Factors affecting inferences on natural mortality and associated environmental effects in state-space age-structured assessment models

Timothy J. Miller[1,2] Gregory L. Britten[3] Elizabeth N. Brooks[2]

Gavin Fay[4] Alexander C. Hansell[2] Christopher M. Legault[2]

Brandon Muffley[5] John Wiedenmann[6]

12 February, 2026

[1]corresponding author: timothy.j.miller@noaa.gov

[2]Northeast Fisheries Science Center, Woods Hole Laboratory, 166 Water Street, Woods Hole, MA 02543 USA

[3]Biology Department, Woods Hole Oceanographic Institution, 266 Woods Hole Rd. Woods Hole, MA, USA

[4]Department of Fisheries Oceanography, School for Marine Science and Technology, University of Massachusetts Dartmouth, 836 S Rodney French Boulevard, New Bedford, MA 02744, USA

[5]Mid-Atlantic Fishery Management Council, 800 North State Street, Suite 201, Dover, DE 19901 USA

[6]Department of Ecology, Evolution, and Natural Resources. Rutgers University

Abstract

Treatment of natural mortality is a major consideration in assessment models and state-space approaches allow estimation of temporal variation in this mortality rate as well as effects of specific covariates. However, there has been no investigation of the reliability of inferences made regarding natural mortality, associated covariate effects, and important assessment output from state-space assessment models. We conducted a large-scale simulation study that considers models fit to data simulated from operating models with alternative assumptions defined by several factors, but we focus on scenarios where there is temporal contrast in fishing pressure and lower uncertainty in population observations (age composition and indices of abundance). We fit estimating models to simulated observations with alternative assumptions on inclusion of environmental effects, estimation of the median natural mortality rate, and the source of temporal variability in the population demography. Our results suggest that estimation of environmental effects on natural mortality is possible and reliable even when the population process error source was misspecified in some scenarios with lower uncertainty in covariate observations, and higher covariate temporal variability.

keywords: state-space assessment models, time-varying natural mortality, bias, AIC

Introduction

State-space population models are now used widely for fisheries stock assessment in Europe, the United States, and Canada (Nielsen and Berg, 2014; Cadigan, 2016; Pedersen and Berg, 2017; Stock and Miller, 2021). Because application of these methods are considered best practice and recommended for the next generation of stock assessment models (Hoyle et al., 2022; Punt, 2023), it is expected their use will only grow globally. An appeal of state-space models lies in their separation of sources of biological and measurement variability by treating latent population characteristics as statistical time series with periodic observations measured with error. Through advances in computational capacity, we can use sophisticated numerical approaches to estimate model parameters as mixed effects (Thorson and Minto, 2015; Kristensen et al., 2016).

State-space stock assessment models, with non-linear functions of latent processes and numerous observation types with different probability distribution assumptions represent one of most complex classes of state-space models. The literature on the effects of various factors on reliability of inferences from state-space assessment models is growing (Li et al., 2024; Miller et al., In review). The importance of contrast in population size and fishing mortality (F) and quality of data used to fit assessment models including the state-space variety is known (Magnusson and Hilborn, 2007; Miller et al., In review). Furthermore, estimation of natural mortality (M), and even temporal variability in M is possible in many scenarios (Lee et al., 2011; Cadigan, 2016; Miller and Hyun, 2018; Miller et al., In review).

The effects of temporal variation in recruitment via undefined or explicit environmental factors have been extensively investigated in both traditional assessment models and state space models (Myers, 1998; Haltuch and Punt, 2011; Johnson et al., 2016; Miller et al., 2016). Reliability of estimating environmental and spawning biomass effects on recruitment in state-space assessment models requires a combination of strong effects, good age composition data quality, contrast in the environmental covariate and lower recruitment variability (Britten et al., 2026; Miller et al., In review).

A critical aspect of fisheries assessment models and their use in management is short-term projections that are used to determine catch advice. While understanding drivers of recruitment is important particularly for subsequent effects on reference points, recruitment in short-term projections typically has little impact on the exploitable biomass in the first few projection years. However, assumptions for M have immediate and larger effects on projected biomass because they affect the abundances at older age classes at the end of the data time series that constitute spawning biomass and catch (Brodziak et al., 2008; Stock et al., 2021).

Because of the effects of M on both biological reference points and short term projections, better understanding sources of variation in M would provide more accurate estimation of abundance and productivity and therefore improved management. Temporal variation in M is less studied than recruitment, but its importance for explaining variability in observations has been demonstrated in state-space assessment models for Atlantic cod and yellowtail flounder (Cadigan, 2016; Stock et al., 2021). Deriso et al. (2008) also demonstrated the importance of several factors affecting M for Pacific herring.

Assessment models could include temporal variation in many aspects of population dynamics or how observations are related to the population. For example the Woods Hole Assessment Model (WHAM) can include process errors treated as random effects for transition in cohorts over time (hereafter referred to as apparent survival), catchability for indices of abundance, selectivity of fishing fleets or indices, movement between regions, or in M (Stock and Miller, 2021; Miller et al., 2025). However, misspecified temporal population process errors could lead to biased population and stock status estimation, and, therefore, poor fisheries management decisions (Legault and Palmer, 2016; Szuwalski et al., 2018). Studies of the reliability of inferences regarding the presence of temporal variability in M are limited. Miller et al. (In review) found AIC could accurately distinguish process errors in apparent survival, but not for those specifically due to M except when uncertainty in population observations (indices, catch and age composition) was low and there was greater temporal variation in M. In their simulation studies looking at models with multiple sources of process error, Li et al. (2024) found including more sources of process error than existed in the operating model was a better model-building approach than excluding them a priori.

Here we conduct a simulation study with operating models (OMs) varying by degree of observation error uncertainty, sources of process error, fishing history, temporal variation in environmental covariates, and magnitude of the effect of the covariate on M. The simulated observations from these OMs are fitted with estimating models that make alternative assumptions for sources of process error, and whether median M and covariate effects are estimated. We evaluate the effects of these factors on convergence of fitted models, whether Akaike's information criterion (AIC) can determine the correct source of process error and correct assumption about covariate effects on M, estimation bias for median M and covariate effects, and the accuracy for estimators of important terminal year M and spawning stock biomass (SSB).

Methods

Our analyses used the Woods Hole Assessment Model (WHAM) to construct both OMs and estimation models (EMs, Miller and Stock, 2020; Stock and Miller, 2021; Miller et al., 2025). The WHAM package has been used extensively to configure OMs and EMs for several other simulation studies (Stock et al., 2021; Legault et al., 2023; Li et al., 2024; Britten et al., 2026; Li et al., 2026a) and is used to assess many commercially important stocks in the Northeast U.S. (e.g., NEFSC, 2022a,b, 2024). We used version 1.0.6.9000, commit 77bbd94 to generate all results.

We completed a simulation study with 288 operating models. The factors defining the configuration of each operating model, described in detail in subsequent sections, include source of population process error (3 levels), index and catch observation uncertainty (2 levels), environmental covariate uncertainty (2 levels), temporal variation in the latent environmental covariate (4 levels), size of the covariate effect on M (3 levels), and fishing history (2 levels). We simulated 100 data sets for each operating model that included simulations of process errors.

For each simulated data set we fit a set of 12 EMs. The factors that distinguish the estimating models, also described in detail below, include source of population process error type (3 levels) whether median M was estimated or assumed known (2 levels), and whether the environmental covariate effect on M was estimated or not (2 levels).

The sources of population process error that were used in the OMs or assumed in the EMs were on recruitment only (R), recruitment and apparent survival (R+S), or recruitment and M (R+M). We did not use the log-normal bias-correction feature for process errors or observations described by Stock and Miller (2021) for OMs and EMs (Li et al., 2026b). Simulations were all carried out on the University of Massachusetts Green High-Performance Computing Cluster. Code for completing the simulations and summarizing results can be found at https://github.com/timjmiller/SSRTWG/ecov_study/mortality.

Operating models

Environmental covariate

In the WHAM model, environmental covariates are assumed to be described as state-space processes with annual observations of the true latent covariate (Miller et al., 2016; Stock and Miller, 2021). In our simulations, the latent covariate is assumed to be a stationary first order autoregressive (AR1) process

$$X_y|X_{y-1} \sim \mathrm{N}\left(\mu_E\left(1-\rho_E\right)+\rho_E X_{y-1},\left(1-\rho_E^2\right)\sigma_E^2\right)$$

with marginal mean $\mu_E = 0$ and variance σ_E^2. The four configurations of the latent environmental covariate in the operating models assume one of two values for the marginal standard deviation $\sigma_E \in \{0.1, 0.5\}$ and for the autocorrelation parameter $\rho_E \in \{0, 0.5\}$.

The observations of the latent environmental covariate are assumed to be unbiased and Gaussian

$$x_y | X_y \sim \mathrm{N}\left(X_y, \sigma_e^2\right)$$

The standard deviation of the environmental observations in the operating models is one of two values $\sigma_e \in \{0.1, 0.5\}$. Figure S1 provides example simulations of the latent covariate and observations under the alternative configurations.

Population

Many of the characteristics of the population biology and structure including the age classes (10 age classes (ages 1 to 10+)), time span (40 years), maturity (Figure S2, top left), growth (Figure S2, top right), time of spawning (1/4 of the year), and recruitment (Figure S2, bottom right) are identical to Miller et al. (In review). The maturity at age is a logistic function with age at 50% maturity (a_{50}) = 2.89 and slope = 0.88 and weight at age is derived from a von Beralanffy growth function where $t_0 = 0$, $L_\infty = 85$, and $k = 0.3$, and a length-weight relationship

$$W_a = \theta_1 L_a^{\theta_2}$$

where $\theta_1 = e^{-12.1}$ and $\theta_2 = 3.2$.

The general model for M in year y is a log-linear function of both covariate effects X_y and process errors $\varepsilon_{M,y}$ and a parameter β_M that defines median M

$$\log M_y = \beta_M + \beta_E X_y + \varepsilon_{M,y}$$

where the process errors are modeled as random effects that may, in general, be autocorrelated normal random variables

$$\varepsilon_{M,y}|\varepsilon_{M,y-1} \sim \mathrm{N}\left(\varepsilon_{M,y-1}, (1-\rho_M^2)\sigma_M^2\right)$$

(Stock and Miller, 2021), but we assume $\rho_M = 0$ in our R+M OMs. We assume the median M rate $e^{\beta_M} = 0.2$ is constant across ages. For R and R+S OMs and EMs, $\varepsilon_{M,y} = 0$. For all R+M OMs, we assume the same standard deviation $\sigma_M = 0.3$, which is estimated in the R+M EMs. The covariate effect is one of 3 alternative values in the operating models, $\beta_\mathrm{E} \in \{0, 0.25, 0.5\}$. The parameters defining the simulated covariate time series, size of the covariate effect, and any M random effects result in a range of different levels of variation in annual values (Figure S3).

We assumed expected recruitment each year is from a Beverton-Holt stock-recruit relationship (SRR)

$$R_y = \frac{a\mathrm{SSB}_{y-1}}{1 + b\mathrm{SSB}_{y-1}}.$$

All biological inputs to calculations of spawning biomass per recruit (i.e., weight, maturity, and M at age) are constant in the R and R+S OMs without covariate effects on M. Therefore, steepness and equilibrium unfished recruitment are also constant over the time period for those OMs (Miller and Brooks, 2021). As in Miller et al. (In review), our assumed biological inputs and selectivity (defined below) with constant M result in equilibrium F that reduces spawning biomass per recruit to 40% of the unfished level is $F_{40\%} = 0.348$. With an assumed unfished recruitment of $R_0 = e^{10}$, setting $F_\mathrm{MSY} = F_{40\%}$ results in a steepness of 0.69 and $a = 0.60$ and $b = 2.4 \times 10^{-5}$. For R+M OMs and all OMs with covariate effects on M, steepness is not constant, but we used the same a and b parameters as other operating models which equates to a steepness and R_0 at the median of the time series models for M and the covariate.

We also used the same two fishing scenarios as Miller et al. (In review) for OMs. In the first scenario, the stock experiences overfishing at $2.5F_\mathrm{MSY}$ for the first 20 years followed by fishing at F_{MSY} for the last 20 years (denoted $2.5F_{\mathrm{MSY}} \rightarrow F_{\mathrm{MSY}}$). In the second scenario, the stock is fished at F_{MSY} for the entire time period (40 years). The magnitude of the overfishing assumptions is intended to reflect estimates of overfishing for Northeast US groundfish stocks from Wiedenmann et al. (2019).

We configured all R, R+S, and R+M OMs with uncorrelated random effects on recruitment with standard deviation on log(recruitment) $\sigma_R = 0.5$. This same assumption was used by Miller et al. (In review) for R+M OMs and other OMs with fishery selectivity and index catchability process errors. For R+S OMs, apparent survival process errors were uncorrelated with $\sigma_{2+} = 0.3$.

Catch and index observations

We define the generation of observations of aggregate (total combined across ages) catch and indices, and corresponding age composition identical to Miller et al. (In review). There is a single fleet operating year round for catch observations with logistic selectivity for the fleet with $a_{50} = 5$ and slope $= 1$ (Figure S2, bottom left). Observations are generated for all 40 years of the model. There are two index time series intended to represent fishery-independent surveys occurring in the spring (0.25 way through the year) and the fall (0.75 way through the year). Catchability of both surveys are assumed to be 0.1. We assumed catch and index age composition observations are generated from a logistic-normal distribution where errors on the multivariate normal scale are independent. The standard deviation parameter is also constant across ages.

Standard deviation for log-aggregate catch was 0.1. There were two levels of observation error variance for aggregate indices and age composition for both indices and fleet catch. A low uncertainty specification assumed standard deviation of both series of log-aggregate index observations was 0.1 and the standard deviation of the logistic-normal for age composition observations was 0.3. In the high uncertainty specification the standard deviation for log-aggregate indices was 0.4 and that for the age composition observations was 1.5. For all estimating models, the standard deviation for log-aggregate observations was assumed known at the true value whereas that for the logistic-normal age composition observations was estimated.

Estimating models

Estimating models were fit to each of 100 simulated data sets from each operating model. There were three factors defining the configuration of each estimating model: 1) whether β_M was estimated or assumed known, 2) whether an environmental effect β_E was estimated or not, and 3) whether the process errors were assumed on recruitment only (R), recruitment and survival (R+S), or recruitment and M (R+M).

The configuration of the process errors in the estimating models generally matched the corresponding options in the operating models. For example, uncorrelated R+S was assumed for both the estimating and operating model. However, R+M EMs did not assume M random effects were uncorrelated (ρ_M was estimated). The environmental covariate observations were included in all estimation models, whether effects on M were estimated or not, to ensure comparability of AIC. All fixed effects parameters for selectivity, catchability, fully-selected F, mean recruitment, initial abundance at age, and variances for logistic-normal age composition distributions were estimated. Any process error variance parameters for recruitment, survival, and M were also estimated. The observation error variance of the environmental observations and aggregate catch and indices were all assumed known at the true values.

Measures of reliability

EM convergence

We measured the frequency of convergence when fitting each EM to the simulated data sets for each OM. There are various ways to assess convergence of the fit (e.g., Carvalho et al., 2021), but we defined successful convergence as the Hessian of the marginal log-likelihood being invertible and providing variance estimates for the fixed effects parameters as recommended by Miller et al. (In review).

AIC for model selection

We measured the frequency of correct model selection using marginal AIC. For a given operating model the set of models that were considered all made the same assumptions on whether or not to estimate β_M or it is assumed at the true value. For model m, the marginal AIC is a function of the marginal log-likelihood maximized with respect to the fixed effects in the model $\boldsymbol{\theta}$ and the number of fixed effects $n(\boldsymbol{\theta})$ estimated,

$$\mathrm{AIC}_m = -2\left[\mathrm{argmax}_{\boldsymbol{\theta}} \log L_m(\boldsymbol{\theta}) - n(\boldsymbol{\theta})\right].$$

All model fits that successfully completed the optimization were used for this set of analyses. We did not condition on convergence as defined above because some lack of convergence would be expected for the correct behavior of more complicated models that include process errors that did not exist in the operating model. For example R+M EMs fit to R OMs would be expected to estimate no variance in the M random effects and the estimated variance parameter going to zero would cause poor convergence but have the same marginal log-likelihood and therefore higher AIC as expected. The correct EM makes the correct assumption for the source of process errors (R, R+S, R+M) and either includes a covariate effect on M when an effect is simulated ($\beta_E \in \{0.25, 0.5\}$ or does not include an effect when one is not simulated ($\beta_E = 0$).

Parameter estimation bias and accuracy

All results here use OM simulations with fits that satisfied the convergence criterion described above. We used this conditioning to reflect how practitioners would proceed in analyses of model fits with real assessment data. That is, practitioners would ensure models converged such that Hessian-based standard errors were available for all model parameter estimates.

We focused on statistical behavior of estimators of the covariate effect on natural mortality ($\widehat{\beta}_E$), the estimator of the median M parameter ($\widehat{\beta}_M$), and estimators of terminal year M ($\widehat{M}$), spawning stock biomass ($\widehat{\text{SSB}}$), and fully-selected F ($\widehat{F}$). In preliminary analyses we examined results for the estimators of all the annual values for M, SSB and F over the whole time series, but we found no appreciable differences in patterns across the various factors defining the OMs and EMs. Furthermore, results for terminal year F were generally inversely related to those for spawning stock biomass, and are provided in the Supplementary Materials and not discussed further.

We calculated errors of $\widehat{\beta}_E$ and $\widehat{\beta}_M$, and the Hessian-based standard error estimators of these parameters ($\widehat{SE}\left(\widehat{\beta}_E\right)$ and $\widehat{SE}\left(\widehat{\beta}_M\right)$) as

$$\delta_i = \widehat{\theta}_i - \theta_i$$

where θ_i is the true value of a given parameter or population attribute for simulation i and $\widehat{\theta}_i$ is the estimate from a model fitted to data from simulation i. For standard errors the true standard error was estimated as the standard deviation of estimates across fits for simulations of a given OM and EM configuration. We calculated the relative errors for terminal year M, SSB, and F

$$\text{RE}_i\left(\theta_i\right) = \frac{\widehat{\theta}_i - \theta_i}{\theta_i}.$$

To measure accuracy we also calculated the root mean square error (RMSE),

$$\text{RMSE}\left(\widehat{\theta}\right) = \sqrt{\frac{1}{n}\sum^{n}\left(\widehat{\theta}_i - \theta_i\right)^2},$$

where n is the number of simulations with converged fits of an EM and OM configuration. The true values for terminal year SSB, and in some OMs M (R+M OMs and any OM with covariate effects on M), vary among simulations. Finally, we recorded whether 95% CIs for $\widehat{\beta}_E$ and $\widehat{\beta}_M$ using Hessian-based standard error estimates for converged EMs that estimated these parameters provided simulations.

Summarizing results across OM and EM attributes

There are numerous OM and EM attributes across all simulation scenarios and, like Miller et al. (In review), we used two methods to summarize the major factors explaining differences in results for each measure of reliability. First, we fit regression models with the responses as the measures of reliability described above and the predictor variables were defined based on OM and EM attributes (e.g., MacKinnon et al., 1995; Wang et al., 2017; Harwell et al., 2018). We used logistic regression for binary indicators of convergence and AIC-based selection of the appropriate assumption of the covariate effect on natural mortality. For indicators of AIC-based selection of EM process error source (multiple categories) we performed multinomial regressions. For other measures of reliability we fit linear regression models to transformed responses. Because relative errors (Eq. **??**) for the various parameters are bounded below at -1, we used a transformation of these values

$$y_i = f\left(\widehat{\theta}_i, \theta_i\right) \tag{1}$$

for relative errors

$$f\left(\widehat{\theta}_i, \theta_i\right) = \log\left[\text{RE}\left(\widehat{\theta}_i, \theta_i\right) + 1\right] \tag{2}$$

for RMSE there is only one value calculated from all simulation fits from a given OM scenario

$$f\left(\widehat{\theta}_i, \theta_i\right) = \log\left[\mathrm{RMSE}\left(\widehat{\theta}\right)\right] \tag{3}$$

We omitted fits where estimated terminal year SSB, M, or F was equal to zero. For all regressions we fit separate models with just individual OM and EM factors included, all factors combined, with all second order interactions, and with all third order interactions. For the multinomial regression, we used the `vglm` function from the VGAM package (Yee, 2008, 2015). We calculated percent reduction of residual deviance (from the null model with no factors) for each of the regression fits. As Miller et al. (In review) note, there is a lack of independence of the results for each EM fitted to the same simulated data set of a given OM and therefore we did not perform formal statistical analyses of effects of OM and EM attributes.

For the second summarization method we used classification and regression trees (Breiman et al., 1984) to illustrate the primary OM and EM attributes and interactions that partition the values for each measure of reliability (e.g., Gonzalez et al., 2018; Collier et al., 2022). We used classification trees for categorical measures (convergence and AIC model selection) and regression trees for the other measures with continuous scales (relative error, RMSE). The response variables were the same as the regressions for the deviance reduction analyses. We used the `rpart` function in the rpart package (Therneau and Atkinson, 2025) to fit trees. Full trees were determined using default settings except that we increased the number of cross-validations to 100. For one classification tree fitted to AIC selection of process error configuration results for R+S OMs, we had to make a small changes to the default prior probabilities (the proportions of each class in the in the data) used in the criterion on whether to allow branches from a given node because the default priors would not produce any branches. For clarity, we manually pruned the full trees to show just the primary branches.

All OM and EM factors we used to examine reductions in deviance and regression trees are given in Table S1. We provide more detailed results for each EM and OM configuration in the Supplementary Materials using similar methods to Miller et al. (In review). We estimated probability of convergence. For AIC-based model selection, we estimated the probability of each EM having lowest AIC. We estimated bias as the median of errors for β_M and β_E and corresonding standard error estimates, and median relative errors for terminal year SSB, F, and M.

