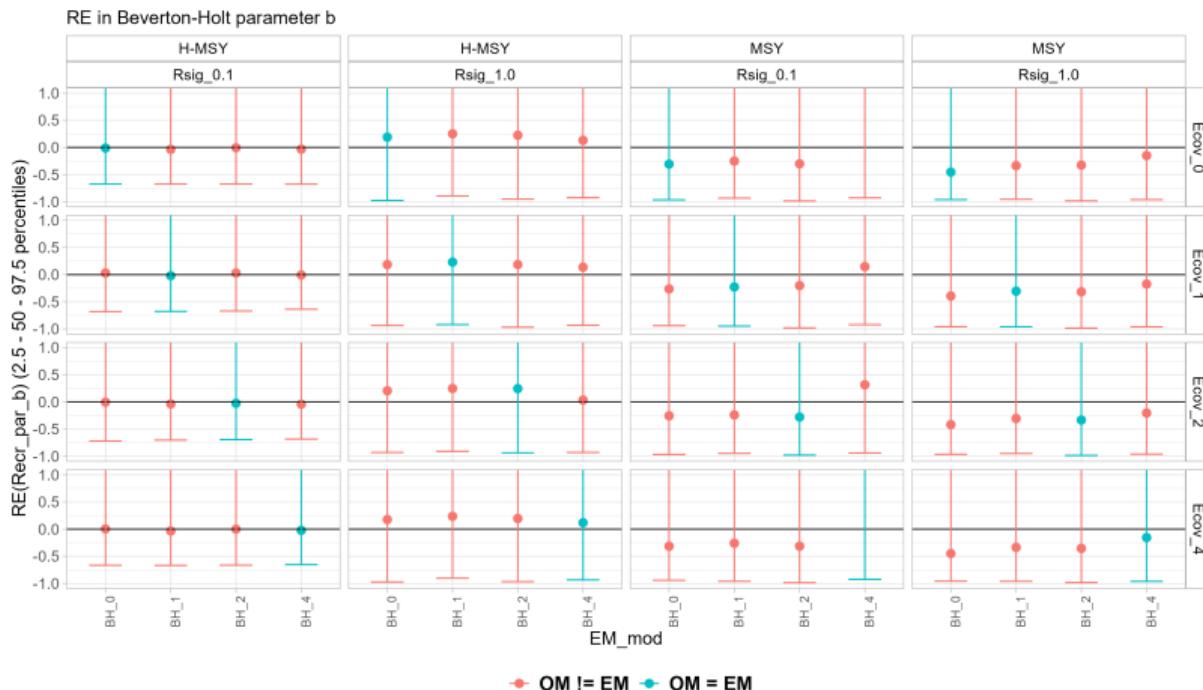
Bias of estimated parameters - SRR a

- Unbiased when $\sigma_R = 0.1$ and high fishing contrast; otherwise biased +/- 0.2 (BH_4 could be much worse)



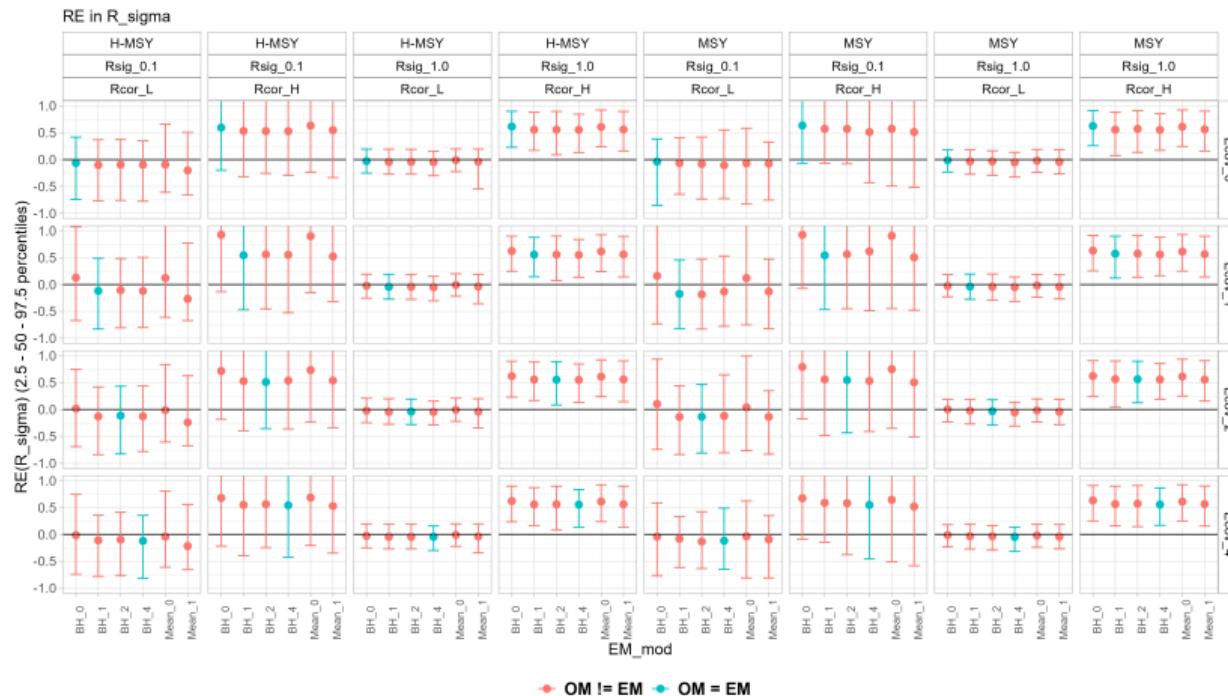
Bias of estimated parameters - SRR b

- Unbiased when $\sigma_R = 0.1$ and high fishing contrast; otherwise biased +/- 0.2 (BH_4 could be much worse)