Results

EM Convergence

For convergence, the EM process error assumption provided the largest percent reduction in deviance (9-25%) of all OM and EM attributes for all three OM process error types (Table 1). Including second order interactions of OM and EM factors also provided further reduction of residual deviance (between 7 and 11%), indicating successful convergence depended on a combination OM and EM attributes.

Classification trees for each OM process error source all had the primary branch defined using the same attribute (EM process error assumption) that provided the largest reduction in deviance (Figure 1). Across all fits for each OM Process error configuration, convergence was best for R+S OMs, but better convergence occurred when the process error was correct for R and R+S OMs. For R+M OMs, R EMs converged well (88%) and good convergence of correctly specified R+M EMs (80.2%) required constant fishing pressure and median natural mortality rate parameter (β_M) to be known. All branches based on observation error of covariates or other assessment data (indices and age composition) showed better convergence with more precise observations. Better convergence also occurred for EMs that did not estimate covariate effects regardless of whether the OM simulated them, EMs that assumed median natural mortality rate was known, and when OMs assumed greater temporal variability in the true environmental covariate. EMs assuming R+M process error (either correctly or not) converged better for OM with constant fishing pressure.

AIC performance

Process error source selection

The level of error in both the population and covariate observations were the attributes that resulted in the largest percent reductions in deviance for AIC selection of the process error configuration for all OM process error types (Table 2). For R+S and R+M OMs, population observation error provided deviance reduction (19% and 13%, respectively) much larger than any other OM and EM factors whereas covariate observation error provided the largest reduction (8%0 for R OMs. Lesser deviance reduction (2-3%) occurred for the true covariate effect size and whether the EM made the correct assumption about including a covariate effect for R OMs and covariate observation error for R+M OMs. Including second and third order interactions, provided the largest reductions in residual deviance for R OMs (24%) and little reduction for R+S and R+M OMs (4% and 6%, respectively).

For all OM process error sources, the attributes defining the primary branches of classification trees matched those that provided the largest reductions in deviance (Figure 2). AIC performed poorly in selecting the correct process error for R+M OMs where lowest AIC occurred for the simpler R EMs. However, accuracy of AIC-based selection was high for R and R+S OMs. In R+S OM simulations where mis-classification occurred, the simpler R EMs had lowest AIC. For R OMs, there was a small decrease in accuracy when the OMs included the largest covariate effect on M, but in those scenarios, AIC was more accurate for the process error assumption when it also incorrectly assumed no covariate effect.

Covariate effect selection

The factors explaining the largest reduction in deviance for AIC selection of the correct covariate effect assumption depended on the magnitude of the true effect assumed in the OM. When OMs had no covariate effect ($\beta_E = 0$) the OM process errror type provided the largest reduction in deviance (4%) (Table 3). When OMs assumed a covariate effect ($\beta_E > 0$), population observation error and magnitude of temporal variability in the true covariate provided the largest reductions in deviance (4% and 8-21%, respectively). The reduction in deviance provided by the magnitude of covariate temporal variability also increased with larger true covariate effect. Including second and third order interactions provided a modest reduction in deviance (5-7%).

The factors defining the primary branches of classification trees for indicators of AIC selecting the correct covariate effect assumption matched the factors providing the largest deviance reductions (Figure 3). Accuracy of AIC-based model selection for the covariate effect assumption was high overall for OMs with no covariate effect (88%) (low Type I error) and low when there was a covariate effect (23% and 35% for OMs with $\beta_E = 0.25$ and 0.5, respectively) (high Type II error), but higher accuracy occurred for certain combinations of OM attributes. When the true covariate effect was greatest ($\beta_E = 0.5$), AIC accuracy was 91% for models fit to OMs with high temporal variability in the covariate and low error in population and covariate observations regardless of whether the EM process error was correct. However, when the covariate effect was weaker ($\beta_E = 0.25$), high accuracy only occurred for a subset of those OMs with R or R+M process errors (76%), and the further conditioning on OMs with high autocorrelation of the covariate provided higher accuracy (86%). When there was no covariate effect, AIC accuracy was best for OMs with the R process error configuration (95%), and for the other two process error configuration, accuracy was improved if the OMs also had lower population observation error and the EM process error assumption was correct (85%).

Covariate effect inferences

Estimation bias

For regression models fit to errors in estimation of the covariate effect β_E, we found very little reduction in deviance for any of the OM and EM factors (Table 4). For the Gaussian model in these regressions, deviance is the sum of squares and there are extremely large estimates of β_E (and errors) for many simulations resulting in extremely large sums of squares and relative differences from the null model are therefore small. Nevertheless, observation error of the covariate provided the either the largest or second largest reductions in deviance of any of the single OM and EM attributes for OM process error configuration types (0.04% to 0.08%). OM fishing history provided reductions of similar scale for R and R+S OMs (0.03 and 0.06%, respectively. Other factors providing similar reductions were true covariate effect size for R OMs (0.05%) and level of population observation error and EM process error assumption for R+S OMs (0.06% for each). Including second and third order interactions also provided relatively large reductions in residual deviance (0.2% to 1.5%).

Median errors for $\widehat{\beta}_E$ across all simulations for each OM process error configuration were negative but close to 0 (-0.1 to -0.07) and primary branches were based on factors that matched those providing the largest reductions in deviance (Figure 4). Branches based on level of error in population or covariate observations provided better median estimation error (closer to 0) with lower uncertainty. Branches based on level of covariate temporal variation provided better median estimation error with higher temporal variation. For R OMs with the largest true covariate effect, estimation error was poor unless there was also high temporal variation and low observation error for the covariate. We observed high estimation error for some R+M OMs with the largest covariate effect size. The detailed results for all OM and EM configurations demonstrate that the median errors get worse with increased values for the true β_E, indicating that the estimates remain close to zero when the true $\beta_E > 0$ (Figures S14 to S16). For R+S OMs, we also found lower estimation error when the EM process error assumption matched.

Standard error estimation bias

Factors providing the largest reductions in deviance for errors in estimation of the standard error of the covariate effect size ($\widehat{\mathrm{SE}}\left(\widehat{\beta}_E\right)$) were generally similar to those that were most important for estimation of β_E itself (Table S2). Level of covariate observation error and EM process error assumption was important for all OM process error configurations (6-13% and 3-4%, respectively). Level of population observation error provided relatively large reductions in deviance for R+S and R+M OMs (3% and 8%, respectively). The OM fishing pressure history provided a lesser reduction in deviance for all OM process error types (2-4%). Like estimation of β_E, relatively large further reductions in deviance also occurred when second and third order interactions of OM and EM attributes were included in the regression models (11-31%).

Median errors in estimation of the standard error for $\widehat{\beta}_E$ across all simulations for a given OM process error configuration were strongly negative (-0.8 to -0.6), but combinations of OM and EM attributes resulted in improved values closer to zero (Figure S4). For all OM process error types, branches with lower covariate observation error provided better median errors in the standard error estimator (closer to 0). Branches where EMs assumed median natural mortality rate was known also generally improved standard error estimation except for R OMs where covariate observation error was high. However, branches with higher temporal variability in the covariate in R and R+S OMs provided improved median errors. For some configurations of R+S OMs, EMs with matching process error assumptions provided improved standard error estimation, but some EMs with the incorrect process error assumptions provided median errors close to 0 for some R and R+M OMs.

Confidence interval bias

For logistic regressions of indicators of confidence intervals including the true value of β_E, factors that provided the largest reductions in deviance were true covariate effect size for R and R+M OMs (11% and 3%, respectively) and EM process error configuration for R+S OMs (7%) (Table 5). Covariate observation error also provided relatively large reduction in deviance for R OMs (8%). Lesser reductions of 1-2% occurred for level of population observation error and covariate temporal variability for all OM process error types. Including second and third order interactions also provided relatively large further reductions in deviance (5-9%).

Over all simulations rates of 95% confidence interval coverage indicated negative bias with rates ranging 85-88% for the different OM process error configurations (Figure 5). Factors defining the primary branches matched the factors that made the largest reductions in deviance for for fitted logistic regression models. For R OMs, the primary branch based on covariate observation error indicated higher rates of confidence interval coverage with lower error, but when that error was higher, higher rates were associated with OMs where no covariate affect was assumed ($\beta_E = 0$). Conditional on those OMs with some covariate effect was simulated with higher observation error, higher rates of coverage occurred with higher error in the population observations. For R+S OMs, higher rates of CI interval coverage were observed for EMs that assumed R+S or R+M process errors.

The more detailed results for each OM and EM configuration show that good confidence interval coverage can be obtained for all of the investigated covariate effect sizes when there is low covariate and population observation error, high covariate temporal variability and contrast in fishing pressure for R and R+S OMs. The EM process error configuration was not a factor for the R OMs, but R+S or R+M EM configurations were necessary for R+S OMs (Figures S21 to S23).

Median natural mortality rate inferences

Estimation bias

Across all OM process error types, regression models fit to errors in natural mortality rate parameter estimation β_M showed largest reductions in deviance from the type of OM fishing pressure (4-17%), EM process error assumption (2-12%), and the EM covariate effect assumption (3-7%, Figure 6). The level of observation error also provided a similar reduction in deviance for R OMs (3%). Relatively large reductions in deviance also occurred with second and third order interactions (7-16%).

Like $\widehat{\beta}_E$, median errors for $\widehat{\beta}_E$ were negative but close to 0 (-0.1 to -0.07) across all simulations for each OM process error configuration (Figure 6). The primary branches for all OM process error types were defined by the type of OM fishing pressure, with median errors closer to zero when there was a change in fishing pressure for R+S and R+M OMs. Branches based on the OM population observation error showed median errors closer to zero with more precise observations. For R+S and R+M OMs, branches that included the matching EM process error assumption provided better median errors.

Standard error estimation bias

The EM process error assumption provided the largest reduction in deviance for regression models fitted to errors in estimation of the standard errors of $\widehat{\beta}_M$ from R (3%) and R+M (7%) OM simulations (Table S3). The type of OM fishing pressure provided the largest reduction (13%) for R+S OMs and a lesser reduction in devaince for R+M OMs (4%). The EM covariate effect assumption also provided similar reductions in deviance for R+S (9%) and R+M (4%) OMs.

The factors defining the primary branches of regression trees matched those providing the largest reductions in deviance for errors in $\widehat{\mathrm{SE}}\left(\widehat{\beta}_M\right)$, but median errors indicated strong negative bias across all OM process error types (Figure S5). Branches with lower population observation error provided better median errors (closer to 0). For R+S OMs, branches where the EM process error assumption was correct provided better median errors.

Detailed results for each OM and EM configuration indicate little bias in certain scenarios where OMs had temporal contrast in fishing pressure and low population observation error (Figures S34 to S36). For R and R+M OMs with those conditions, all EM process error assumptions provided median errors close to zero except for OMs with the true covariate effect strongest and the true covariate time series was autocorrelated and had high temporal variation. For R+S OMs with those conditions, only EMs with the correct process error assumption exhibited low (absolute) median error.

Confidence interval bias

For logistic regressions of indicators of confidence intervals including β_M, factors that provided the largest reductions in deviance were the OM fishing history and level of OM population observation error for all OM process error types (0.3-6%) (Table 7). Including second and third order interactions provided the largest relative increase in deviance reduction for R+S (8-11%) and R+M OMs (5-8%) and lesser increases for R OMs (3-5%).

The factors defining the primary branches of classification trees for indicators of confidence interval inclusion of β_M matched those providing the largest reductions in deviance for each of the OM process error types ((Figure 7). Across all simulations, rates of 95% confidence interval coverage indicated negative bias with rates ranging 83-87% for the different OM process error types. Branches based on contrast in OM fishing history and lower population observation error indicated less bias in confidence interval coverage except R+S OMs in conditions where the EM process error was incorrectly assumed to be R only.

The more detailed results for each OM and EM configuration show that good confidence interval coverage can be obtained for many R and R+S OM scenarios with low population observation error and contrast in fishing pressure for R and R+S OMs, but coverage was generally unreliable for all R+M OM scenarios (Figures S21 to S23).

In the same EM OM combination we investigated above for $\widehat{\beta}_E$, we observed an opposite negative correlation of $\widehat{\beta}_M$ and its standard error estimates (Figure S40), which would result in CIs being too narrow when $\widehat{\beta}_M$ values are larger than average.

Terminal year natural morality rate

Estimation bias

Regressions of errors in estimation of terminal year natural mortality rate show largest reductions in deviance from whether the EM treats β_M as known (Table 8). For R and R+S OMs, annual natural mortality rates are constant therefore terminal M is known when β_M is assumed known and there will be no error in any of those EMs. Beyond the β_M assumption, the OM fishing history, level of population observation error, and EM assumptions for covariate effect and process error all provided reductions that were also of importance.

Regression trees for errors in estimation of terminal year M have primary branches based on the whether β_M was assumed known or estimated fro all OM process error types (Figure 9). For R and R+S OMs, the branches where β_M is assumed known have no error as expected because yearly M is equal to e^{β}_{M}. For the same branch in R+M OMs, there was also very little error in terminal year M. When M was estimated, median errors were near zero for many scenarios regardless of EM process error configuration. However, median error was problematic for all of the conditional subsets of results where β_M was estimated in regression trees for R+S OMs. Detailed results for each OM and EM attribute show that there was little evidence of bias for R+S OMs when population observation error was low and there was contrast in fishing pressure (Figure S47). In R+M OMs, median errors indicated little bias overall with temporal contrast in fishing pressure, although detailed results demonstrated best reliability when there was also lower population observation error (Figure S48).

Estimation accuracy

Table S4

Like bias of terminal M, whether when median M was known or estimated provided the largest reduction in deviance (22-45%) for accuracy of estimates as measured by RMSE for all OM process error types (Table S4). Including second and third order interactions also provided large further reductions in deviance (21-38%).

Also like bias, regression trees fit to log-RMSE of terminal M showed better accuracy when EMS assumed the median M was known rather than estimated (Figure S6). Better accuracy was also generally associated with lower population observation error and temporal contrast in fishing pressure. For R+S OMs, R+S EMs provided better accuracy than incorrect process error assumptions.

Terminal year spawning stock biomass

Estimation bias

Fits of regression models to log-scale errors of terminal SSB resulted in largest reducitons in deviance for OM fishing history (2-3%), population observation error (1% for R and R+M OMs), and EM median M assumption (2%; Table 9). The type of EM process error assumed also provided a relevant reduction in deviance for R+S OMs (1%). Including second and third order interactions of OM and EM attributes provided further relevant deviance reductions of 5-10%.

The primary branches of regression trees fitted to log-scale errors in terminal SSB were based on type of OM fishing history for all OM process error types, but differences in median errors between branches were small (Figure 8). For R+S and R+M OMs, branches that assumed median M was known indicated less bias than those with median M estimated. For R+S

OMs with constant fishing pressure and median M estimated, less bias was indicated when the EM process error assumption was correct.

Estimation accuracy

Largest reductions in deviance for accuracy of terminal year SSB as measured by RMSE were provided by whether median M was known or estimated (20-21%) and whether there was temporal contrast in fishing pressure (20-33%) across all OM process error types (Table S5). Lesser reductions were also provided by the EM process error assumption for R and R+M OMs. Inclusion of second and third order interactions of OM and EM attributes also provided important reductions in deviance (26-40%).

Primary branches of regression trees fit to log-RMSE of terminal year SSB were based on the same factors that provided the largest reduction in deviance (Figure S7). All branches with temporal change in fishing pressure provided better accuracy than branches with constant fishing pressure. Similarly, better accuracy occurred for branches with median M assumed known rather than estimated, and lower population observation error.

Discussion

Our simulation study demonstrated that estimation of environmental effects on M is possible and reliable in certain scenarios even when the process error was misspecified (e.g., R and R+S EMs fit to R+M OM). In many of these same OMs, frequency of convergence of fitted models did not appear to suffer when covariate effects on M were estimated even when there was no effect simulated. However, these scenarios are information rich in that there was contrast in population size (due to changes in fishing pressure) and the covariate affecting the population, and there was low uncertainty in population and covariate observations. Previous research has shown that estimation of a constant M parameter requires contrast in time series and informative data (Lee et al., 2011), so it is no surprise that estimation of these effects also requires relatively good information via more precise observations and higher contrast in the covariate time series.

Even though estimation of covariate effects was unbiased in many scenarios, AIC could only reliably detect covariate effects for R and R+M OMs with contrast in covariates and low covariate uncertainty. In those scenarios where the covariate effect could be detected, CI coverage for the covariate effect was often biased even when there was little or no bias in the estimators of the effect and its standard error. Cadigan et al. (2024) found CI coverage to be biased for SSB and F estimation in a state-space model in some scenarios, but they attributed the poor coverage to bias in Hessian-based standard error estimation, and their simulations held any process error random effects constant. The coverage bias we observed, at least for some OM-EM combinations, may be related to correlation of the estimators of the covariate effect and the corresponding standard error and therefore consideration of other methods of calculating CIs may be warranted (e.g., those based on profile likelihood and/or Monte Carlo sampling of the log-likelihood surface).

Miller et al. (In review) investigated R+M OMs with two levels of M process error variability ($\sigma_M \in \{0.1, 0.5\}$) and only found AIC able to accurately distinguish R+M process errors with the higher level of process error variability ($\sigma_M = 0.5$). We assumed $\sigma_M = 0.3$, intermediate to the values investigated by Miller et al. (In review), and found process error inferences unreliable for the source of process error, indicating that the level of variability required for detecting M process errors must be greater than $\sigma_M = 0.3$, but may still be less than 0.5. In our results and those by Miller et al. (In review), AIC typically chose R EMs which would indicate the fitted R+M EMs would estimate no variability in the M process errors. Future studies like ours where R+M OMs are simulated with greater variability in M process errors would better inform reliability of covariate effect inferences under such scenarios.

Deriso et al. (2008) attempted to estimate process errors as well as covariate effects with

M for Pacific herring, but similarly found no variability in M, suggesting there was too much uncertainty in the available observations relative to the true temporal variability in M. Given that we found covariate effect inferences using R EMs was reliable in R+M OMs with apparently little variability in M process errors, the findings of Deriso et al. (2008) on covariate effects for Pacific herring would presumably also be robust to true low variability in M. However, they did not account for uncertainty in covariate observations, some of which would probably have substantial uncertainty (e.g., competition and predation covariates). We found higher covariate observation uncertainty to cause true covariate effects to be less detectable using AIC, but we did not investigate the implications for incorrectly assuming no covariate uncertainty for covariate inferences.

Any bias or poor accuracy for annual SSB estimation was primarily a function of whether or not the median M parameter was estimated or known and the types of process errors, rather than the treatment of the covariate effects on M. For example, we found estimation of SSB was better when the EM had the process error correctly specified for R+S OMs. Fortunately, our results and those by Miller et al. (In review) demonstrate that marginal AIC seems to be a good tool for determining whether this source of process error should be included in the model. However, the reliability of the estimation of SSB does break down in the less ideal scenarios when there is higher population observation error, and lack of contrast in fishing pressure (e.g., Figures S53 to S55).

The R+S and R+M EMs both include process error for the survival of cohorts and would be expected to perform similarly, and they did in our simulations when the OM and EM matched. However, we found that the biases using R+M EMs for R+S OMs were generally worse than using R+S EMs for R+M OMs. Additionally, there are implications for biological reference points for R+M EMs (e.g., Legault and Palmer, 2016) that are not present with R+S EMs. So, we recommend following the suggestion from Li et al. (2024) to prefer R+S EMs over R+M OMs unless strong biological evidence is present to support a particular R+M OM. Such support could be found through both biological understanding (e.g., Trijoulet et al., 2020) as well as statistical properties such as a large delta AIC for R+M compared to R+S associated with greater temporal variability in natural mortality (Miller et al., In review).

The ability to accurately infer covariate effects on M in some realistic situations indicates that such investigations may be fruitful. Ability to make inferences could improve further when WHAM is extended to incorporate tagging data (**?**). Tagging data can greatly inform natural mortality estimation (Pollock et al., 1991; Hampton, 2000), and this impact on M estimation should also apply to estimation of covariate effects or unexplained temporal variation in M. Given our findings and planned future WHAM development, we expect investigations of and accounting for covariate effects on M to become more common within the fisheries stock assessment process. At the same time, it will be equally important to conduct research that will improve our understanding of how best to measure the depletion of stocks and determine catch advice for these stocks with covariate effects on M.

Higher rates of coverage could be due to positive bias in SE estimates which would provide larger CIs including the true value more often. However, the true standard error of the estimator of β_M and β_E is unknown and way well be a function of the realized time series of random effects for each simulation. That is, the standard error estimate from the hessian is conditional on the realized time series and the variability it is estimating is the variability of observations conditional on the realized time series. Therefore, the true standard error could be approximated for each realized time series by conditionally simulating the observations and holding the random effects constant. It would be interesting to see how much the true SE varies with alternative realizations of the random effects. This is related to work by Cadigan and colleagues. Confidence interval coverage error could similarly be a function of the realized time series, but assessing the bias of the coverage rate over simulations should still indicate whether the rates are good integrating over that variation.

Should point out that regression fits and classificaiton/regression trees model the means of data subsets and therefore results are sensitive to extreme values. The ggplots and the values we present in nodes of regression trees are medians which are more robust to extreme values. This would explain patterns in medians of the ggplots being somewhat different than the results of the regressions and regression trees. This was what lead us to analyse converged estimates of β_E and β_M rather than all values from fitted models. Using median regression and extending regression trees to use median absolute deviations as loss function may be useful approaches (e.g., Chaudhuri and Loh, 2002). Similarly random forests that include quantile regression methods may be a worthwhile alternative to summarizing relative importance of alternative OM and EM attributes (Meinshausen and Ridgeway, 2006).