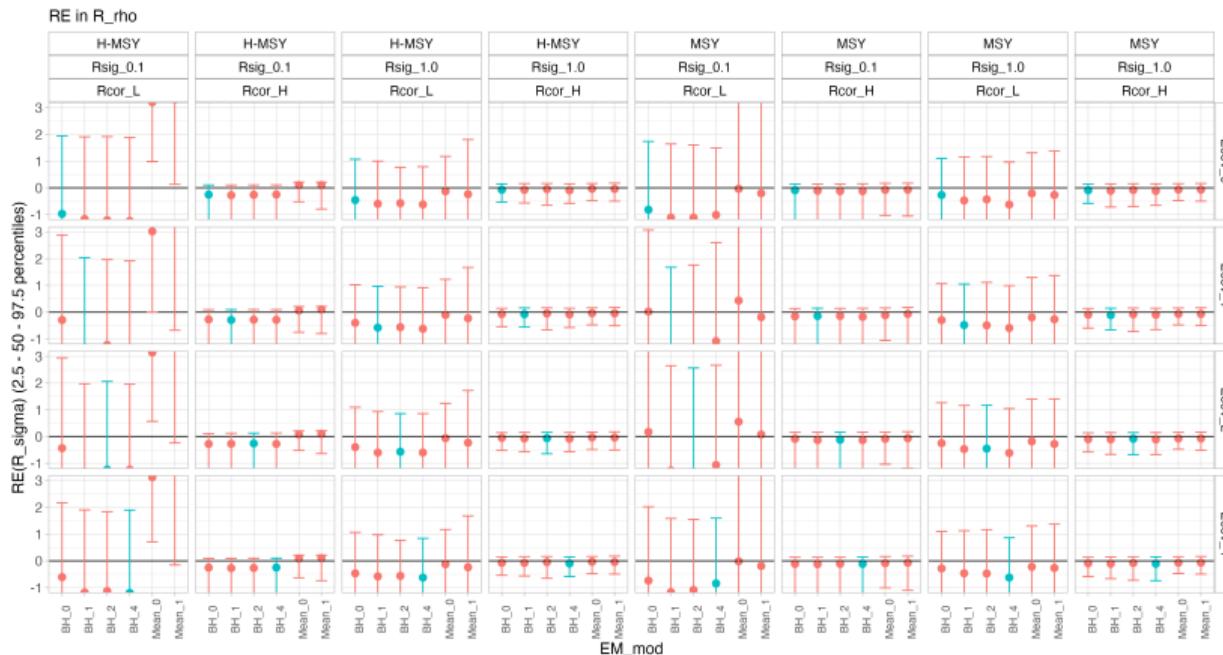
Bias of estimated parameters - σ_R

- Bias pattern mirroring between σ_R and correlation



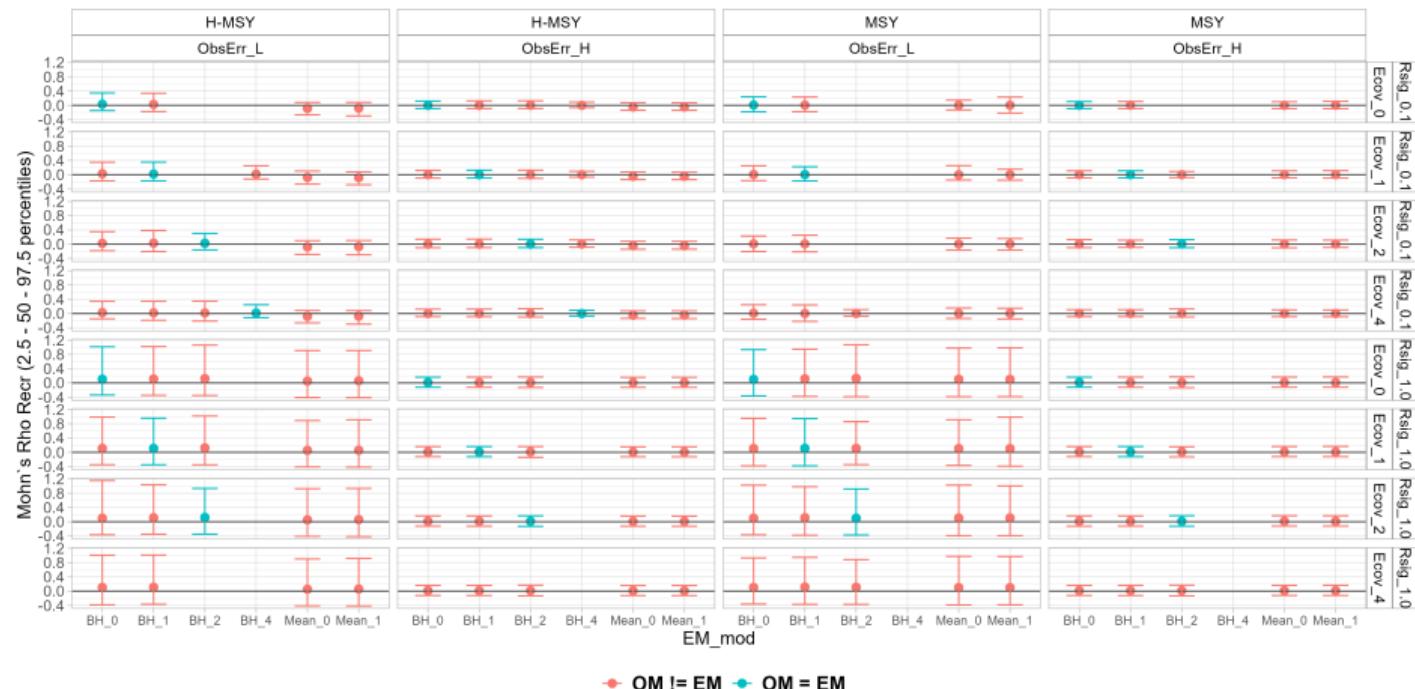
Bias of estimated parameters - ρ_R

- Bias pattern mirroring between σ_R and correlation
- For most informative case (H-MSY, $\sigma_R=0.1$, low recruitment correlation), mean SRR has large positive bias



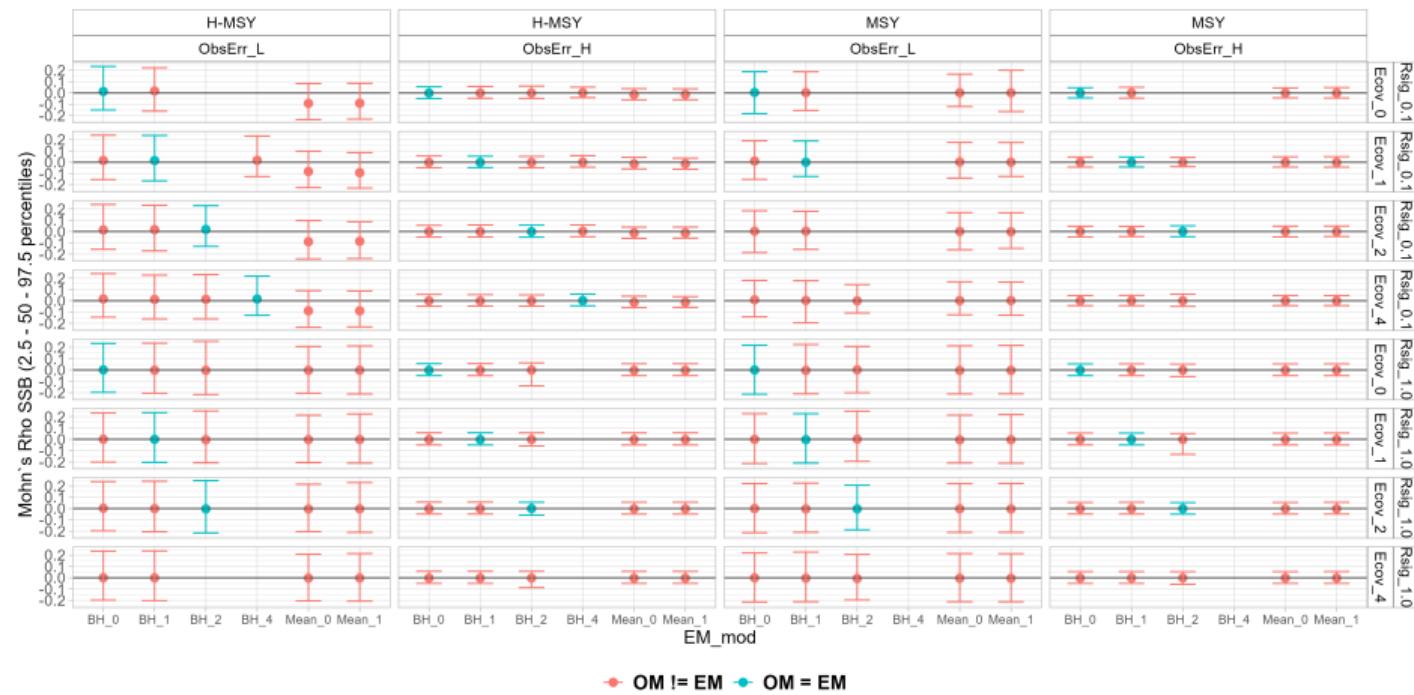
Mohn's ρ - Recruitment (7 year peel)