Acknowledgements

This work was funded by NOAA Fisheries Northeast Fisheries Science Center.

References

Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J. 1984. Classification and Regression Trees. Chapman and Hall/CRC, New York, NY USA. doi:10.1201/9781315139470. 368p.

Britten, G., Brooks, E.N., and Miller, T.J. 2026. Factors affecting the inference on stock-recruitment relationships in state-space assessment models. Canadian Journal of Fisheries and Aquatic Sciences doi:10.1139/cjfas-2025-0210.

Brodziak, J., Cadrin, S.X., Legault, C.M., and Murawski, S.A. 2008. Goals and strategies for rebuilding New England groundfish stocks. Fisheries Research **94**(3): 355–366. doi:10.1016/j.fishres.2008.03.008.

Cadigan, N.G. 2016. A state-space stock assessment model for northern cod, including under-reported catches and variable natural mortality rates. Canadian Journal of Fisheries and Aquatic Sciences **73**(2): 296–308. doi:10.1139/cjfas-2015-0047.

Cadigan, N.G., Albertsen, C.M., Zheng, N., and Nielsen, A. 2024. Are state-space stock assessment model confidence intervals accurate? case studies with sam and barents sea stocks. Fisheries Research **272**: 106950. doi:10.1016/j.fishres.2024.106950.

Carvalho, F., Winker, H., Courtney, D., Kapur, M., Kell, L., Cardinale, M., Schirripa, M., Kitakado, T., Yemane, D., Piner, K.R., Maunder, M.N., Taylor, I., Wetzel, C.R., Doering, K., Johnson, K.F., and Methot, R.D. 2021. A cookbook for using model diagnostics in integrated stock assessments. Fisheries Research **240**: 105959. doi:10.1016/j.fishres.2021.105959.

Chaudhuri, P. and Loh, W.Y. 2002. Nonparametric estimation of conditional quantiles using quantile regression trees. Bernoulli **8**(5): 561 – 576.

Collier, Z.K., Zhang, H., and Soyoye, O. 2022. Alternative methods for interpreting Monte Carlo experiments. Communications in Statistics - Simulation and Computation pp. 1–16. doi:10.1080/03610918.2022.2082474.

Deriso, R.B., Maunder, M.N., and Pearson, W.H. 2008. Incorporating covariates into fisheries stock assessment models with application to pacific herring. Ecological Applications **18**(5): 1270–1286. doi:10.1890/07-0708.1.

Gonzalez, O., O'Rourke, H.P., Wurpts, I.C., and Grimm, K.J. 2018. Analyzing Monte Carlo simulation studies with classification and regression trees. Structural Equation Modeling: A Multidisciplinary Journal **25**(3): 403–413. doi:10.1080/10705511.2017.1369353.

Haltuch, M.A. and Punt, A.E. 2011. The promises and pitfalls of including decadal-scale climate forcing of recruitment in groundfish stock assessment. Canadian Journal of Fisheries and Aquatic Sciences **68**(5): 912–926. doi:10.1139/f2011-030.

Hampton, J. 2000. Natural mortality rates in tropical tunas: size really does matter. Canadian Journal of Fisheries and Aquatic Sciences **57**(5): 1002–1010. doi:10.1139/f99-287.

Harwell, M., Kohli, N., and Peralta-Torres, Y. 2018. A survey of reporting practices of computer simulation studies in statistical research. The American Statistician **72**(4): 321–327. doi:10.1080/00031305.2017.1342692.

Hoyle, S.D., Maunder, M.N., Punt, A.E., Mace, P.M., Devine, J.A., and A'mar, Z.T. 2022. Preface: Developing the next generation of stock assessment software. Fisheries Research **246**: 106176.

Johnson, K.F., Councill, E., Thorson, J.T., Brooks, E., Methot, R.D., and Punt, A.E. 2016. Can autocorrelated recruitment be estimated using integrated assessment models and how does it affect population forecasts? Fisheries Research **183**: 222–232. doi: 10.1016/j.fishres.2016.06.004.

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.

Lee, H.H., Maunder, M.N., Piner, K.R., and Methot, R.D. 2011. Estimating natural mortality within a fisheries stock assessment model: An evaluation using simulation analysis based on twelve stock assessments. Fisheries Research **109**(1): 89–94. doi: 10.1016/j.fishres.2011.01.021.

Legault, C.M. and Palmer, M.C. 2016. In what direction should the fishing mortality target change when natural mortality increases within an assessment? Canadian Journal of Fisheries and Aquatic Sciences **73**(3): 349–357. doi:10.1139/cjfas-2015-0232.

Legault, C.M., Wiedenmann, J., Deroba, J.J., Fay, G., Miller, T.J., Brooks, E.N., Bell, R.J., Langan, J.A., Cournane, J.M., Jones, A.W., and Muffley, B. 2023. Data-rich but model-resistant: an evaluation of data-limited methods to manage fisheries with failed age-based stock assessments. Canadian Journal of Fisheries and Aquatic Sciences **80**(1): 27–42. doi:10.1139/cjfas-2022-0045.

Li, C., Deroba, J.J., Miller, T.J., Legault, C.M., and Perretti, C.T. 2024. An evaluation of common stock assessment diagnostic tools for choosing among state-space models with multiple random effects processes. Fisheries Research **273**: 106968. doi:10.1016/j.fishres.2024.106968.

Li, C., Deroba, J.J., Berger, A.M., Goethel, D.R., Langseth, B.J., Schueller, A.M., and Miller, T.J. 2026a. Random effects on numbers-at-age transitions implictly account for movement dynamics and improve performance within a state-space stock assessment. Canadian Journal of Fisheries and Aquatic Sciences **83**: 1–18. doi:10.1139/cjfas-2025-0092.

Li, C., Deroba, J.J., Miller, T.J., Legault, C.M., and Perretti, C. 2026b. Guidance on bias-correction of log-normal random effects and observations in state-space assessment models. Canadian Journal of Fisheries and Aquatic Sciences **83**: 1–11. doi:10.1139/cjfas-2025-0093.

MacKinnon, D.P., Warsi, G., and Dwyer, J.H. 1995. A simulation study of mediated effect measures. Multivariate Behavioral Research **30**(1): 41–62. doi:10.1207/s15327906mbr3001_3.

Magnusson, A. and Hilborn, R. 2007. What makes fisheries data informative? Fish and Fisheries **8**(4): 337–358. doi:https://doi.org/10.1111/j.1467-2979.2007.00258.x.

Meinshausen, N. and Ridgeway, G. 2006. Quantile regression forests. Journal of Machine Learning Research **7**(6): 983—-999.

Miller, T.J. and Brooks, E.N. 2021. Steepness is a slippery slope. Fish and Fisheries **22**(3): 634–645. doi:10.1111/faf.12534.

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. and Stock, B.C. 2020. The Woods Hole Assessment Model (WHAM).

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., Curti, K.L., and Hansell, A.C. 2025. Space for WHAM: a multi-region, multi-stock generalization of the Woods Hole Assessment Model with an application to black sea bass. Canadian Journal of Fisheries and Aquatic Sciences **82**. doi:10.1139/cjfas-2025-0097.

Miller, T.J., Britten, G., Brooks, E.N., Fay, G., Hansell, A., Legault, C.M., Li, C., Muffley, B., Stock, B.C., and Wiedenmann, J. In review. An investigation of factors affecting inferences from and reliability of state-space age-structured assessment models. For Canadian Journal of Fisheries and Aquatic Sciences, State-space assessment models special collection.

Myers, R.A. 1998. When do environment–recruitment correlations work? Reviews in Fish Biology and Fisheries **8**(3): 229–249. doi:10.1023/A:1008828730759.

NEFSC 2022a. Final report of the haddock research track assessment working group. Available at https://s3.us-east-1.amazonaws.com/nefmc.org/14b_EGB_Research_Track_Haddock_WG_Report_DRAFT.pdf.

NEFSC 2022b. Report of the american plaice research track working group. Available at https://s3.us-east-1.amazonaws.com/nefmc.org/2_American-Plaice-WG-Report.pdf.

NEFSC 2024. Butterfish research track assessment report. US Dept Commer Northeast Fish Sci Cent Ref Doc. 24-03; 191 p.

Nielsen, A. and Berg, C.W. 2014. Estimation of time-varying selectivity in stock assessments using state-space models. Fisheries Research **158**: 96–101. doi:10.1016/j.fishres.2014.01.014.

Pedersen, M.W. and Berg, C.W. 2017. A stochastic surplus production model in continuous time. Fish and Fisheries **18**(2): 226–243. doi:10.1111/faf.12174.

Pollock, K., Hoenig, J., and Jones, C. 1991. Estimation of fishing and natural mortality when a tagging study is combined with a creel survey or port sampling. American Fisheries Society Symposium **12**: 423–434.

Punt, A.E. 2023. Those who fail to learn from history are condemned to repeat it: A perspective on current stock assessment good practices and the consequences of not following them. Fisheries Research **261**: 106642.

Stock, B.C. and Miller, T.J. 2021. 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. Fisheries Research **240**: 105967. doi:10.1016/j.fishres.2021.105967.

Stock, B.C., Xu, H., Miller, T.J., Thorson, J.T., and Nye, J.A. 2021. Implementing two-dimensional autocorrelation in either survival or natural mortality improves a state-space assessment model for Southern New England-Mid Atlantic yellowtail flounder. Fisheries Research **237**: 105873. doi:10.1016/j.fishres.2021.105873.

Szuwalski, C.S., Ianelli, J.N., and Punt, A.E. 2018. Reducing retrospective patterns in stock assessment and impacts on management performance. ICES Journal of Marine Science **75**(2): 596–609. doi:10.1093/icesjms/fsx159.

Therneau, T. and Atkinson, B. 2025. rpart: Recursive Partitioning and Regression Trees. R package version 4.1.25.

Thorson, J.T. and Minto, C. 2015. Mixed effects: a unifying framework for statistical modelling in fisheries biology. ICES Journal of Marine Science **72**(5): 1245–1256. doi: 10.1093/icesjms/fsu213.

Trijoulet, V., Fay, G., and Miller, T.J. 2020. Performance of a state-space multispecies model: What are the consequences of ignoring predation and process errors in stock assessments? Journal of Applied Ecology **57**(1): 121–135. doi:10.1111/1365-2664.13515.

Wang, S., Cadigan, N.G., and Benoît, H.P. 2017. Inference about regression parameters using highly stratified survey count data with over-dispersion and repeated measurements. Journal of Applied Statistics **44**(6): 1013–1030. doi:10.1080/02664763.2016.1191622.

Wiedenmann, J., Free, C.M., and Jensen, O.P. 2019. Evaluating the performance of data-limited methods for setting catch targets through application to data-rich stocks: A case study using northeast U.S. fish stocks. Fisheries Research **209**(1): 129–142. doi:10.1016/j.fishres.2018.09.018.

Yee, T.W. 2008. The VGAM package. R News **8**(2): 28–39.

Yee, T.W. 2015. Vector Generalized Linear and Additive Models: With an Implementation in R. Springer, New York, NY USA. doi:10.1007/978-1-4939-2818-7. 589p.

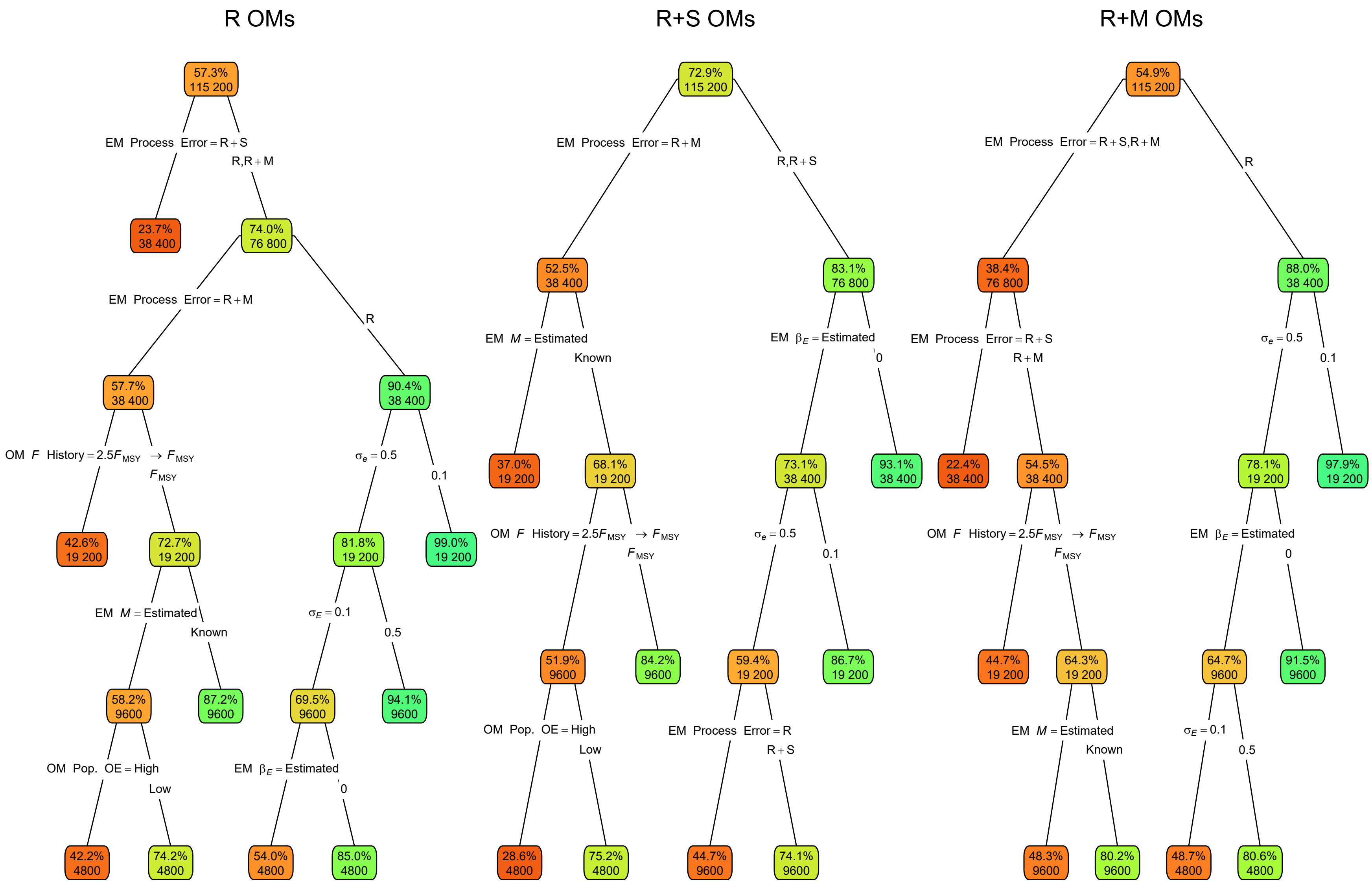

Fig. 1. Classification trees indicating primary factors determining convergence as defined by providing Hessian-based standard errors for R, R+S, and R+M OMs. Nodes denote percent convergence (top) and number of fits (bottom) for the corresponding subset. Lower or higher convergence rates are indicated by more red or green polygons, respectively

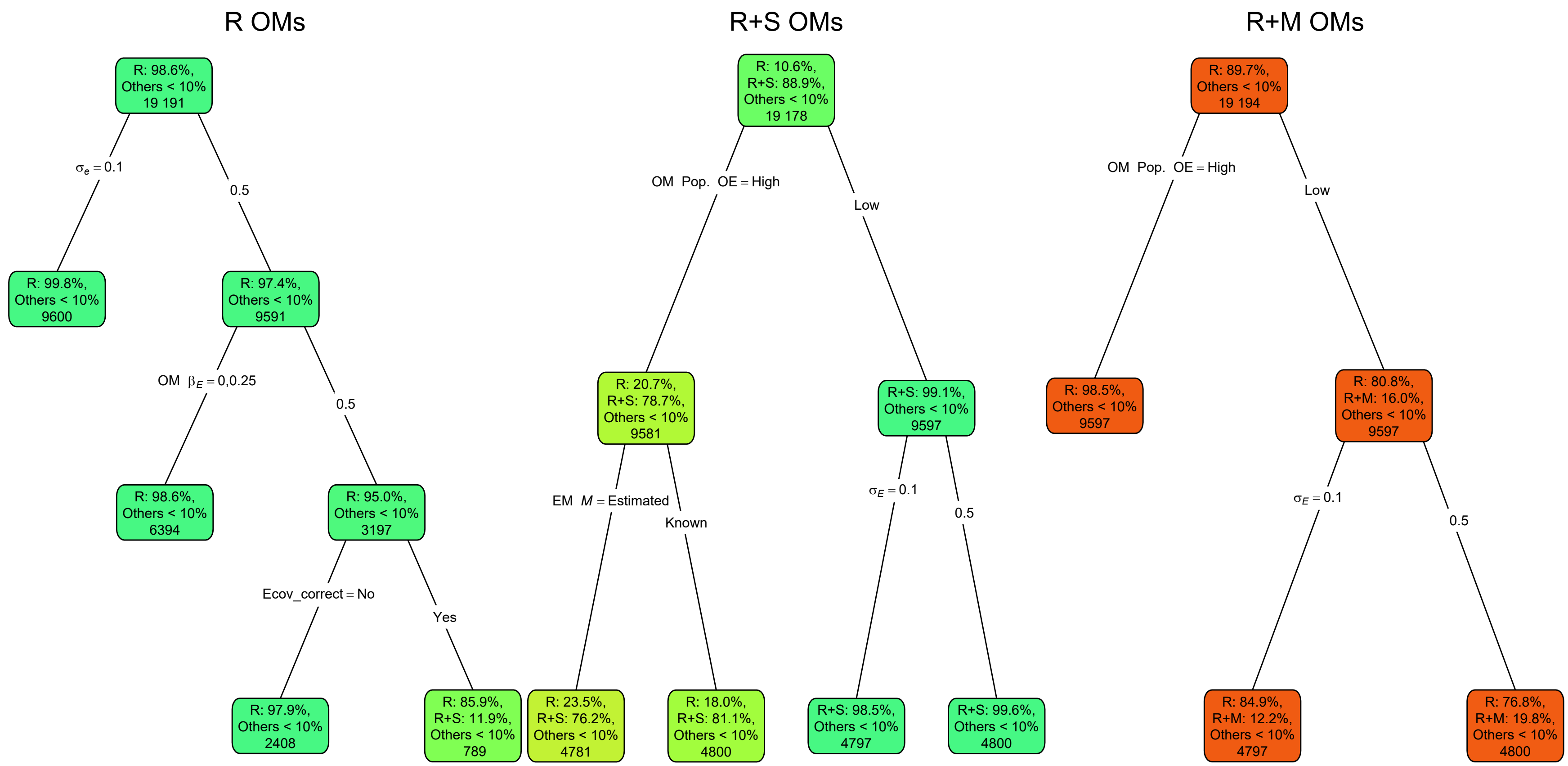

Fig. 2. Classification trees indicating primary factors determining which EM process error assumption provides the lowest AIC for R, R+S, and R+M OMs. Each node shows the percentage of EM process error models with lowest AIC and number of observations for the corresponding subset. Lower or higher accuracy of the process error assumption are indicated by more red or green polygons, respectively.

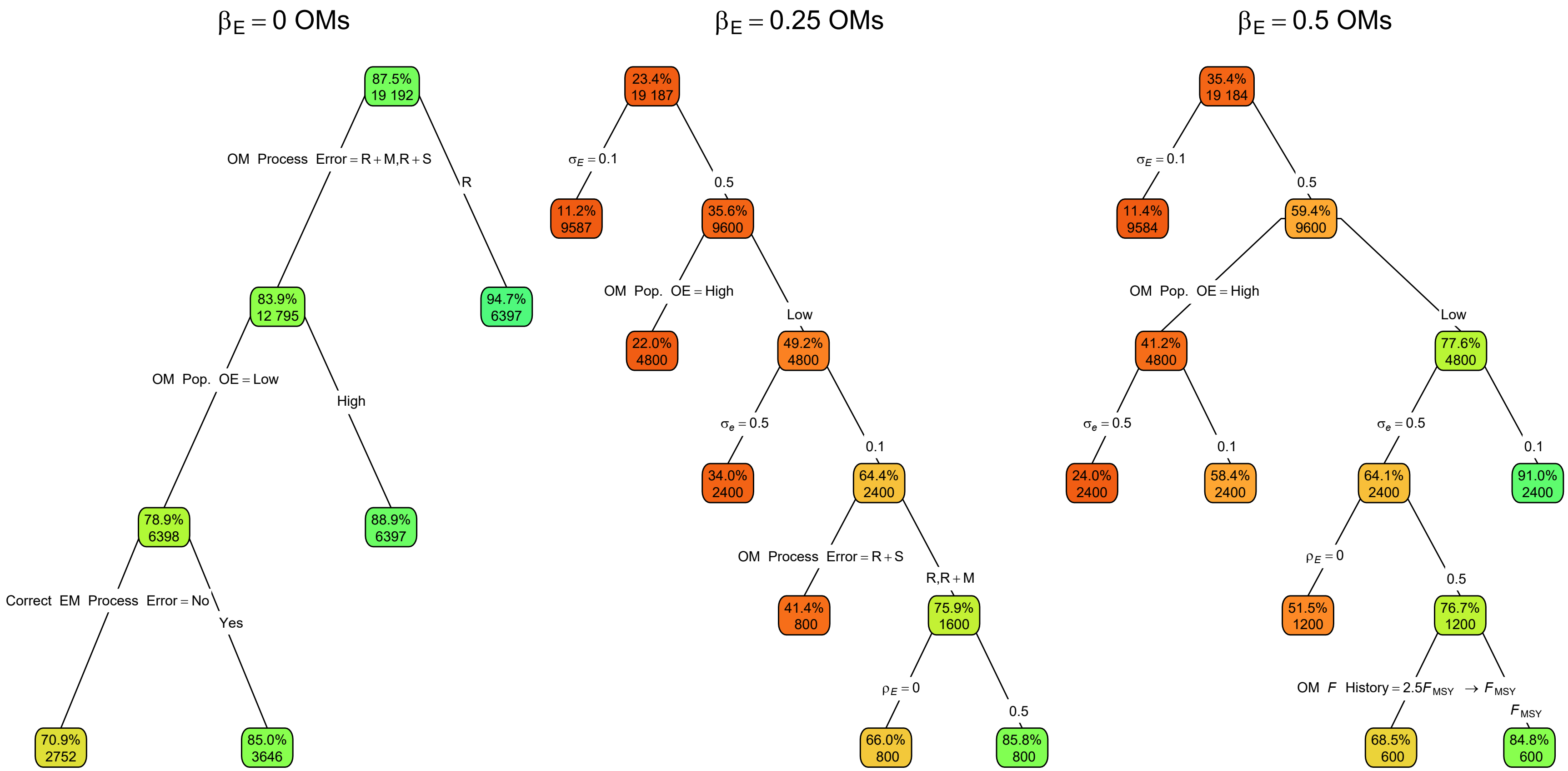

Fig. 3. Classification trees indicating primary factors determining whether the correct EM assumption for covariate effect on natural mortality (no effect with OM $\beta_E = 0$ or estimated effect when OM $\beta_E > 0$) provides the lowest AIC for OMs assuming values for β_E of 0, 0.25, and 0.5. Nodes denote the percentage of EMs with correct assumption and lowest AIC (top), and number of observations (bottom) for the corresponding subset. Lower or higher accuracy of the process error assumption are indicated by more red or green polygons, respectively.