- Missing symbol indicates failed peels for one or more EM fits
- σ_R and observation error were influential factors



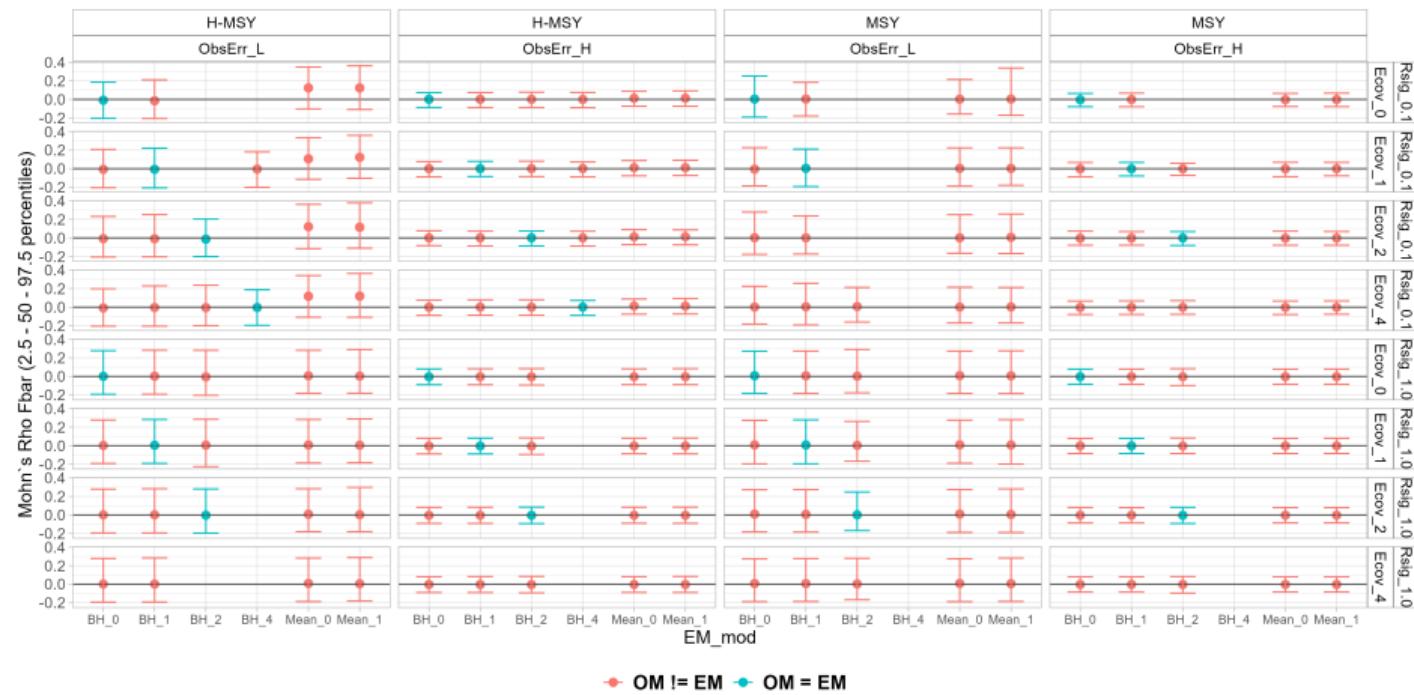
Mohn's ρ - SSB (7 year peel)

- observation error was influential factor



Mohn's ρ - Fbar (7 year peel)

- observation error was influential factor



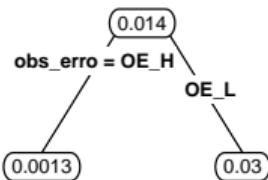
Background
ooooInputs
oooOM
oooooo EM
ooAnalysis & Results
ooooooooooooooooooooAlert
ooResults
oooooooooooo●ooooooooooooConclusions
ooFuture Work
ooo

Mohn's ρ - Classification Tree for Recruitment, SSB, Fbar

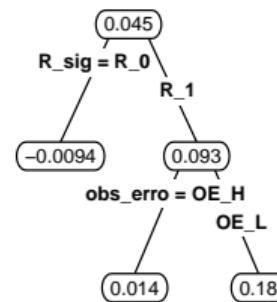
a) SSB

(-0.0056)

b) Fbar



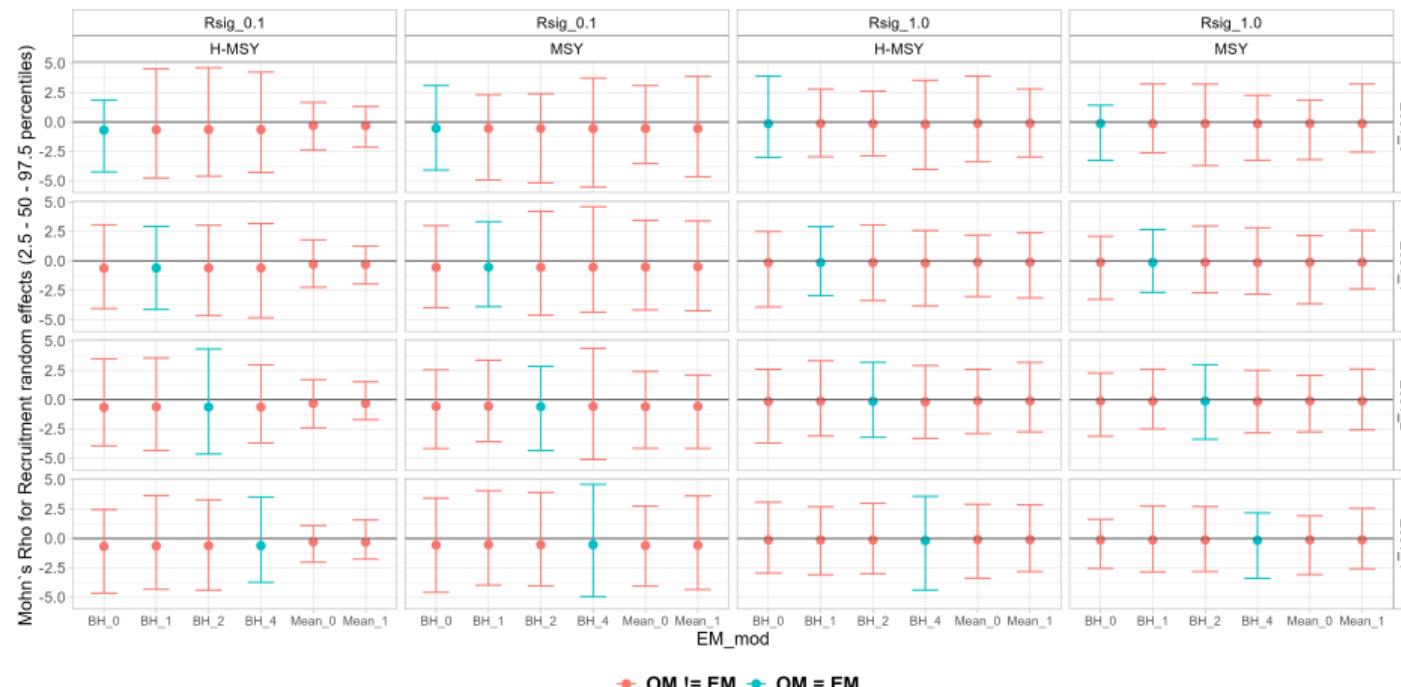
c) Recr



Mohn's rho

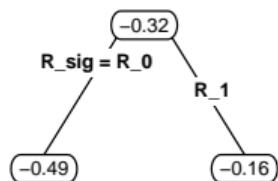
Mohn's ρ - Random Effects on Recruitment (7 year peel)

- Negative bias when σ_R is low
 - Expected worse ρ from incorrect EM, but was not the case



Mohn's ρ - Classification Tree for Random Effects on Recruitment

R_dev



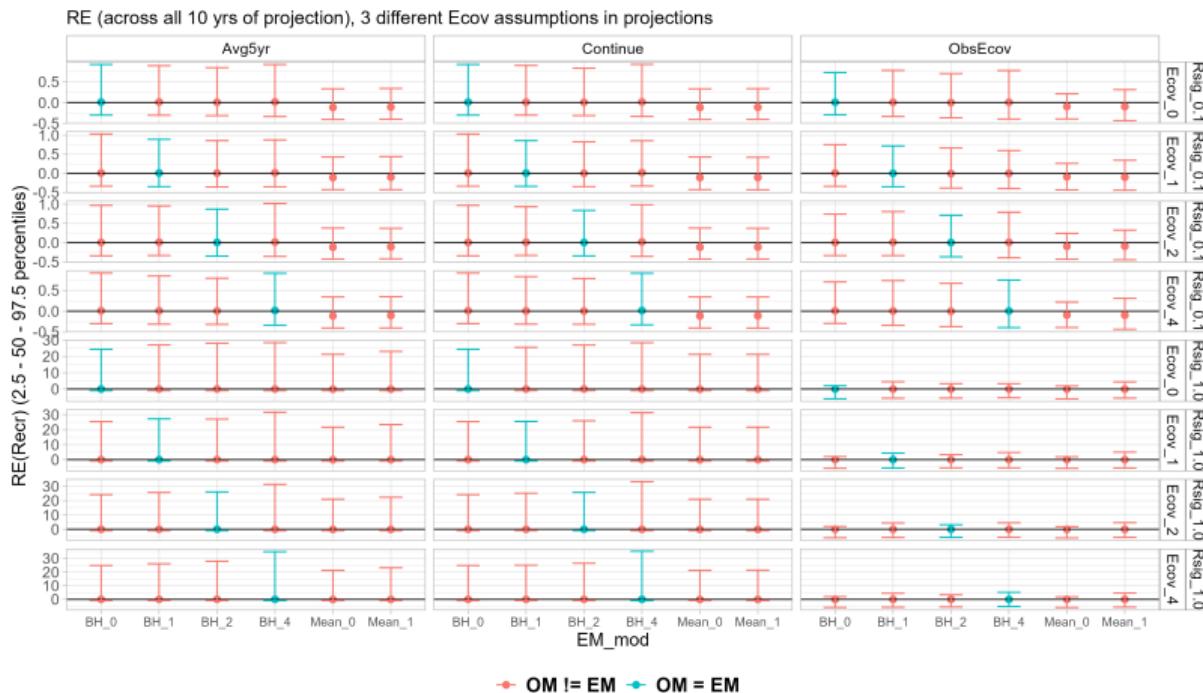
Mohn's rho on recruitment random effects

Projections: Specifications and Ecov assumptions

- Projections were made from the 10th peel of each fitted model (Terminal year - 10)
- A 5-year average of maturity, weight at age, and selectivity specified (time-invariant anyway)
- F in the projection was set at the true OM F value
- The environmental index was treated 3 ways
 1. A 5-year average of the most recent index value was held constant in the 10 year projection ('Avg5yr')
 2. The estimated process continued in the 10 year projection ('Continue')
 3. The observed environmental index was specified in the 10 year projection ('ObsEcov')
- Projected Recruitment, SSB, and Catch were compared to the true OM values in the projection

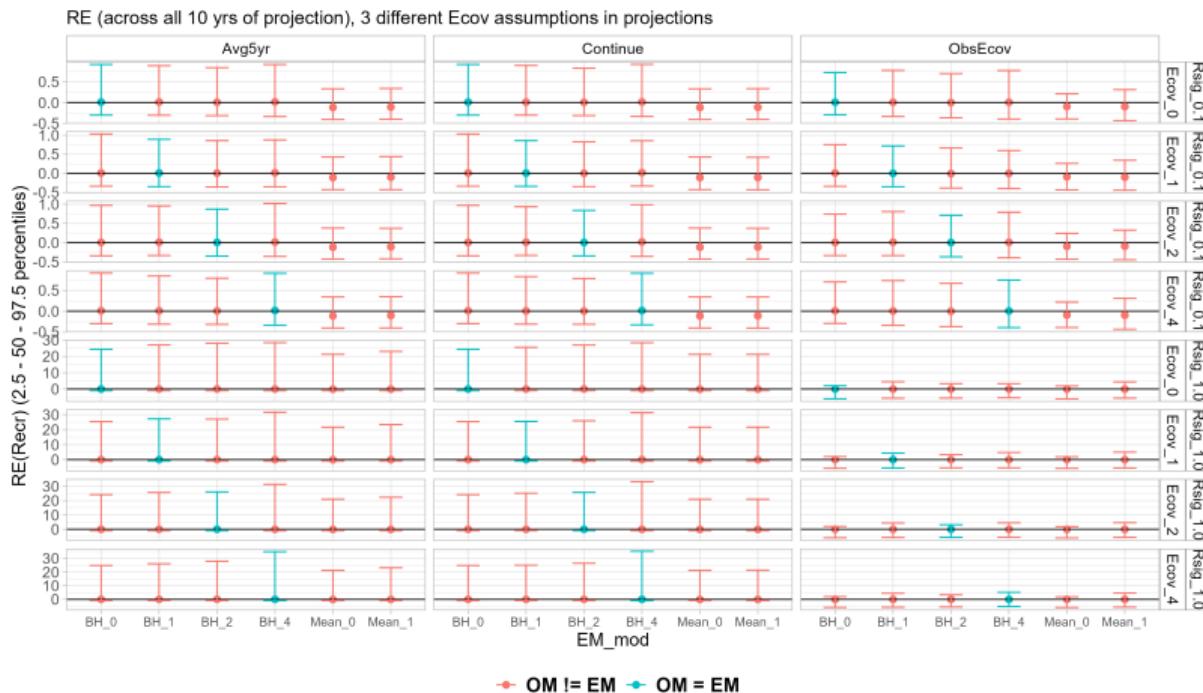
Projection performance - Relative Error Recruitment (first 3 years)

- Minimal bias (median at most -0.1 for mean SRR)
- Quantile range much larger for BH