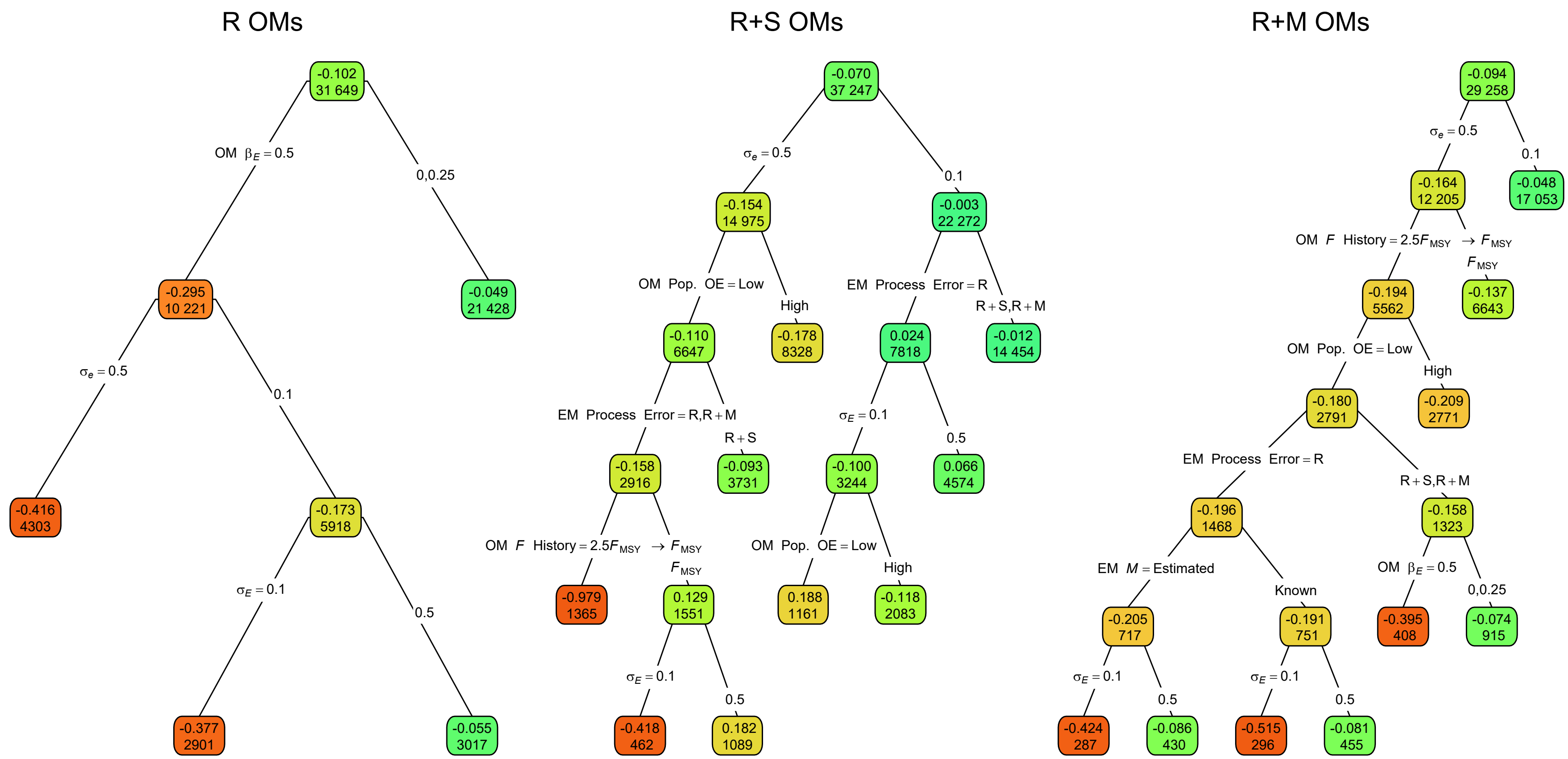

Fig. 4. Regression trees indicating primary factors determining reductions in sums of squares of errors in estimation measured by Eq. **??** for the covariate effect on natural mortality for R, R+S, and R+M OMs. Each node shows the median error (top) and number of observations (bottom) for the corresponding subset. Lower or higher median absolute errors of the process error assumption are indicated by more green or red polygons, respectively.

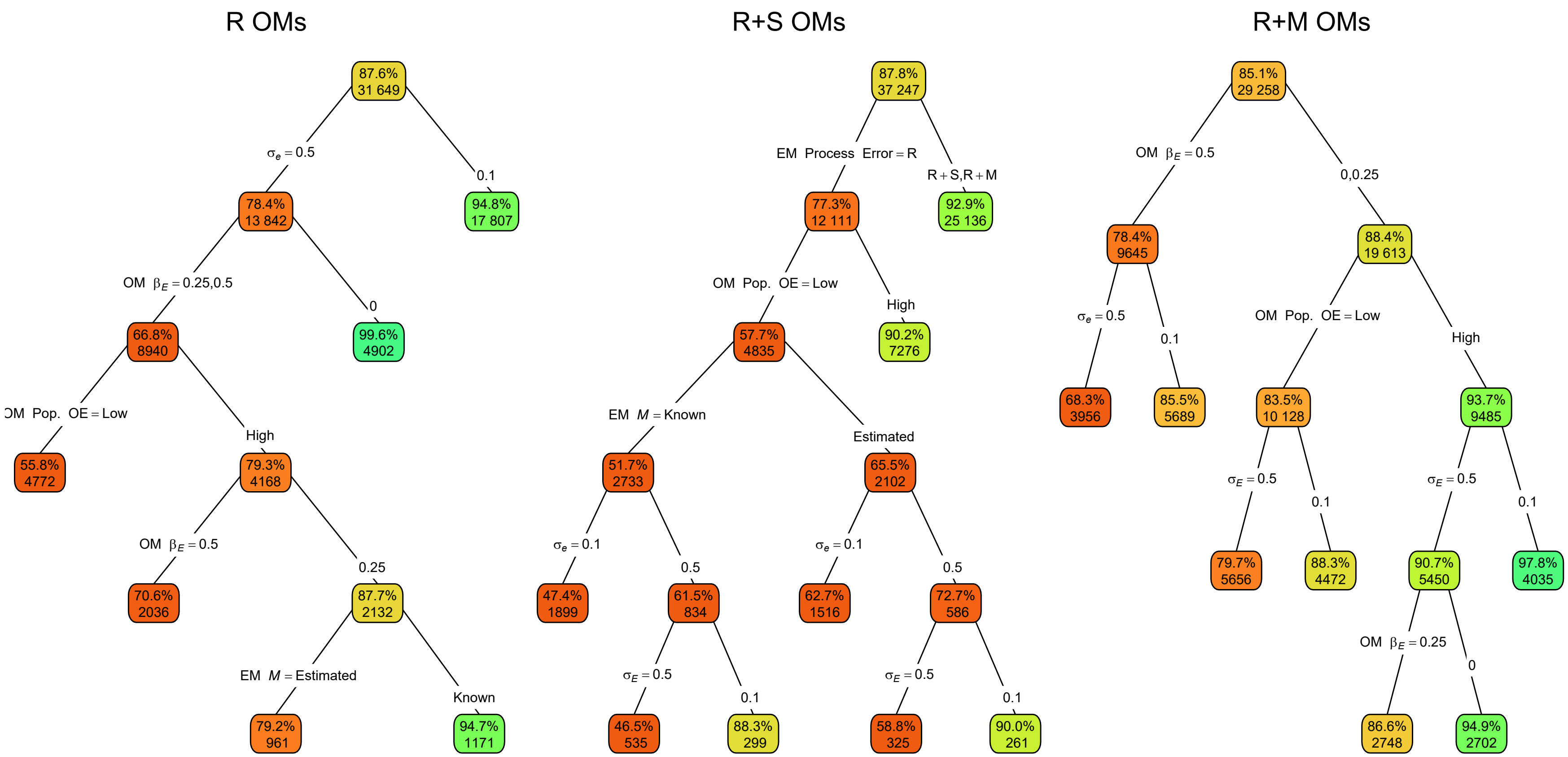

Fig. 5. Classification trees indicating primary factors determining whether confidence intervals for estimated covariate effects on natural mortalityin EMs fitted to R, R+S, and R+M OMs included the true value. Each node shows the median error (top) and number of observations (bottom) for the corresponding subset. Lower or higher median absolute errors of the process error assumption are indicated by more green or red polygons, respectively.

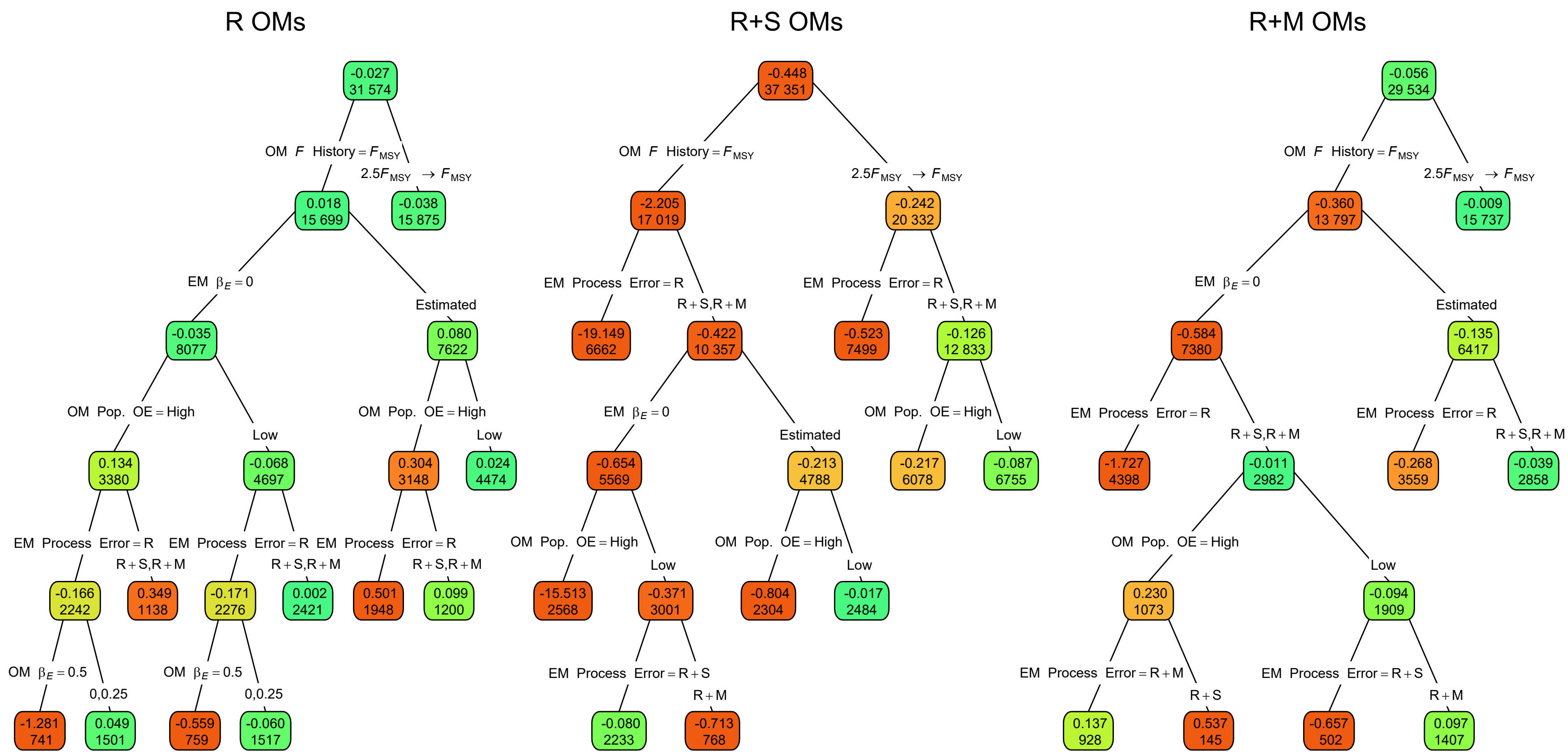

Fig. 6. Regression trees indicating primary factors determining reductions in sums of squares of errors in estimation measured by Eq. **??** for the median natural mortality rate parameter in EMs fitted to R, R+S, and R+M OMs. Each node shows the median error (top) and number of observations (bottom) for the corresponding subset. Lower or higher median absolute errors of the process error assumption are indicated by more green or red polygons, respectively.

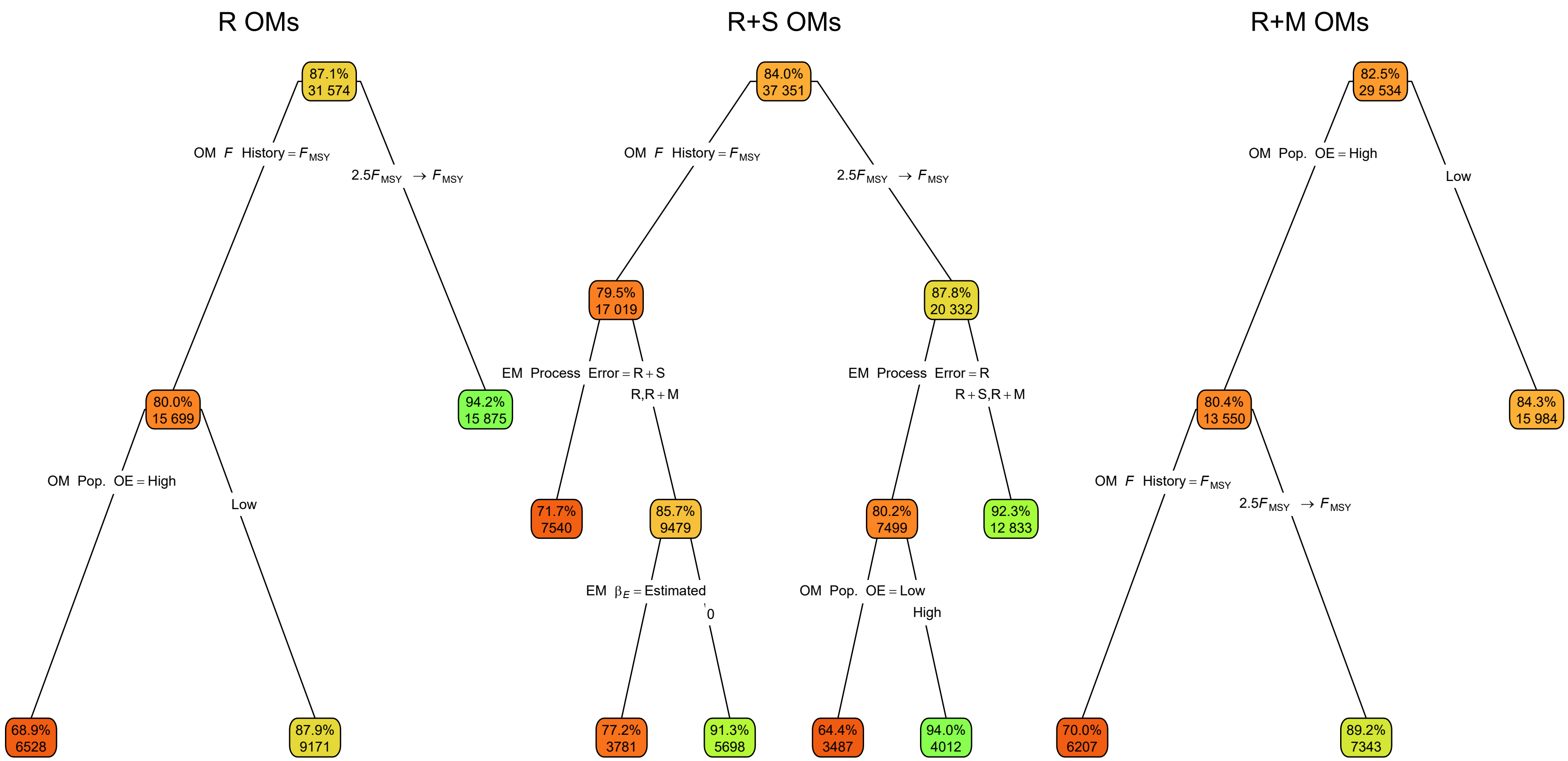

Fig. 7. Classification trees indicating primary factors determining whether confidence intervals for estimated median natural mortality rate parameter in EMs fitted to R, R+S, and R+M OMs included the true value. Each node shows the median error (top) and number of observations (bottom) for the corresponding subset. Lower or higher median absolute errors of the process error assumption are indicated by more green or red polygons, respectively.

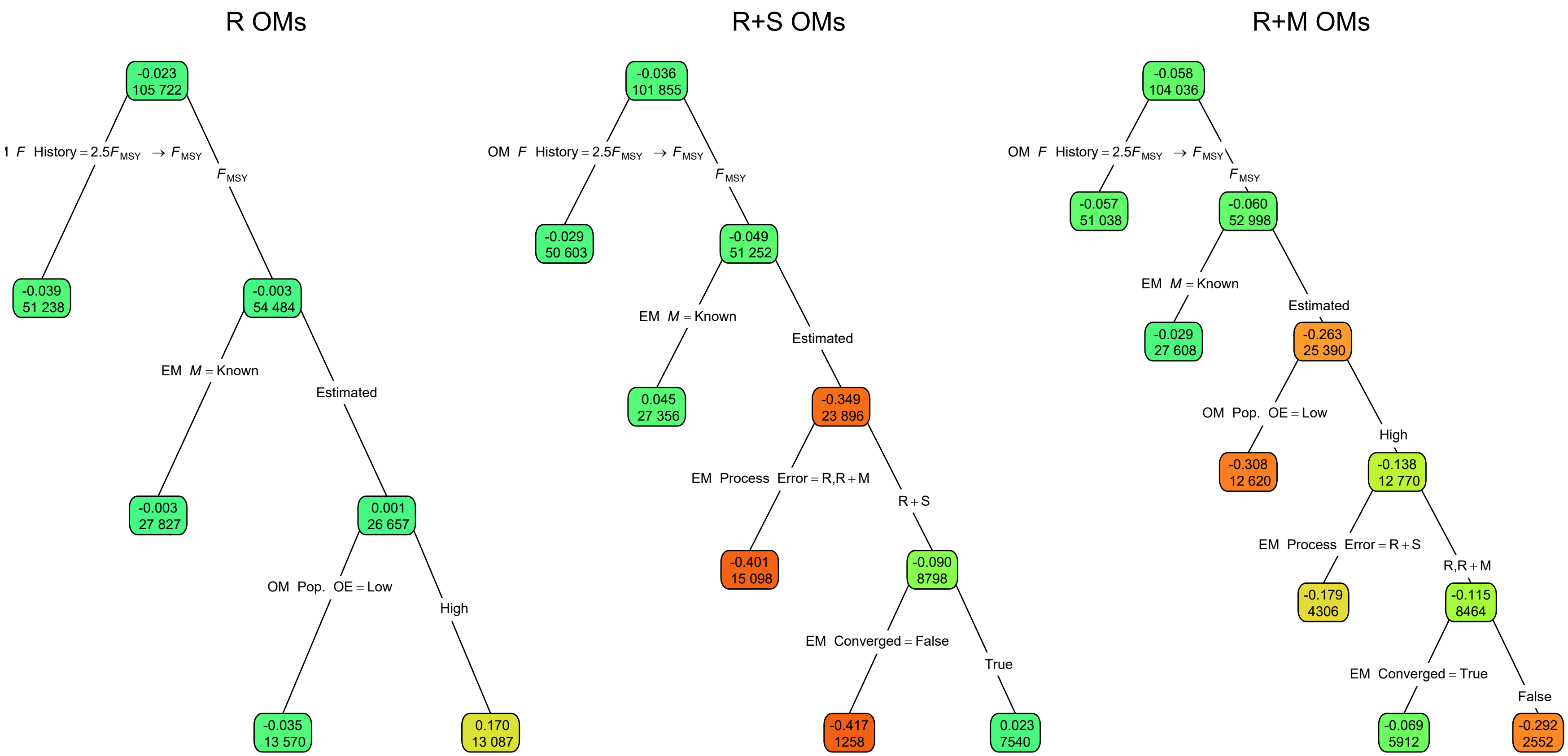

Fig. 8. Regression trees indicating primary factors determining reductions in sums of squares of errors in estimation measured by Eq. **??** for the terminal year SSB in EMs fitted to R, R+S, and R+M OMs. Each node shows the median error (top) and number of observations (bottom) for the corresponding subset. Lower or higher median absolute errors of the process error assumption are indicated by more green or red polygons, respectively.

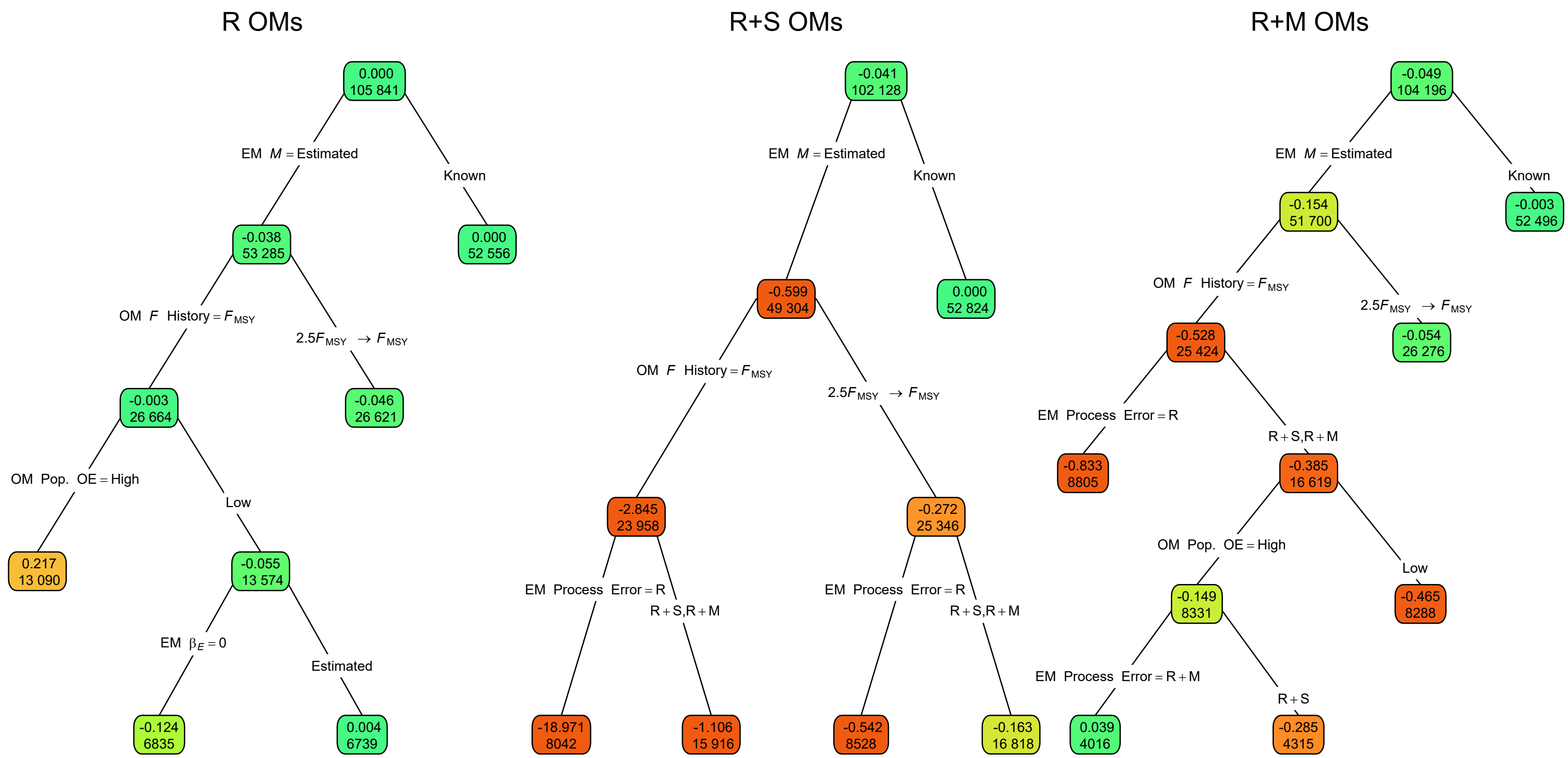

Fig. 9. Regression trees indicating primary factors determining reductions in sums of squares of errors in estimation measured by Eq. **??** for the terminal year natural mortality rate in EMs fitted to R, R+S, and R+M OMs. Each node shows the median error (top) and number of observations (bottom) for the corresponding subset. Lower or higher median absolute errors of the process error assumption are indicated by more green or red polygons, respectively.

Table 1. For each OM process error source (columns), percent reduction in deviance for logistic regression models fit to indicators of convergence (providing Hessian-based standard errors) with each OM and EM factor (rows) included individually, combined, and with second and third order interactions.

Factor	R	R+S	R+M
OM F history	0.86	<0.01	0.24
OM Obs. Error	0.56	0.03	0.26
OM σ_e	0.66	2.99	0.93
OM σ_E	0.53	2.22	0.62
OM ρ_E	<0.01	0.02	<0.01
OM β_E	0.02	<0.01	<0.01
EM Process Error	24.53	9.33	23.06
EM β_E assumption	0.16	2.96	0.51
EM M Assumption	0.18	2.83	0.40
All factors	28.92	22.67	27.34
+ All Two Way	36.15	33.54	34.03
+ All Three Way	37.95	36.58	35.57

Table 2. For each OM process error source (columns), percent reduction in deviance for multinomial logistic regression models fit to indicators of EM process error assumption with lowest AIC with each OM and EM factor (rows) included individually, combined, and with second and third order interactions.

Factor	R	R+S	R+M
OM F history	0.23	0.06	0.30
OM Obs. Error	0.54	18.58	13.43
OM σ_e	8.41	0.05	2.23
OM σ_E	0.17	0.08	0.59
OM ρ_E	1.29	0.03	0.06
OM β_E	3.42	0.04	0.38
EM M Assumption	0.35	0.34	0.48
EM β_E Assumption Correct	2.35	0.22	0.78
All factors	21.97	19.29	19.07
+ All Two Way	38.25	21.13	24.73
+ All Three Way	46.26	22.85	26.97

Table 3. For each OM covariate effect assumption (columns), percent reduction in deviance for logistic regression models fit to indicators of correct EM covariate effect assumption (no effect with OM $\beta_E = 0$ or estimated effect when OM $\beta_E > 0$) with lowest AIC with each OM and EM factor (rows) included individually, combined, and with second and third order interactions.

Factor	$\beta_E = 0$	$\beta_E = 0.25$	$\beta_E = 0.5$
OM F history	<0.01	0.09	0.18
OM Obs. Error	1.19	4.06	4.64
OM Process Error	4.11	0.71	0.25
OM σ_e	0.66	2.00	1.97
OM σ_E	0.35	7.97	20.66
OM ρ_E	0.21	1.23	1.48
EM M Assumption	0.23	0.15	0.15
EM Process Error Assumption Correct	1.95	0.39	0.06
All factors	6.95	17.65	33.44
+ All Two Way	10.58	23.21	38.36
+ All Three Way	11.83	24.95	39.79

Table 4. For each OM process error source (columns), percent reduction in deviance for linear regression models fit to errors in estimation measured by Eq. **??** for the covariate effect on natural mortality ($\widehat{\beta}_E$) with each OM and EM factor (rows) included individually, combined, and with second and third order interactions.

Factor	R	R+S	R+M
OM F history	0.03	0.06	0.02
OM Obs. Error	<0.01	0.06	0.02
OM σ_e	0.04	0.08	0.06
OM σ_E	<0.01	0.02	0.01
OM ρ_E	<0.01	0.01	<0.01
OM β_E	0.05	<0.01	<0.01
EM Process Error	0.01	0.06	0.01
EM M assumption	0.02	0.02	<0.01
All factors	0.18	0.37	0.14
+ All Two Way	0.50	1.11	0.35
+ All Three Way	1.38	1.88	0.66

Table 5. For each OM process error source (columns), percent reduction in deviance for logistic regression models fit to indicators of whether confidence intervals for covariate effect estimates includes the true value with each OM and EM factor (rows) included individually, combined, and with second and third order interactions.

Factor	R	R+S	R+M
OM F history	0.24	0.05	0.01
OM Obs. Error	1.36	1.91	1.78
OM σ_e	8.26	0.24	0.92
OM σ_E	0.92	1.70	1.56
OM ρ_E	0.03	0.05	0.06
OM β_E	11.10	0.68	2.72
EM Process Error	0.01	6.51	0.27
EM M assumption	0.25	0.29	0.06
All factors	23.61	12.16	7.78
+ All Two Way	30.17	18.08	12.63
+ All Three Way	32.58	20.62	14.87

Table 6. For each OM process error source (columns), percent reduction in deviance for linear regression models fit to errors in estimation measured by Eq. **??** for the median natural mortality rate parameter ($\widehat{\beta}_M$) with each OM and EM factor (rows) included individually, combined, and with second and third order interactions.

Factor	R	R+S	R+M
OM F history	4.03	16.52	6.51
OM Obs. Error	2.74	0.54	0.64
OM σ_e	0.01	<0.01	0.01
OM σ_E	0.18	0.02	0.15
OM ρ_E	0.07	0.05	<0.01
OM β_E	0.37	0.04	0.10
EM Process Error	2.31	11.99	2.65
EM β_E assumption	2.83	7.20	3.25
All factors	12.14	33.71	12.81
+ All Two Way	21.74	46.46	19.84
+ All Three Way	26.38	49.41	22.24

Table 7. For each OM process error source (columns), percent reduction in deviance for logistic regression models fit to indicators of whether confidence intervals for median natural mortality rate estimates includes the true value with each OM and EM factor (rows) included individually, combined, and with second and third order interactions.

Factor	R	R+S	R+M
OM F history	6.08	1.46	0.14
OM Obs. Error	1.48	0.41	0.29
OM σ_e	0.01	0.06	0.01
OM σ_E	0.13	0.02	0.04
OM ρ_E	<0.01	<0.01	0.02
OM β_E	0.08	0.11	0.09
EM Process Error	0.53	0.08	0.04
EM β_E assumption	<0.01	0.19	0.01
All factors	8.46	2.32	0.62
+ All Two Way	11.97	10.57	5.99
+ All Three Way	13.35	13.51	8.44

Table 8. For each OM process error source (columns), percent reduction in deviance for linear regression models fit to errors in estimation measured by Eq. **??** for the terminal year natural mortality with each OM and EM factor (rows) included individually, combined, and with second and third order interactions.

Factor	R	R+S	R+M
Convergence	<0.01	0.18	0.01
OM F history	0.66	2.69	1.00
OM Obs. Error	0.29	0.43	0.16
$OM\sigma_e$	<0.01	<0.01	<0.01
$OM\sigma_E$	0.02	<0.01	<0.01
$OM\rho_E$	<0.01	0.01	<0.01
OM β_E	0.03	0.01	0.01
EM Process Error	0.22	1.42	0.43
$EM\beta_E$ assumption	0.35	1.07	0.39
EM M assumption	0.84	5.10	1.22
All factors	2.56	10.83	3.37
+ All Two Way	5.41	18.75	6.38
+ All Three Way	7.93	22.49	8.76

Table 9. For each OM process error source (columns), percent reduction in deviance for linear regression models fit to errors in estimation measured by Eq. **??** for the terminal year SSB with each OM and EM factor (rows) included individually, combined, and with second and third order interactions.

Factor	R	R+S	R+M
Convergence	0.02	0.13	<0.01
OM F history	2.08	2.74	1.75
OM Obs. Error	1.09	0.18	1.19
$OM\sigma_e$	0.01	<0.01	<0.01
$OM\sigma_E$	<0.01	<0.01	<0.01
$OM\rho_E$	<0.01	<0.01	<0.01
OM β_E	<0.01	0.01	<0.01
EM Process Error	0.52	1.05	0.57
$EM\beta_E$ assumption	<0.01	0.05	0.01
EM M assumption	1.76	2.12	1.51
All factors	5.97	6.58	5.21
+ All Two Way	12.81	13.00	11.97
+ All Three Way	16.41	15.57	15.73

Supplemental Materials

Referenced Figures and Tables

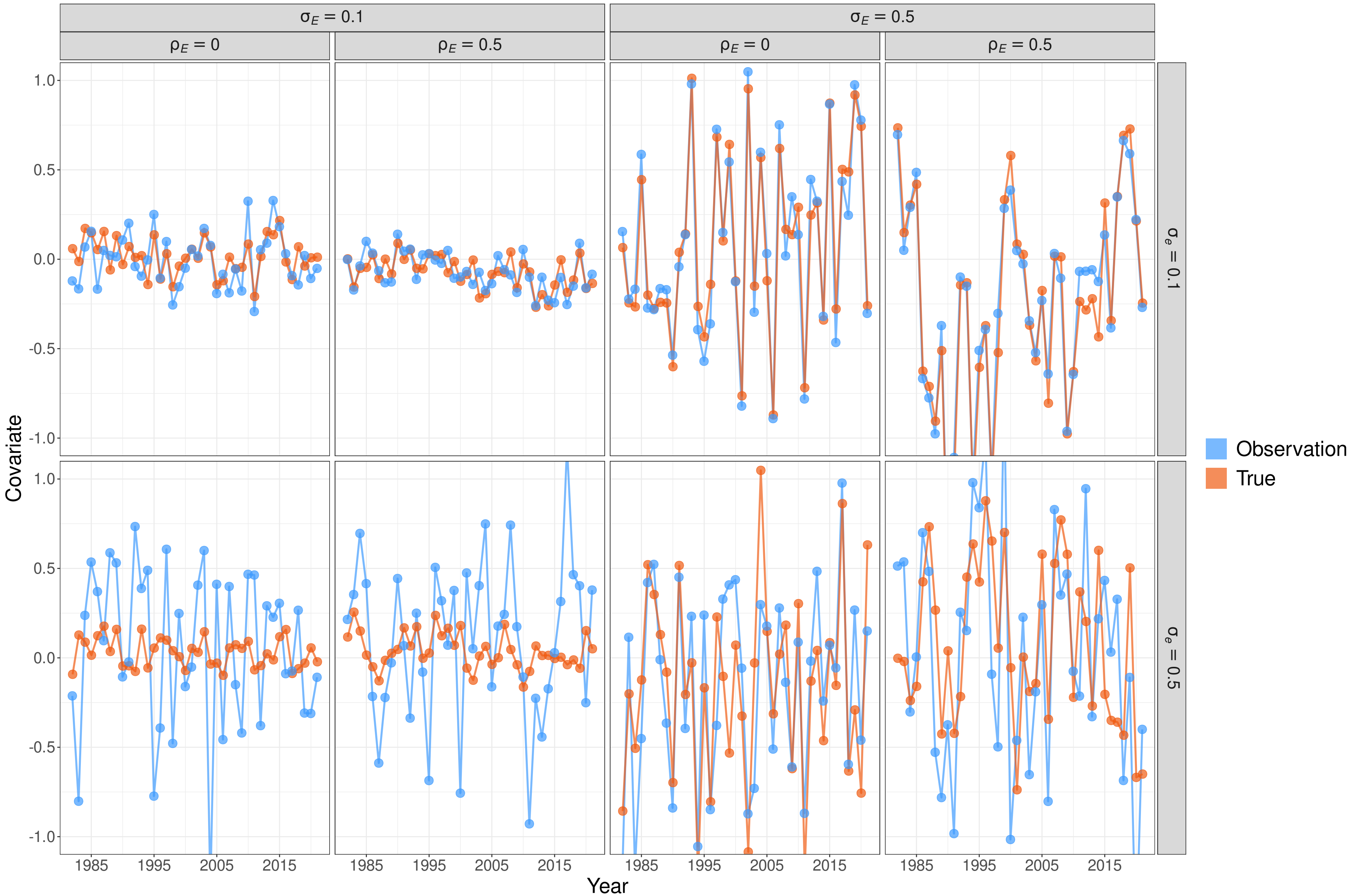

Fig. S1. Example simulations of environmental covariate latent processes and observations with different levels of observation error, and different assumptions about variability of the latent process.

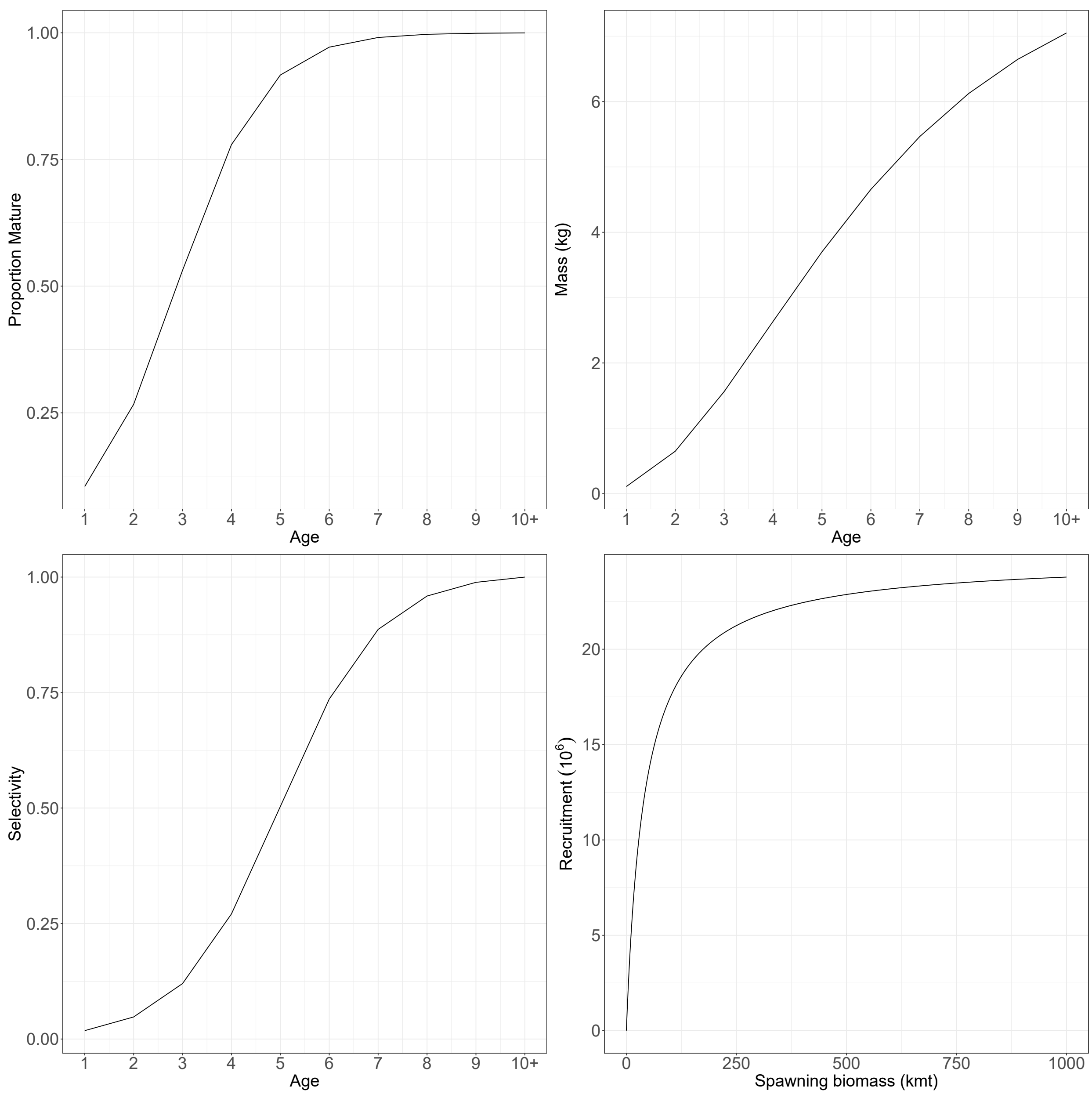

Fig. S2. The proportion mature at age, weight at age, fleet and index selectivity at age, and Beverton-Holt stock-recruit relationship assumed for the population in all operating models. For operating models with random effects on fleet selectivity, this represents the selectivity at the mean of the random effects.