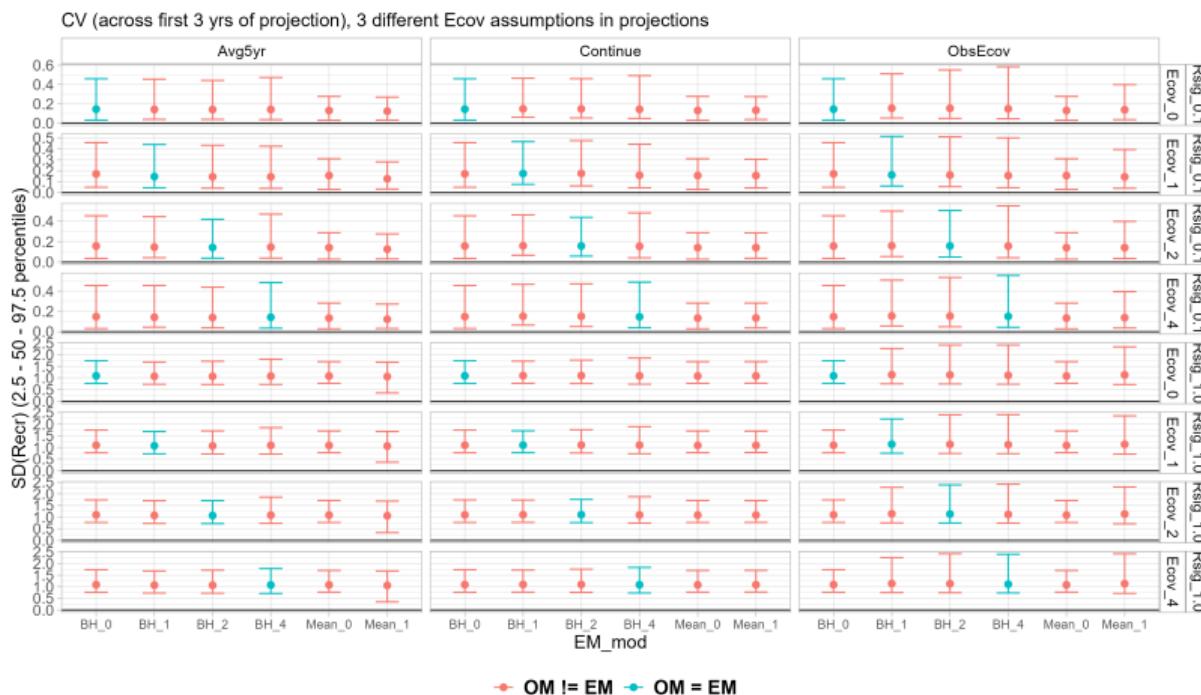
Projection performance - Relative Error Recruitment (all 10 years)

- Minimal bias (median at most -0.1 for mean SRR)
- Quantile range much larger for BH



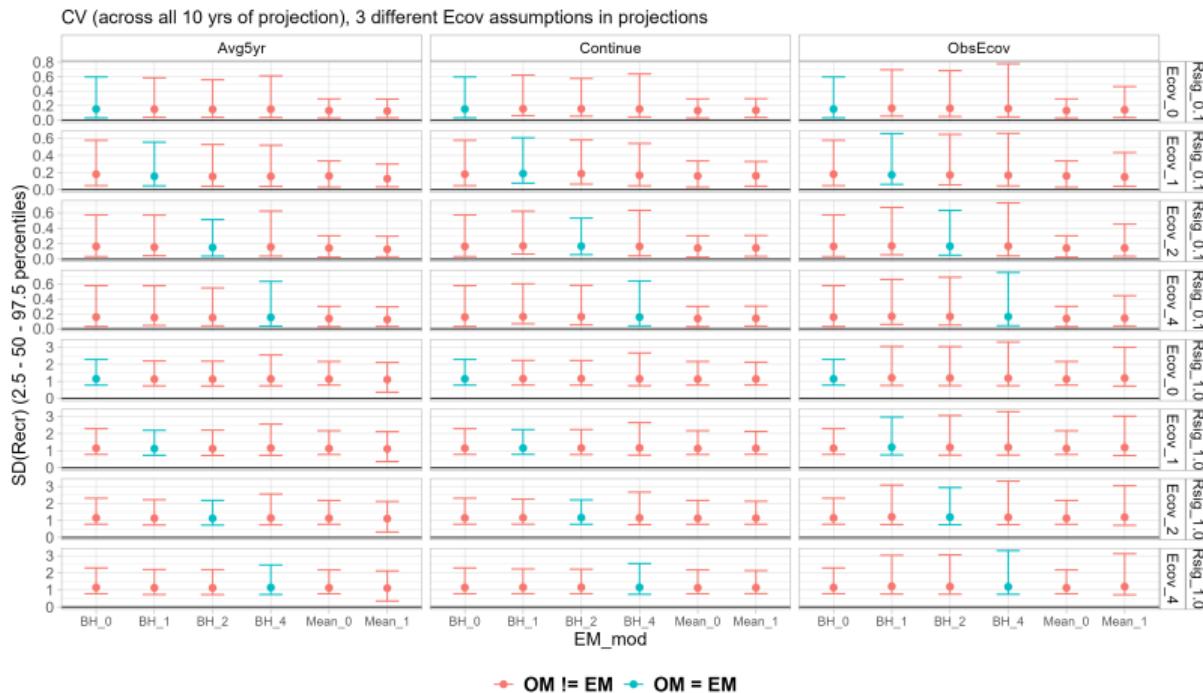
Projection performance - Recruitment Uncertainty (first 3 years)

- CV slightly lower for mean SRR, usually narrower quantile range



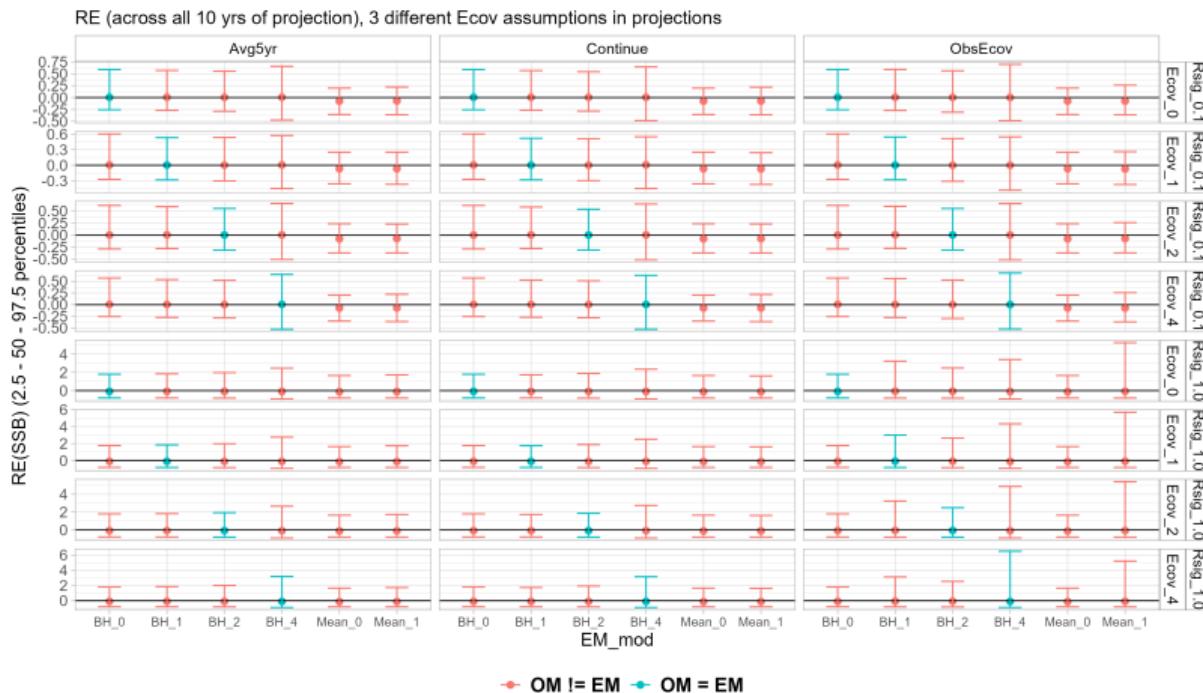
Projection performance - Recruitment Uncertainty (all 10 years)