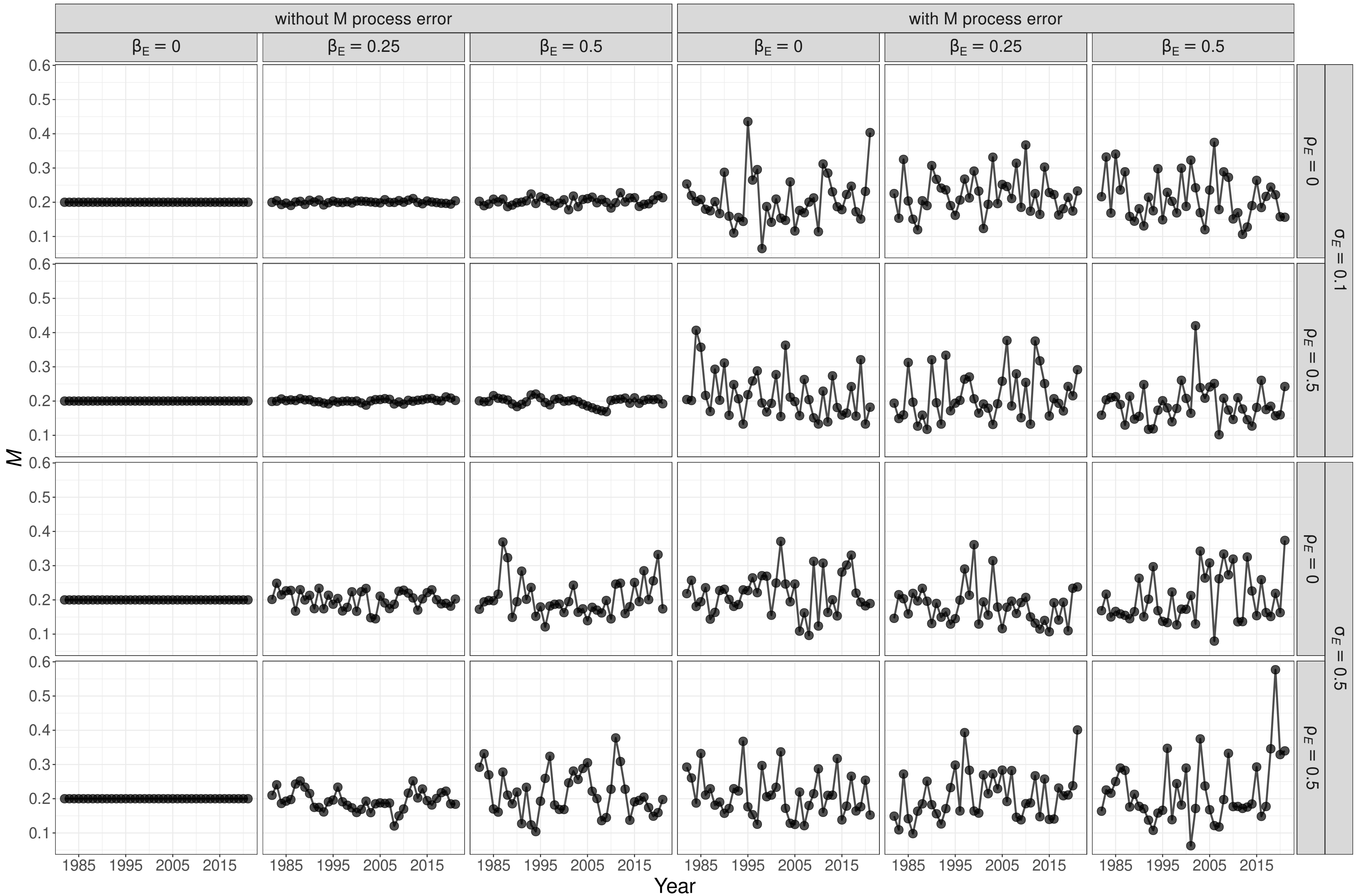

Fig. S3. Example simulations of annual natural mortality rates that may be a function of a temporally varying environmental covariate and autoregressive random effects.

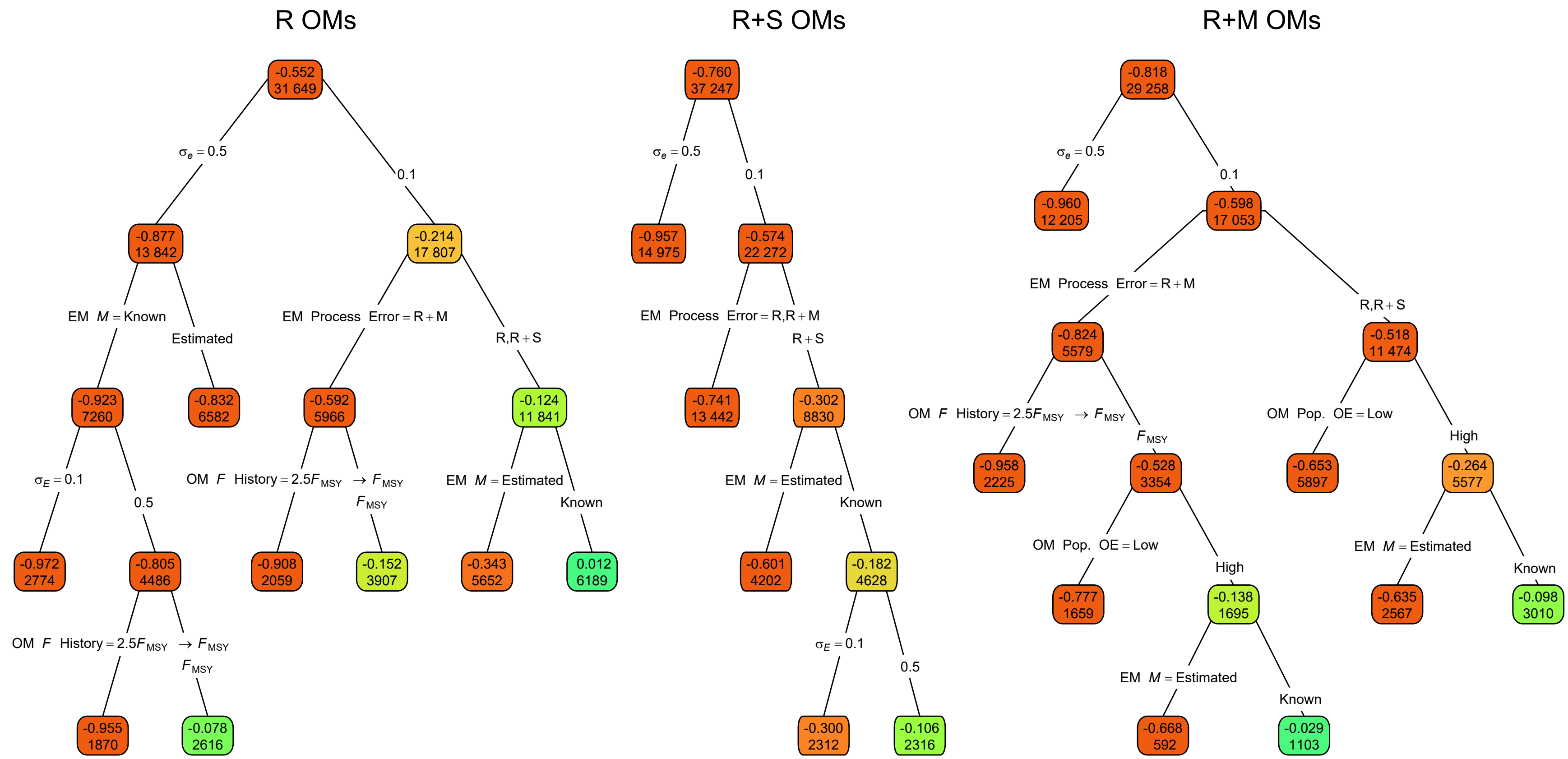

Fig. S4. Regression trees indicating primary factors determining reductions in sums of squares of errors in estimation measured by Eq. ?? for the Hessian-based standard error estimates for covariate effect on natural mortality $\left(\widehat{\mathrm{SE}}\left(\widehat{\beta}_E\right)\right)$ for R, R+S, and R+M OMs. Each node shows the median error (top) and number of observations (bottom) for the corresponding subset. Lower or higher median absolute errors of the process error assumption are indicated by more green or red polygons, respectively.

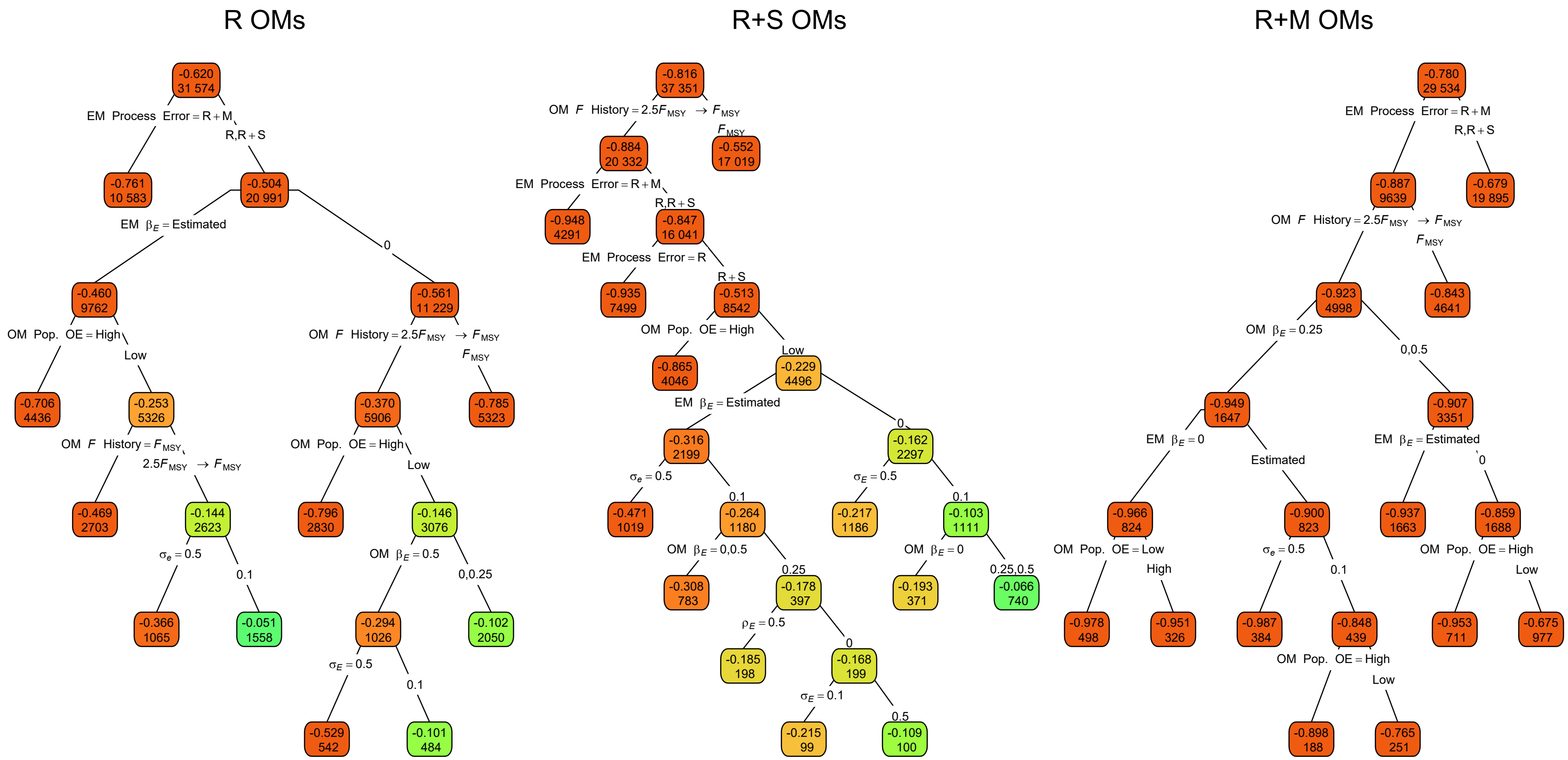

Fig. S5. Regression trees indicating primary factors determining reductions in sums of squares of errors in estimation measured by Eq. ?? for the Hessian-based standard error estimates for median natural mortality rate parameter $(\widehat{\mathrm{SE}}\left(\widehat{\beta}_M\right))$ in EMs fitted to R, R+S, and R+M OMs. Each node shows the median error (top) and number of observations (bottom) for the corresponding subset. Lower or higher median absolute errors of the process error assumption are indicated by more green or red polygons, respectively.

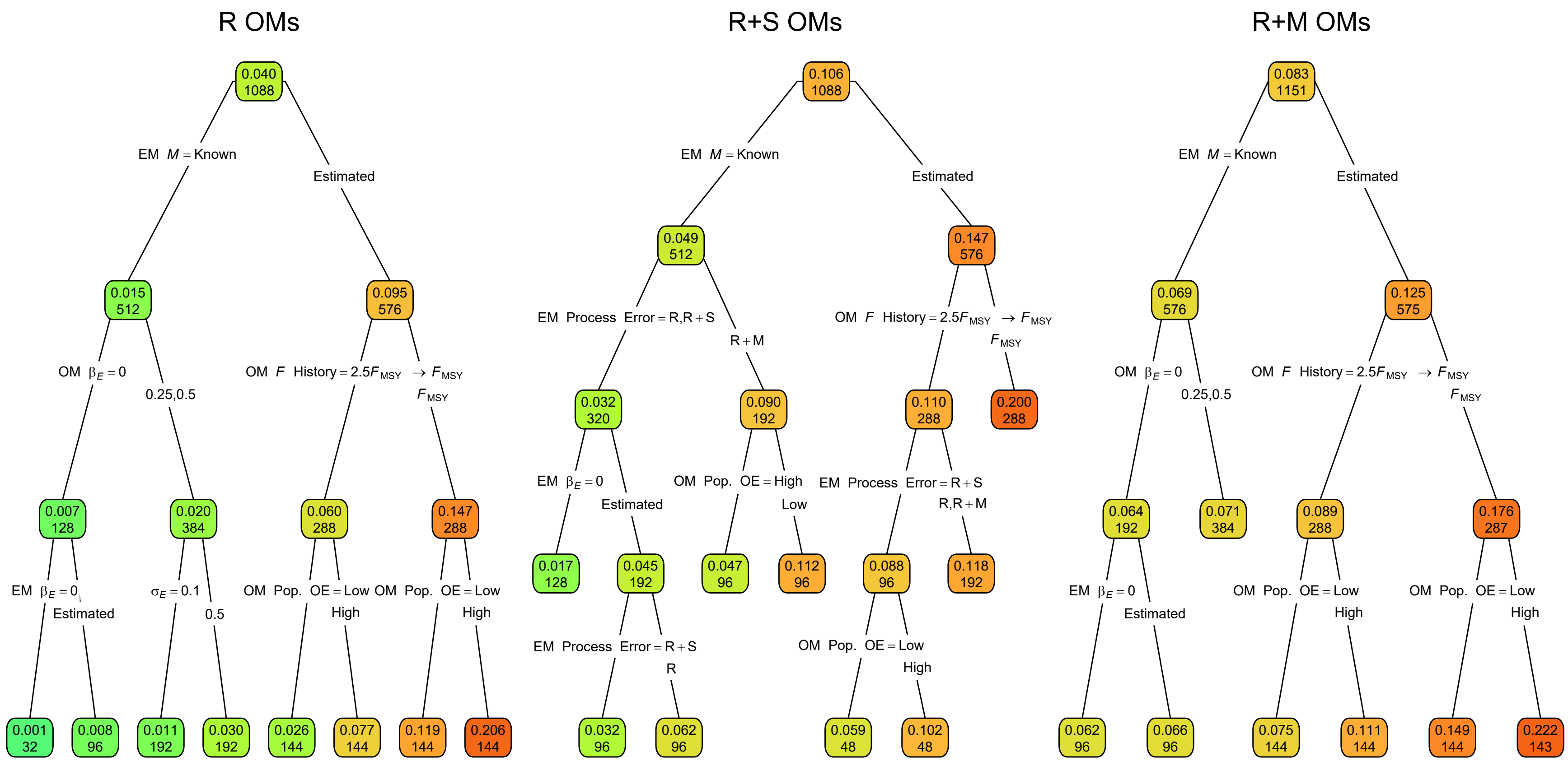

Fig. S6. Regression trees indicating primary factors determining reductions in sums of squares of errors in estimation measured by Eq. **??** for the RMSE of terminal year natural mortality rate in EMs fitted to R, R+S, and R+M OMs. Each node shows the median error (top) and number of observations (bottom) for the corresponding subset. Lower or higher median absolute errors of the process error assumption are indicated by more green or red polygons, respectively.

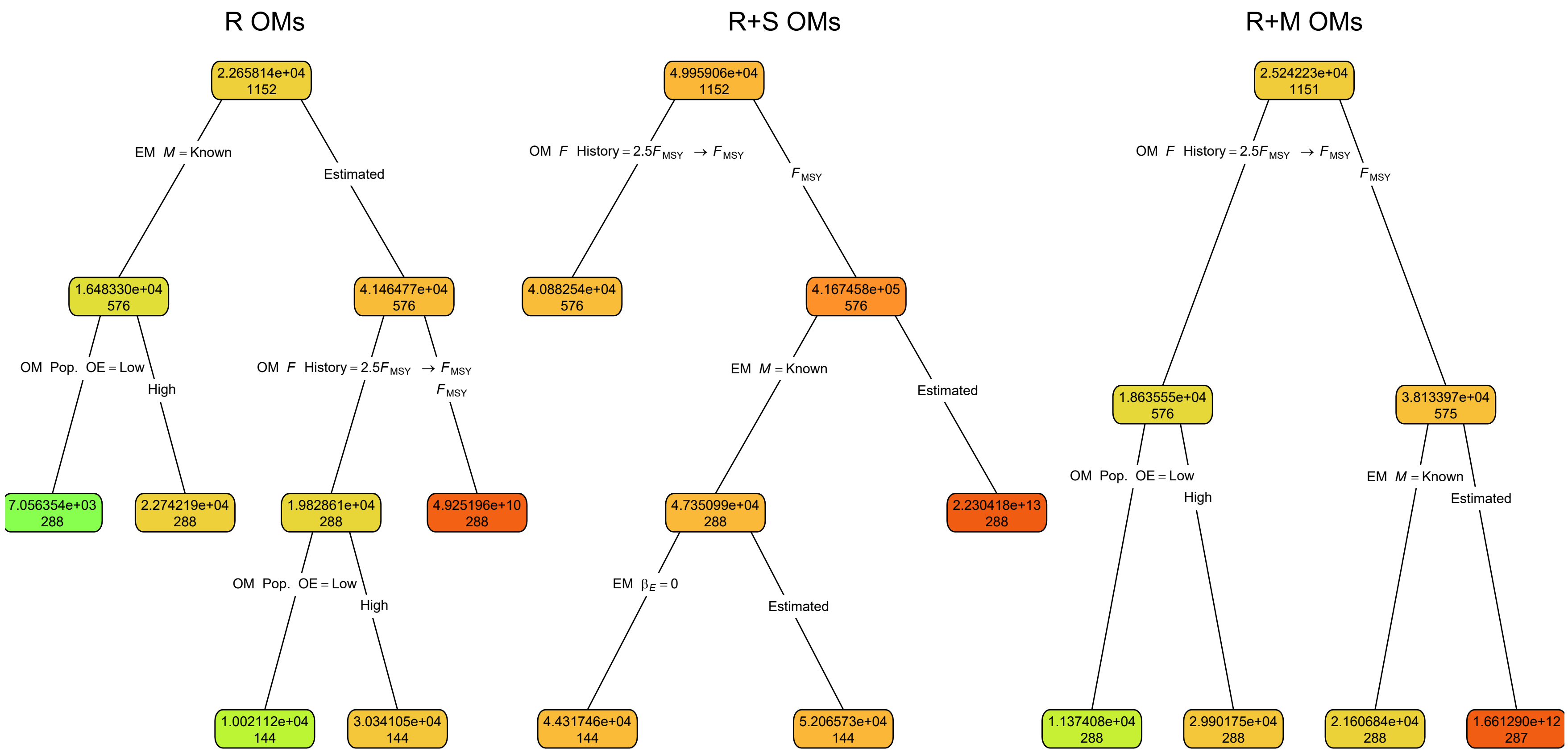

Fig. S7. Regression trees indicating primary factors determining reductions in sums of squares of errors in estimation measured by Eq. ?? for the RMSE of terminal year SSB in EMs fitted to R, R+S, and R+M OMs. Each node shows the median error (top) and number of observations (bottom) for the corresponding subset. Lower or higher median absolute errors of the process error assumption are indicated by more green or red polygons, respectively.

Table S1. OM and EM attributes used in analyses of deviance reduction and classification and regression trees.

Factor	Levels
OM Process error	R, R+S, R+M
OM Fishing History	F_{MSY}, $2.5F_{\text{MSY}} \rightarrow F_{\text{MSY}}$
OM Covariate effect size (β_E)	0, 0.25, 0.5
OM Population Observation Error	Low, High
OM Covariate Observation Error	Low ($\sigma_e = 0.1$), High ($\sigma_e = 0.5$)
OM Covariate Process Error SD (σ_E)	0.1, 0.5
OM Covariate Process Error Correlation (ρ_E)	0, 0.5
EM Process Error	R, R+S, R+M
EM Covariate Effect	None ($\beta_E = 0$), β_E estimated
EM Median Natural Mortality Rate Parameter	Known, Estimated

Convergence results

Table S2. For each OM process error source (columns), percent reduction in deviance for linear regression models fit to errors in estimation measured by Eq. **??** for the Hessian-based standard error of the estimates of the covariate effect on natural mortality ($\widehat{\mathrm{SE}}\left(\widehat{\beta}_E\right)$) with each OM and EM factor (rows) included individually, combined, and with second and third order interactions.

Factor	R	R+S	R+M
OM F history	1.84	2.68	4.00
OM Obs. Error	0.06	2.74	8.04
OM σ_e	6.27	7.28	12.83
OM σ_E	0.48	1.05	0.46
OM ρ_E	0.17	0.01	0.02
OM β_E	0.37	0.01	0.05
EM Process Error	2.90	3.89	2.98
EM M assumption	0.15	0.34	0.47
All factors	13.01	20.95	29.24
+ All Two Way	28.58	32.25	42.77
+ All Three Way	44.44	42.83	51.84

Table S3. For each OM process error source (columns), percent reduction in deviance for linear regression models fit to errors in estimation measured by Eq. **??** for the Hessian-based standard error of the estimates of the median natural mortality rate parameter ($\widehat{\mathrm{SE}}\left(\widehat{\beta}_M\right)$) with each OM and EM factor (rows) included individually, combined, and with second and third order interactions.

Factor	R	R+S	R+M
OM F history	0.55	12.91	3.65
OM Obs. Error	0.98	0.01	0.07
OM σ_e	<0.01	0.03	0.11
OM σ_E	0.13	0.03	0.19
OM ρ_E	<0.01	0.02	<0.01
OM β_E	0.12	0.07	0.23
EM Process Error	3.16	8.79	6.83
EM β_E assumption	0.18	8.96	3.81
All factors	5.32	28.39	14.71
+ All Two Way	11.80	44.67	22.70
+ All Three Way	17.83	51.54	29.46

Table S4. For each OM process error source (columns), percent reduction in deviance for linear regression models fit to RMSE of terminal year M measured by Eq. **??** with each OM and EM factor (rows) included individually, combined, and with second and third order interactions.

Factor	R	R+S	R+M
OM F history	0.26	2.21	14.65
OM Obs. Error	1.65	0.71	4.68
$OM\sigma_e$	0.02	0.16	0.01
$OM\sigma_E$	1.06	1.23	0.84
$OM\rho_E$	0.02	0.01	<0.01
OM β_E	3.73	1.94	1.14
EM Process Error	2.04	3.85	0.87
$EM\beta_E$ assumption	0.19	0.59	0.07
EM M assumption	22.28	45.15	44.70
All factors	32.83	56.71	66.99
+ All Two Way	57.65	84.17	87.96
+ All Three Way	74.93	94.57	91.94

Table S5. For each OM process error source (columns), percent reduction in deviance for linear regression models fit to RMSE of terminal year SSB measured by Eq. **??** with each OM and EM factor (rows) included individually, combined, and with second and third order interactions.

Factor	R	R+S	R+M
OM F history	19.88	32.67	20.92
OM Obs. Error	2.21	0.11	1.48
$OM\sigma_e$	0.05	0.16	0.03
$OM\sigma_E$	0.08	0.16	0.02
$OM\rho_E$	0.03	<0.01	0.02
OM β_E	0.05	0.03	0.14
EM Process Error	5.72	0.13	5.73
$EM\beta_E$ assumption	0.03	0.79	0.05
EM M assumption	20.05	21.11	19.43
All factors	48.09	55.17	47.82
+ All Two Way	79.09	80.98	78.40
+ All Three Way	88.30	89.49	87.67

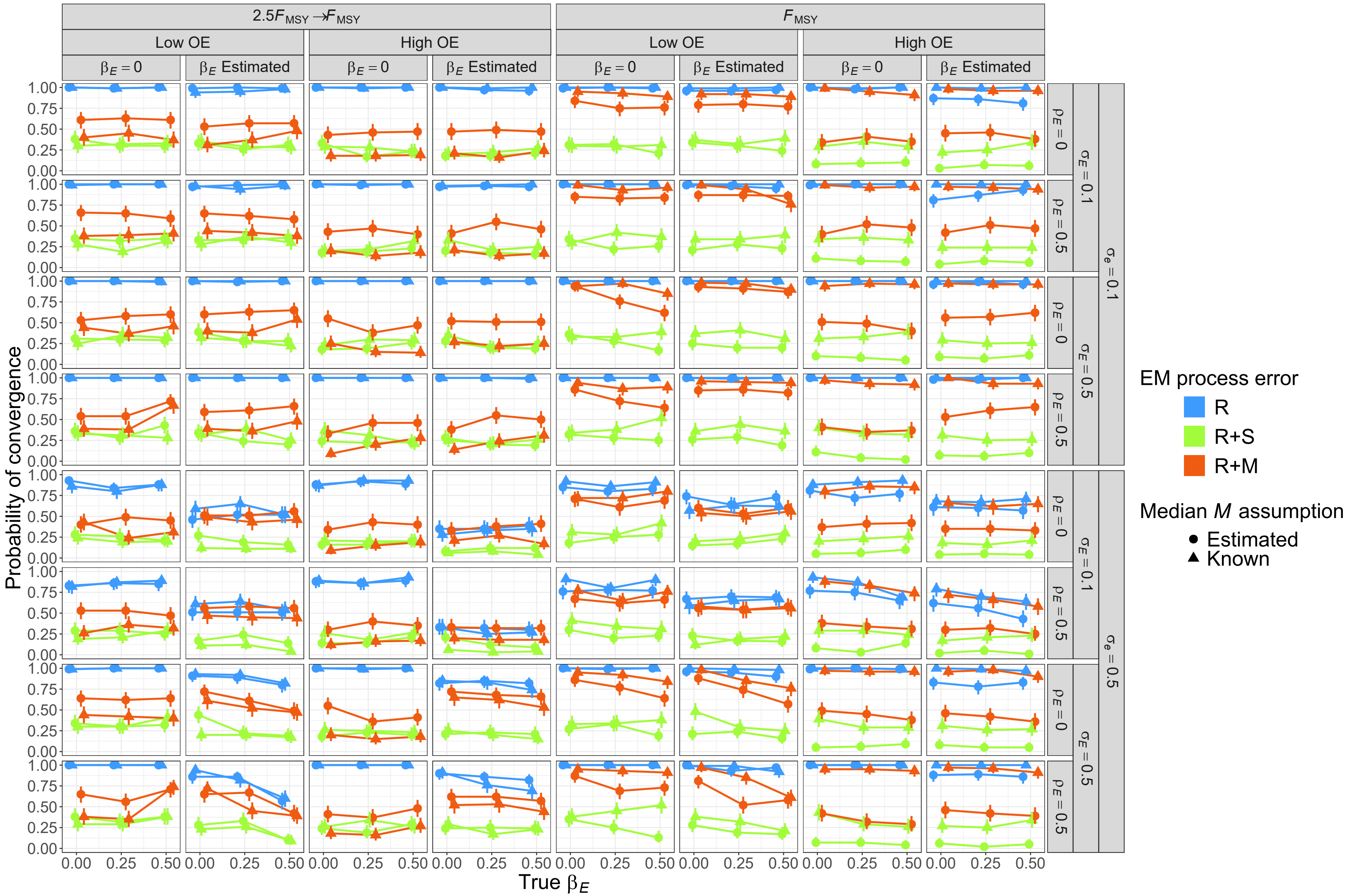

Fig. S8. Estimated probability of fits providing Hessian-based standard errors for EMs assuming alternative process error, that estimate or assume known median natural mortality, and that estimate or assume no covariate effect on median natural mortality when fitted to R OMs and three levels of true covariate effect on median natural mortality (x axis). Vertical lines represent 95% confidence intervals.

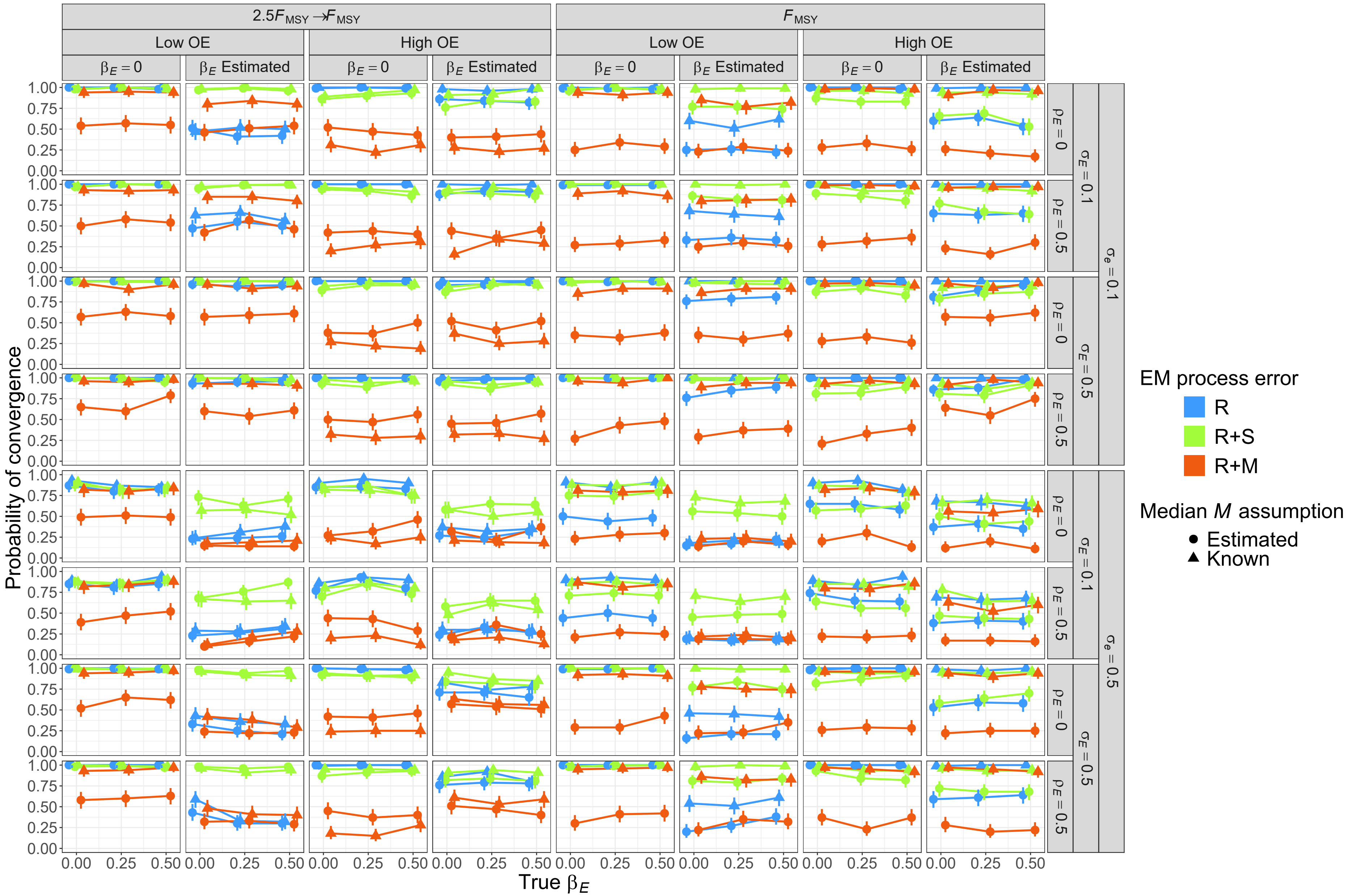

Fig. S9. Estimated probability of fits providing Hessian-based standard errors for EMs assuming alternative process error, that estimate or assume known median natural mortality, and that estimate or assume no covariate effect on median natural mortality when fitted to R+S OMs and three levels of true covariate effect on median natural mortality (x axis). Vertical lines represent 95% confidence intervals.

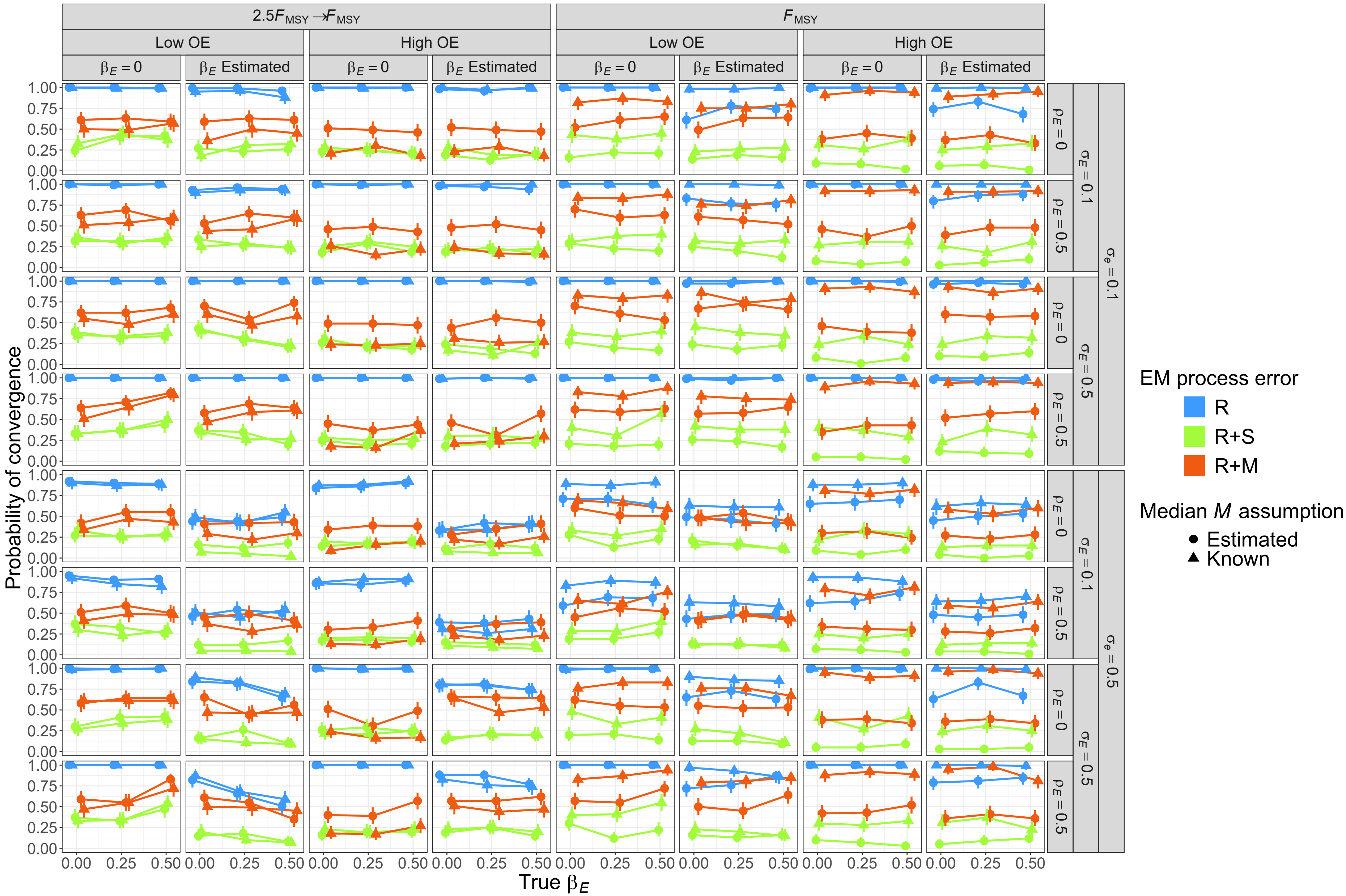

Fig. S10. Estimated probability of fits providing Hessian-based standard errors for EMs assuming alternative process error, that estimate or assume known median natural mortality, and that estimate or assume no covariate effect on median natural mortality when fitted to R+M OMs and three levels of true covariate effect on median natural mortality (x axis). Vertical lines represent 95% confidence intervals.

AIC results

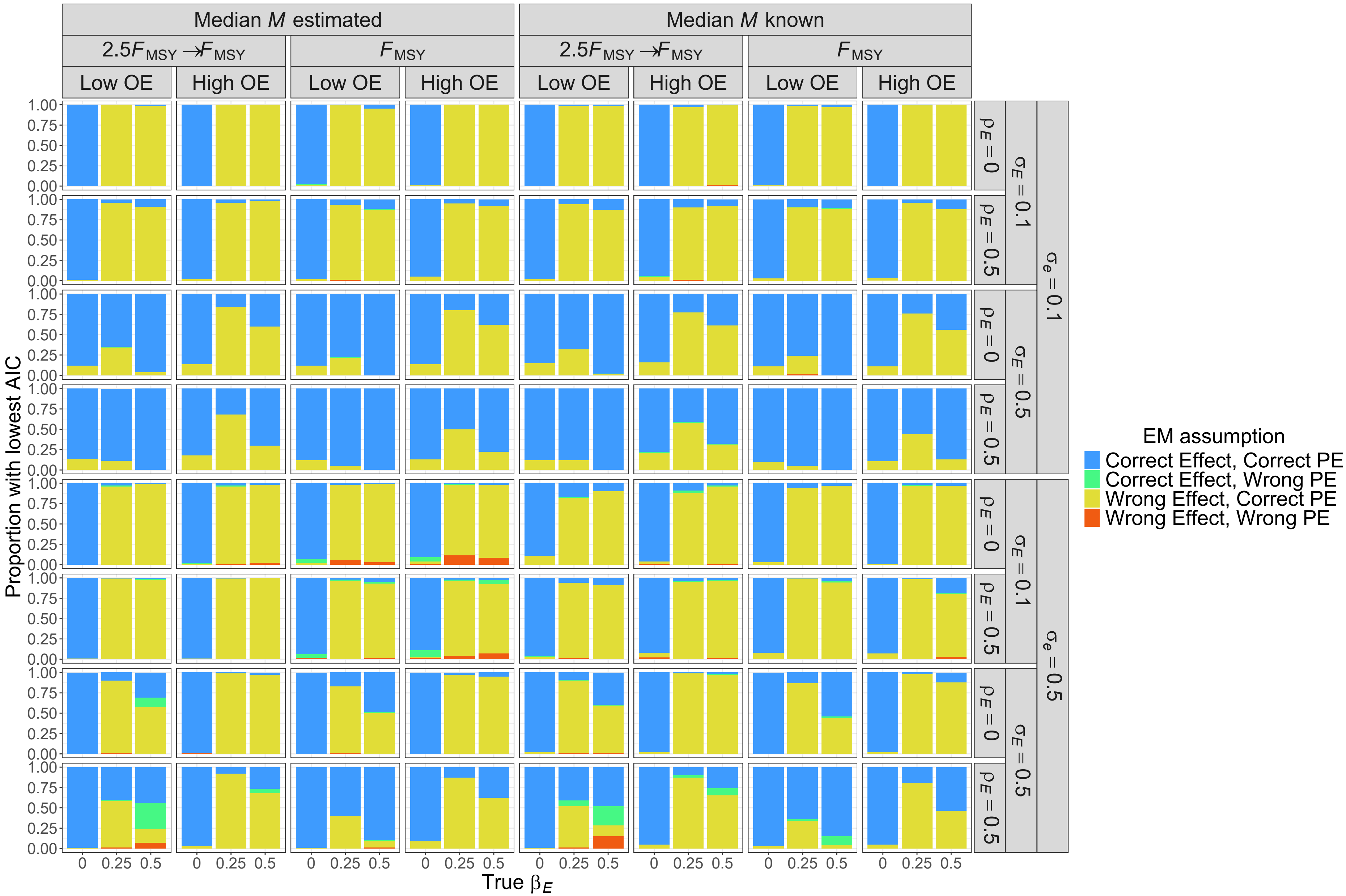

Fig. S11. Proportion of simulated data sets for R OMs where the EM type (treatment of environmental covariate and assumed process error type) had the lowest AIC.

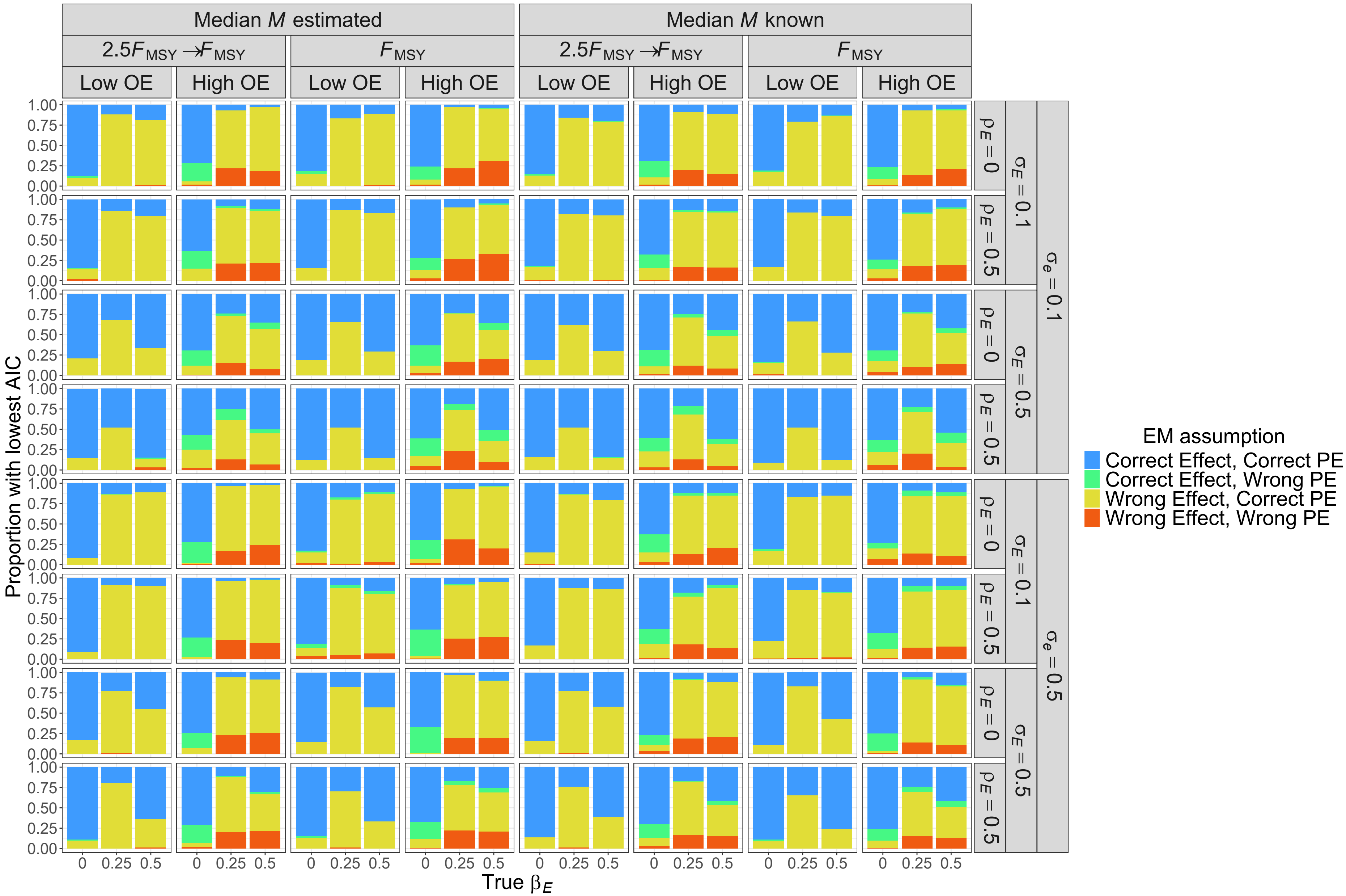

Fig. S12. Proportion of simulated data sets for R+S OMs where the EM type (treatment of environmental covariate and assumed process error type) had the lowest AIC.