- CV slightly lower for mean SRR, usually narrower quantile range



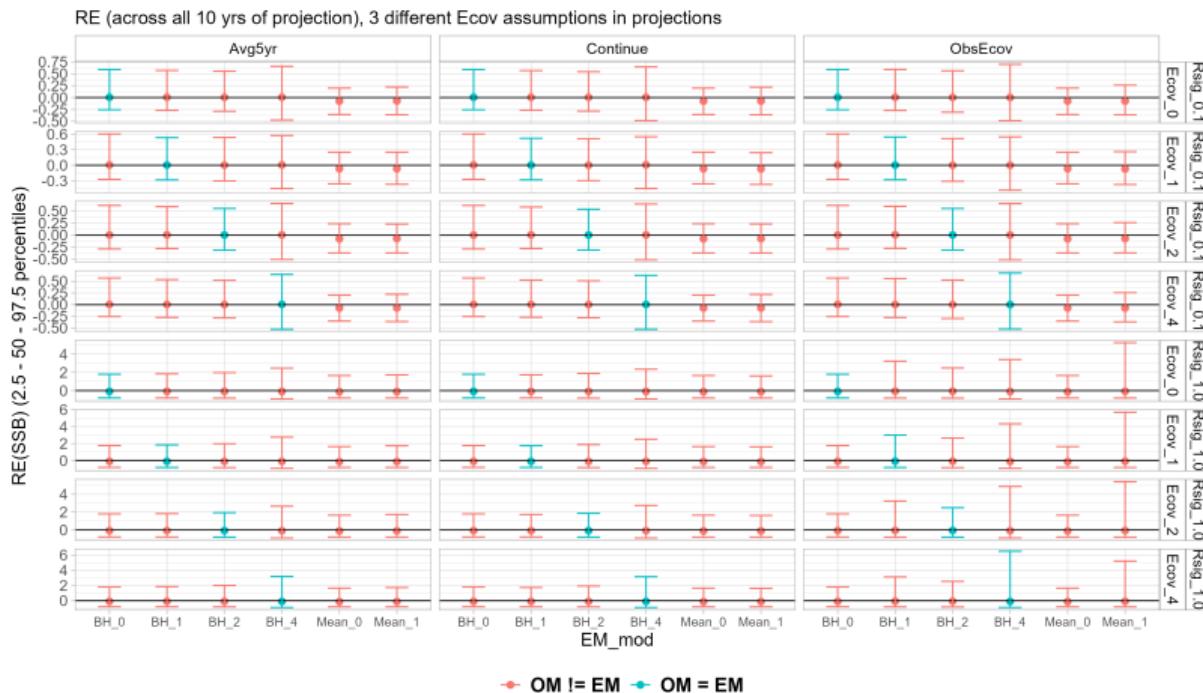
Projection performance - Relative Error SSB (first 3 years)

- Minimal bias (median at most -0.1 for mean SRR)
- Quantile range much larger for BH



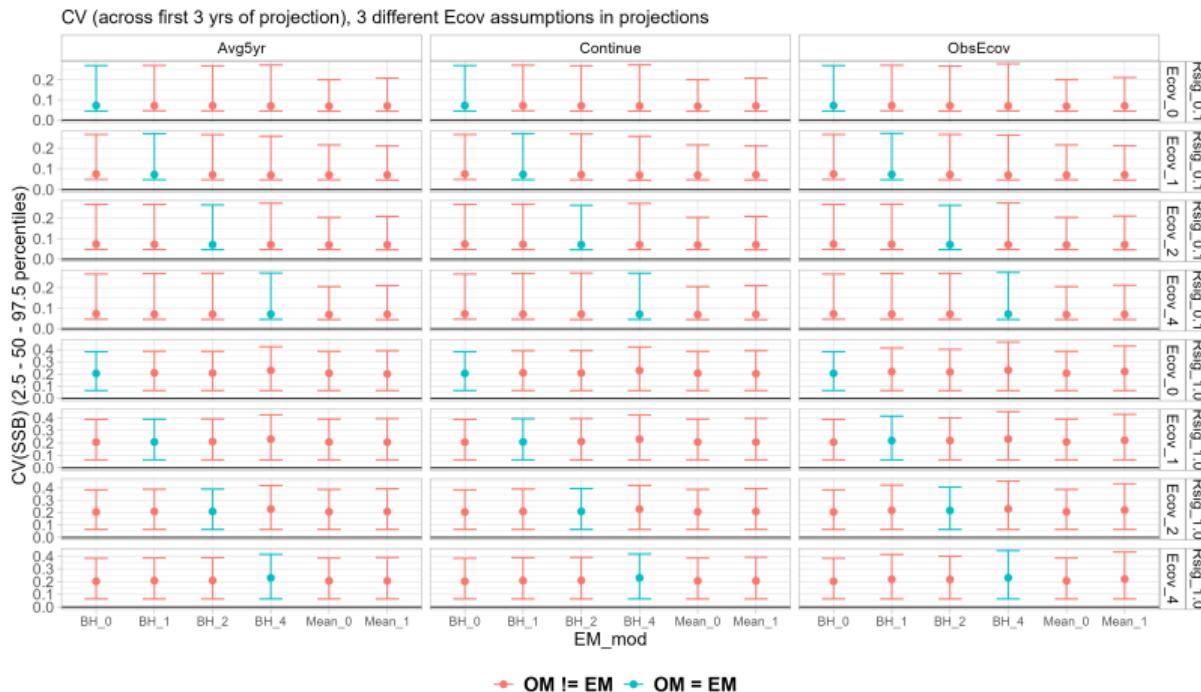
Projection performance - Relative Error SSB (all 10 years)

- Minimal bias (median at most -0.1 for mean SRR)
- Quantile range much larger for BH



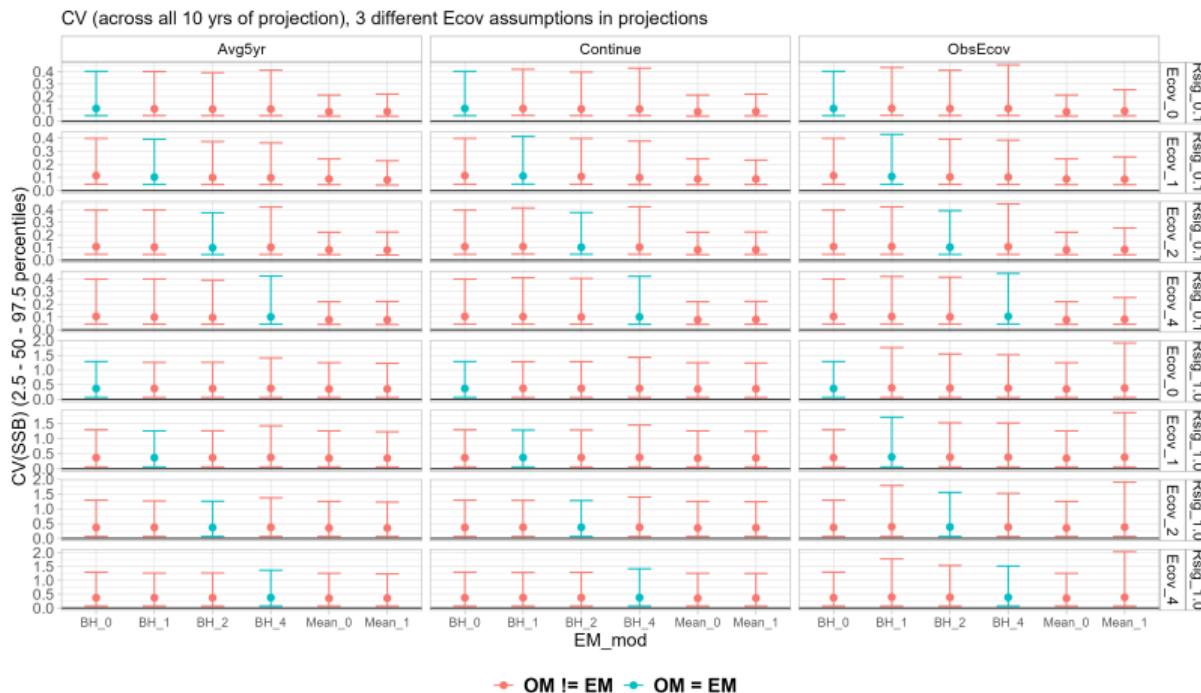
Projection performance - SSB Uncertainty (first 3 years)

- CV slightly lower for mean SRR, usually narrower quantile range



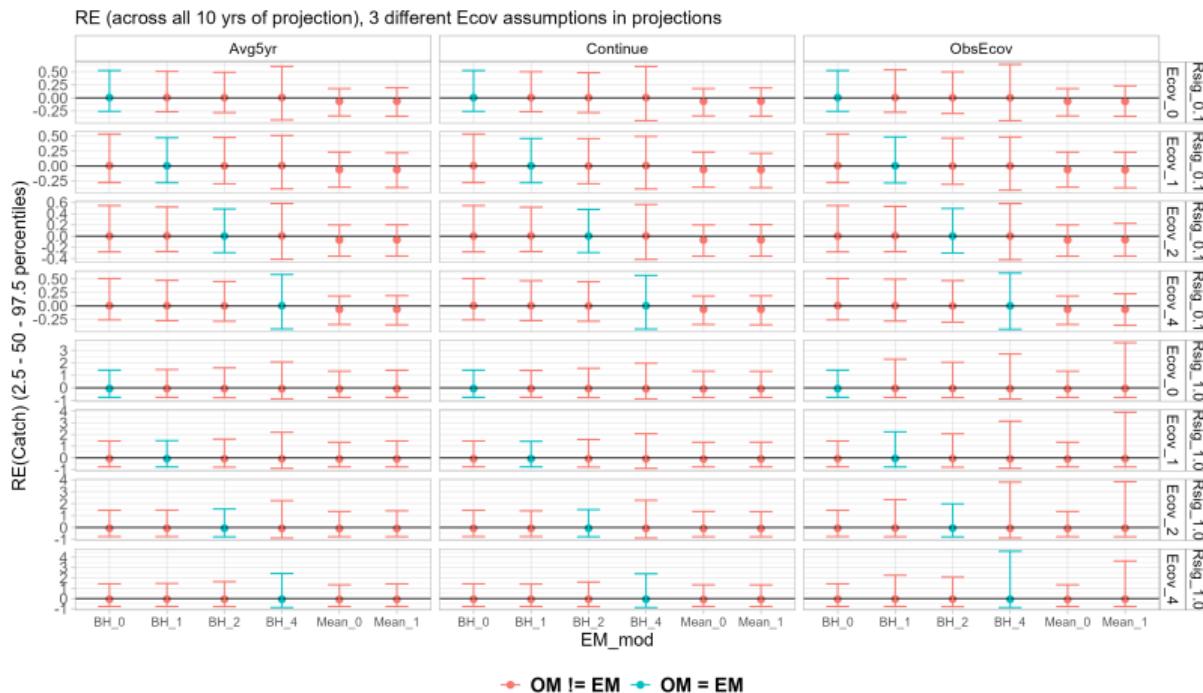
Projection performance - SSB Uncertainty (all 10 years)

- CV slightly lower for mean SRR, usually narrower quantile range



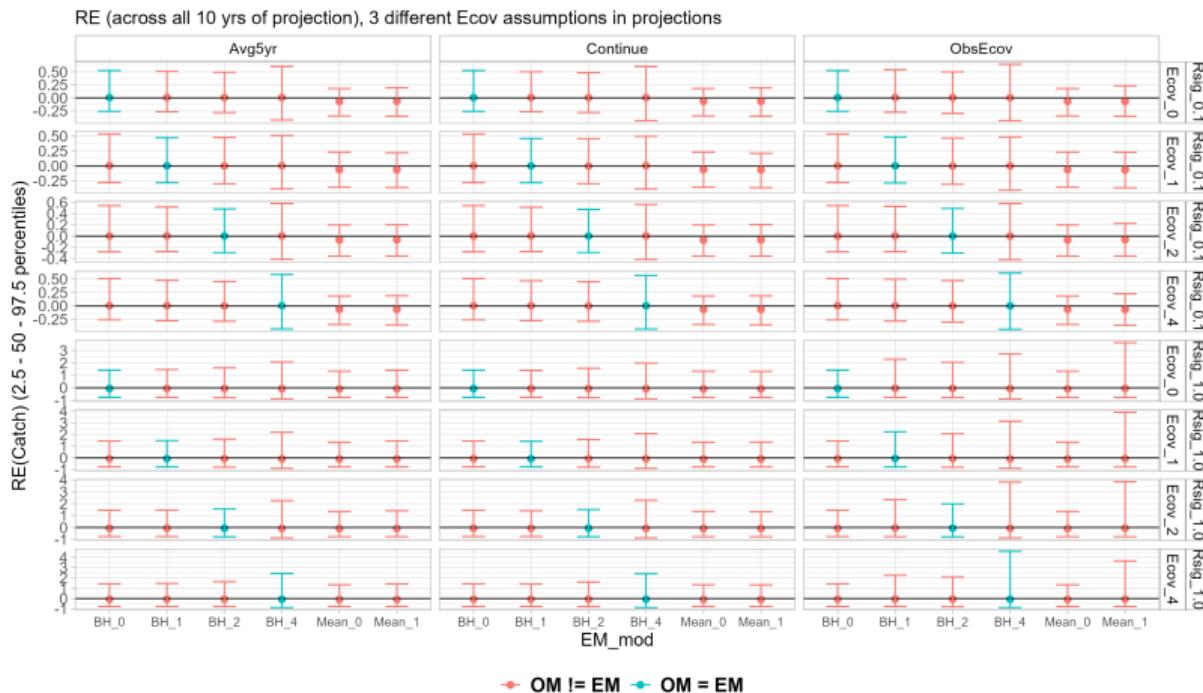
Projection performance - Relative Error Catch (first 3 years)

- Minimal bias (median at most -0.1 for mean SRR)
- Quantile range much larger for BH



Projection performance - Relative Error Catch (all 10 years)

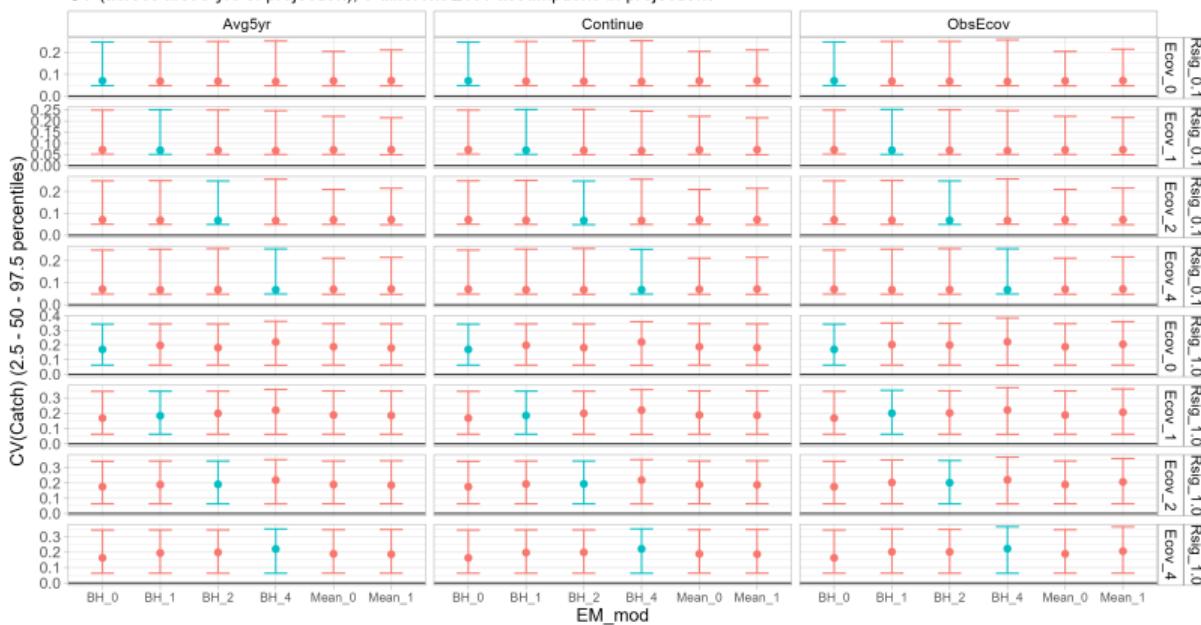
- Minimal bias (median at most -0.1 for mean SRR)
- Quantile range much larger for BH



Projection performance - Catch Uncertainty (first 3 years)

- CV slightly lower for mean SRR, usually narrower quantile range

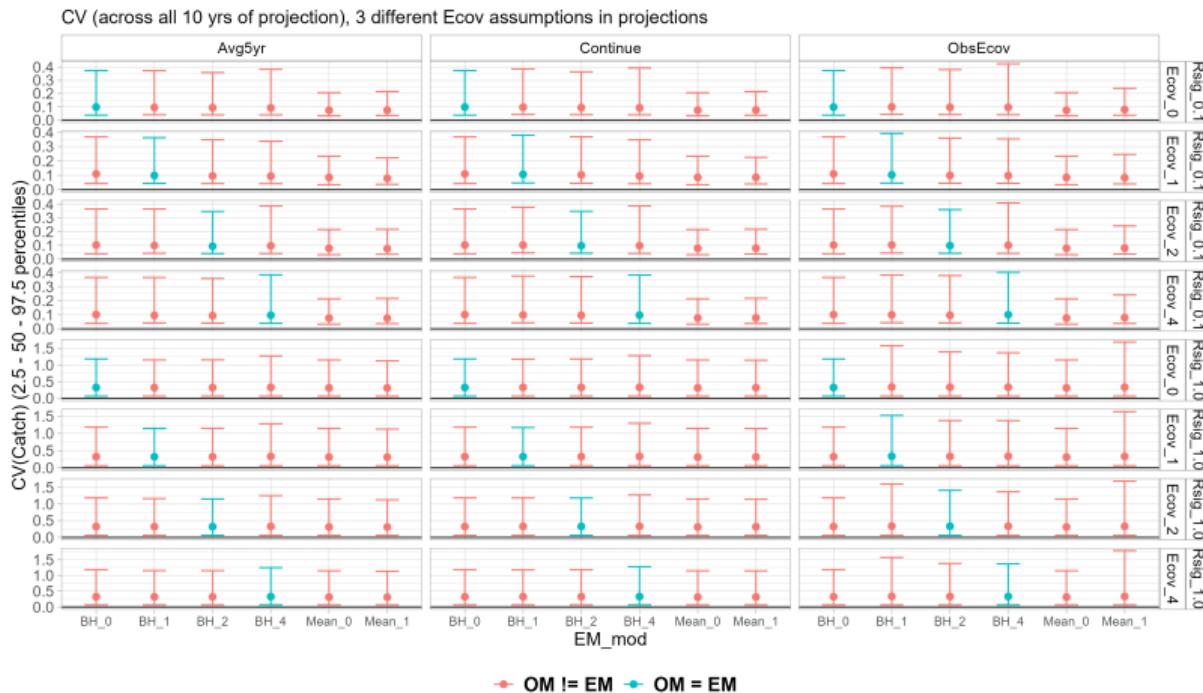
CV (across first 3 yrs of projection), 3 different Ecov assumptions in projections



— OM != EM — OM = EM

Projection performance - Catch Uncertainty (all 10 years)

- CV slightly lower for mean SRR, usually narrower quantile range



Projection Summary

- Minimal difference in median bias
- Recruitment is highly uncertain in projections, exacerbated by length of projection
- SSB and Catch also show increased uncertainty in longer projections
- Long projections are expected to be very imprecise, recommend shorter projections
- Tighter quantile range and slightly lower CV for mean SRR

Background
oooo

Inputs
ooo

OM
oooooo EM
oo

Analysis & Results
oooooooooooooooooooo

Alert
oo

Results
oooooooooooooooooooo

Conclusions
●○

Future Work
ooo

Conclusions

Background
oooo

Inputs
ooo

OM
oooooo EM
oo

Analysis & Results
oooooooooooooooooooooooooooo

Alert
oo

Results
oooooooooooooooooooooooooooo

Conclusions
oo

Future Work
ooo

Take-aways

- Stock-recruitment model identification was poor, requiring unrealistic OM scenarios
- EMs with mean SRR and recruitment random effects generally performed similarly well as OM=EM, including for cases with an SR relationship and/or an ecov effect
- In conclusion, we recommend recruitment random effects from a mean SRR as the default EM in WHAM

Background
oooo

Inputs
ooo

OM
oooooo EM
oo

Analysis & Results
oooooooooooooooooooo

Alert
oo

Results
oooooooooooooooooooo

Conclusions
oo

Future Work
●oo

Future Work

Background
oooo

Inputs
ooo

OM
oooooo EM
oo

Analysis & Results
oooooooooooooooooooo

Alert
oo

Results
oooooooooooooooooooo

Conclusions
oo

Future Work
o●o

Recommendations for further analyses

- In WP-1, we recommended exploring additional levels of $\sigma_R=0.5$; however, we note that WP1-Appendix looked at $\sigma_R=0.5$ and found no difference

Acknowledgements

- This work could not have been completed without support from the Microsoft Cooperative Research and Development Agreement (CRADA) and NOAA's National Cloud Program Office (OCIO), as well as MIT (... greg to fill in)
- We thank other members of the SSRTWG for thoughtful comments during earlier discussions and presentations of this work