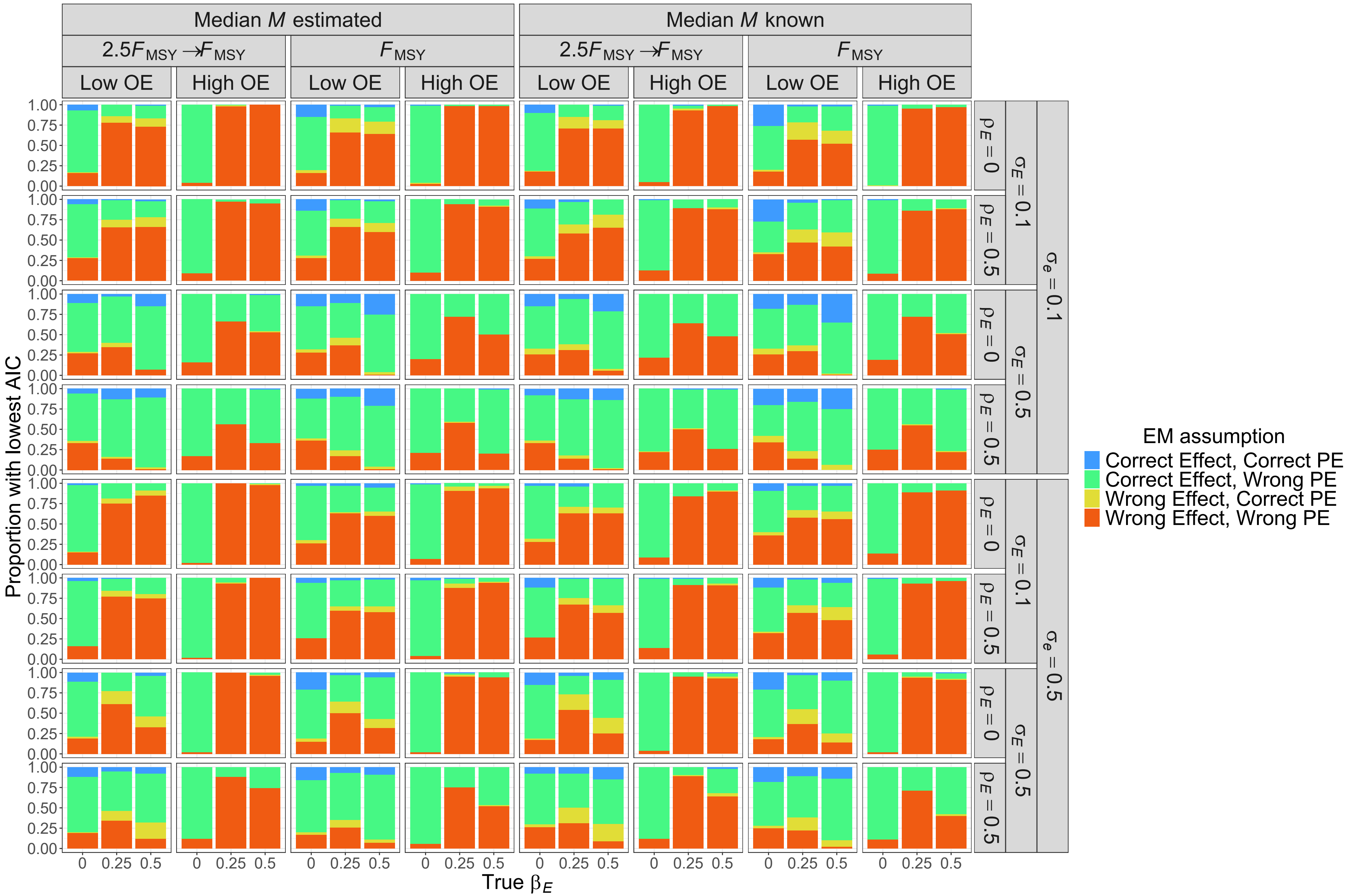

Fig. S13. Proportion of simulated data sets for R+M OMs where the EM type (treatment of environmental covariate and assumed process error type) had the lowest AIC.

Covariate effect bias

Table S6. For each OM process error source (columns), percent reduction in deviance for linear regression models fit to errors in estimation measured by Eq. **??** for the covariate effect on natural mortality (β_E) with each OM and EM factor (rows) included individually, combined, and with second and third order interactions. Includes results from all unconverged and converged models.

Factor	R	R+S	R+M
OM F history	<0.01	0.01	<0.01
OM Obs. Error	<0.01	<0.01	<0.01
OM σ_e	<0.01	<0.01	<0.01
OM σ_E	<0.01	<0.01	<0.01
OM ρ_E	<0.01	<0.01	<0.01
OM β_E	<0.01	0.01	<0.01
EM Convergence	<0.01	<0.01	<0.01
EM Process Error	<0.01	0.01	0.01
EM M assumption	<0.01	<0.01	<0.01
All factors	0.02	0.03	0.03
+ All Two Way	0.16	0.20	0.18
+ All Three Way	0.58	0.73	0.59

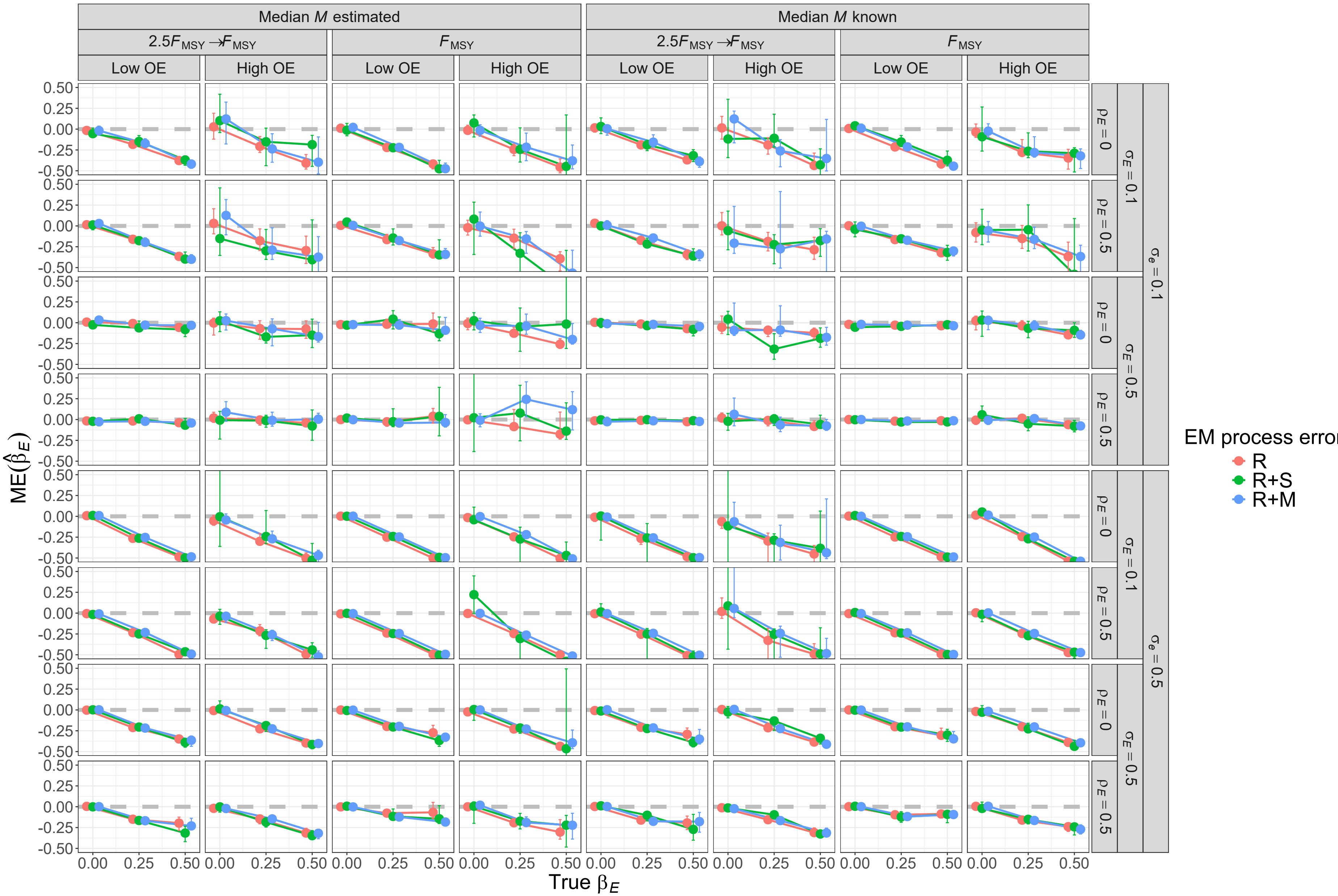

Fig. S14. For R OMs, median error (ME) of estimates of environmental effect on natural mortality β_E from fitting EMs with alternative process error assumptions and treatment of median natural mortality (e_M^{β} known or estimated). Vertical lines represent 95% confidence intervals.

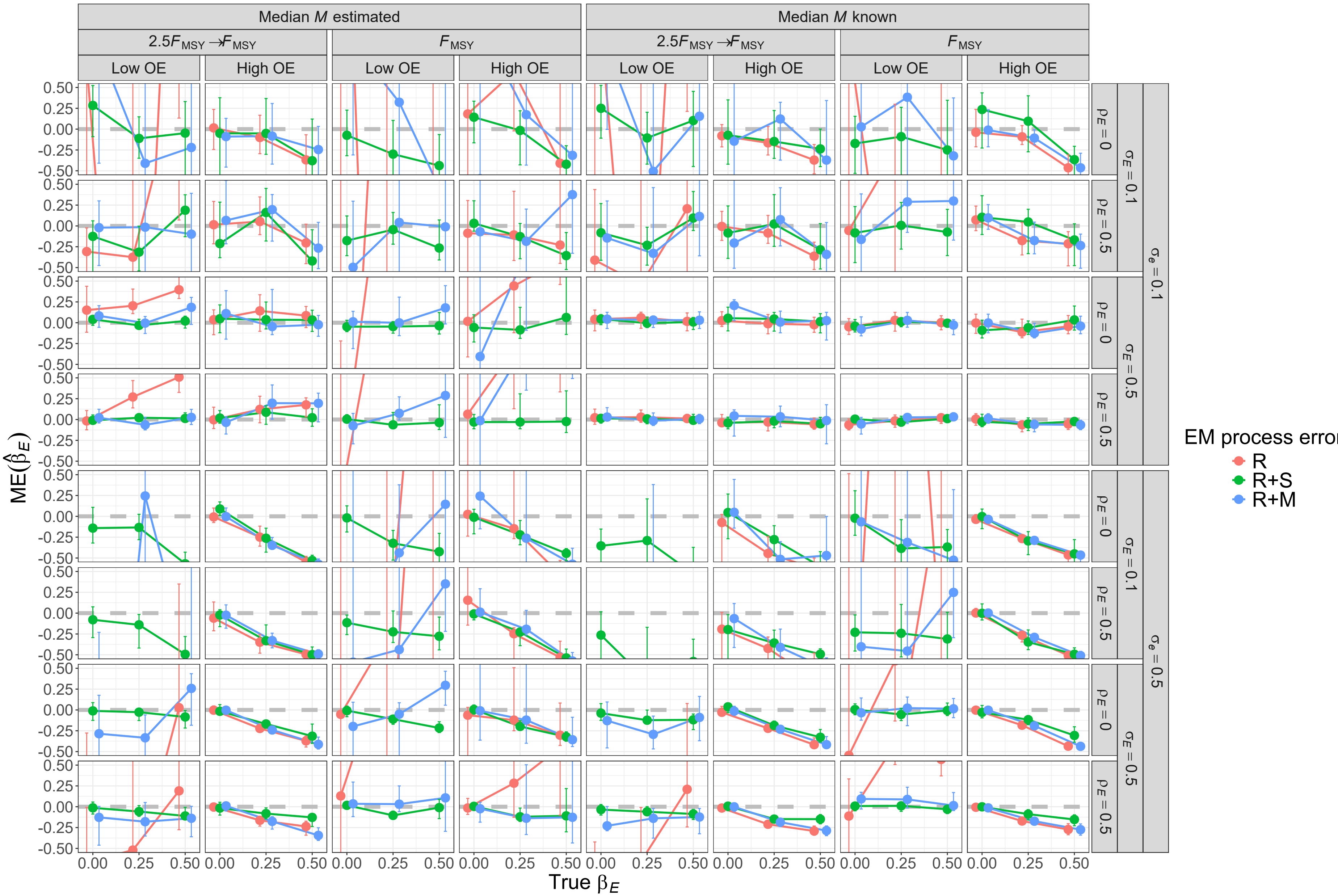

Fig. S15. For R+S OMs, median error (ME) of estimates of environmental effect on natural mortality β_E from fitting EMs with alternative process error assumptions and treatment of median natural mortality (e^{β}_{M} known or estimated). Vertical lines represent 95% confidence intervals.

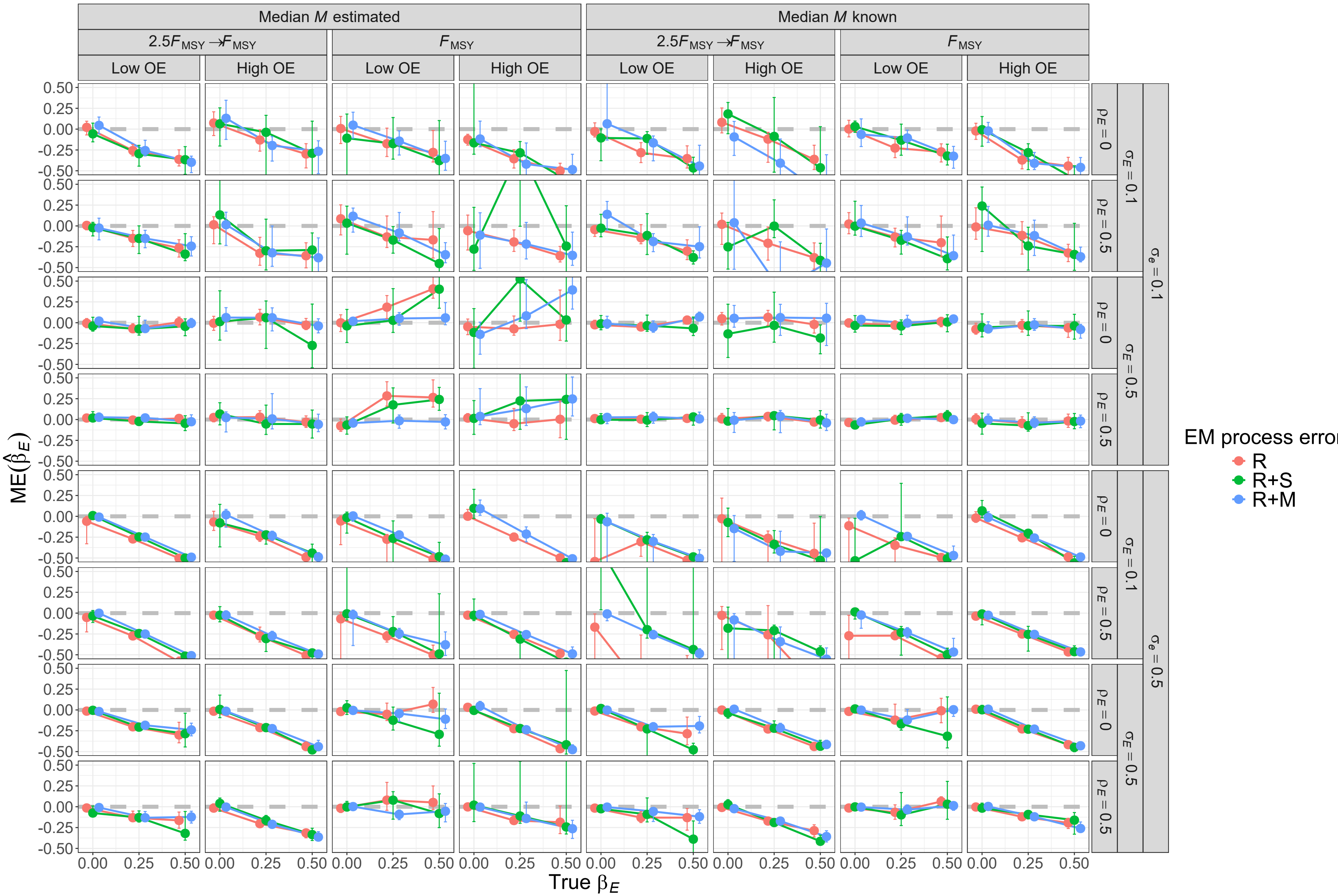

Fig. S16. For R+M OMs, median error (ME) of estimates of environmental effect on natural mortality β_E from fitting EMs with alternative process error assumptions and treatment of median natural mortality (e_M^{β} known or estimated). Vertical lines represent 95% confidence intervals.

Covariate effect standard error estimation bias

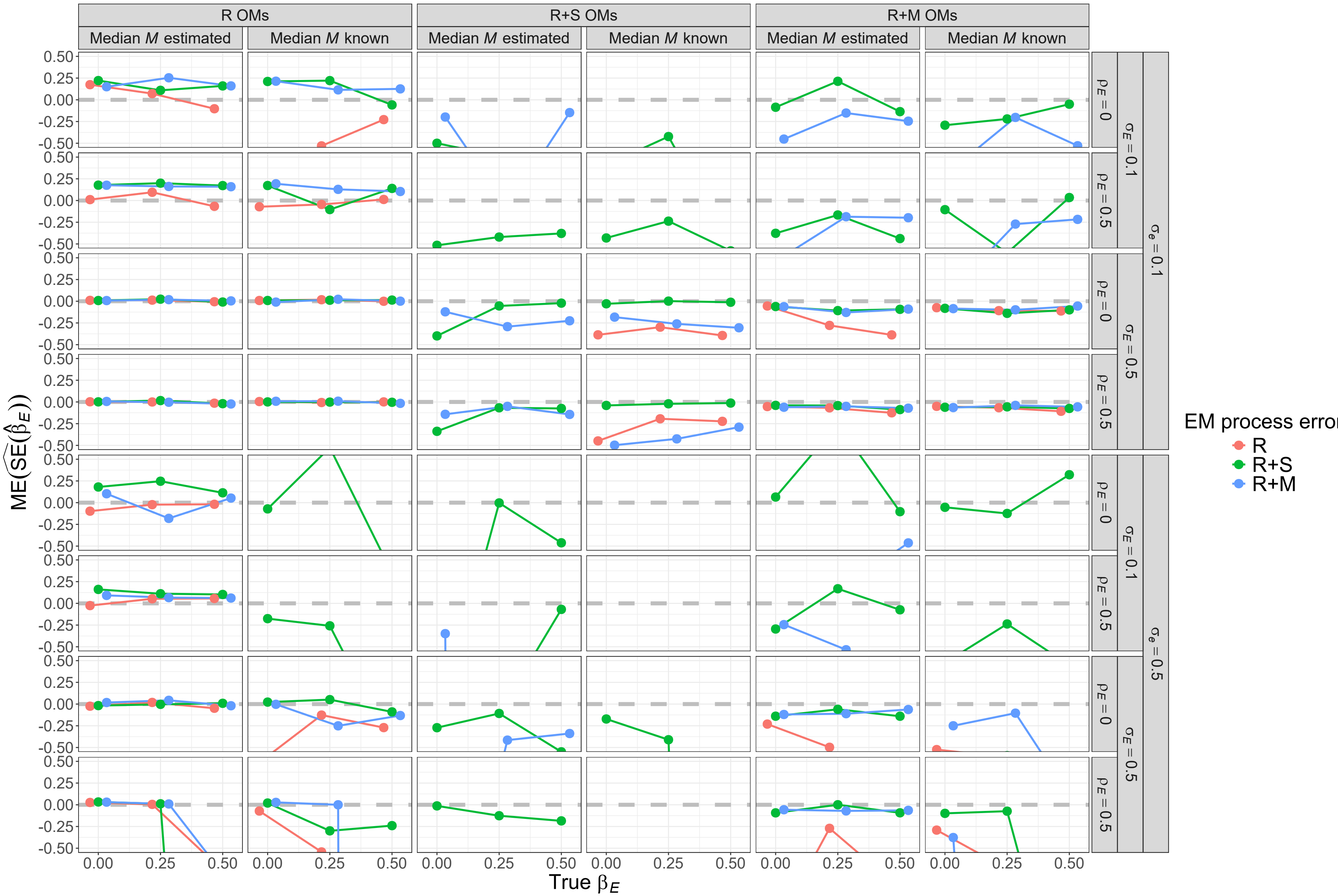

Fig. S17. Median error (ME) of Hessian-based estimates of standard error for covariate effect on natural mortality β_E from fitting EMs with alternative process error assumptions and treatment of median natural mortality (e^{β}_{M} known or estimated). All OMs had low observation error and contrast in fishing mortality. True standard error is defined as the mean of the standard error estimates accross converged fits to simulated data sets for a given OM scenario.

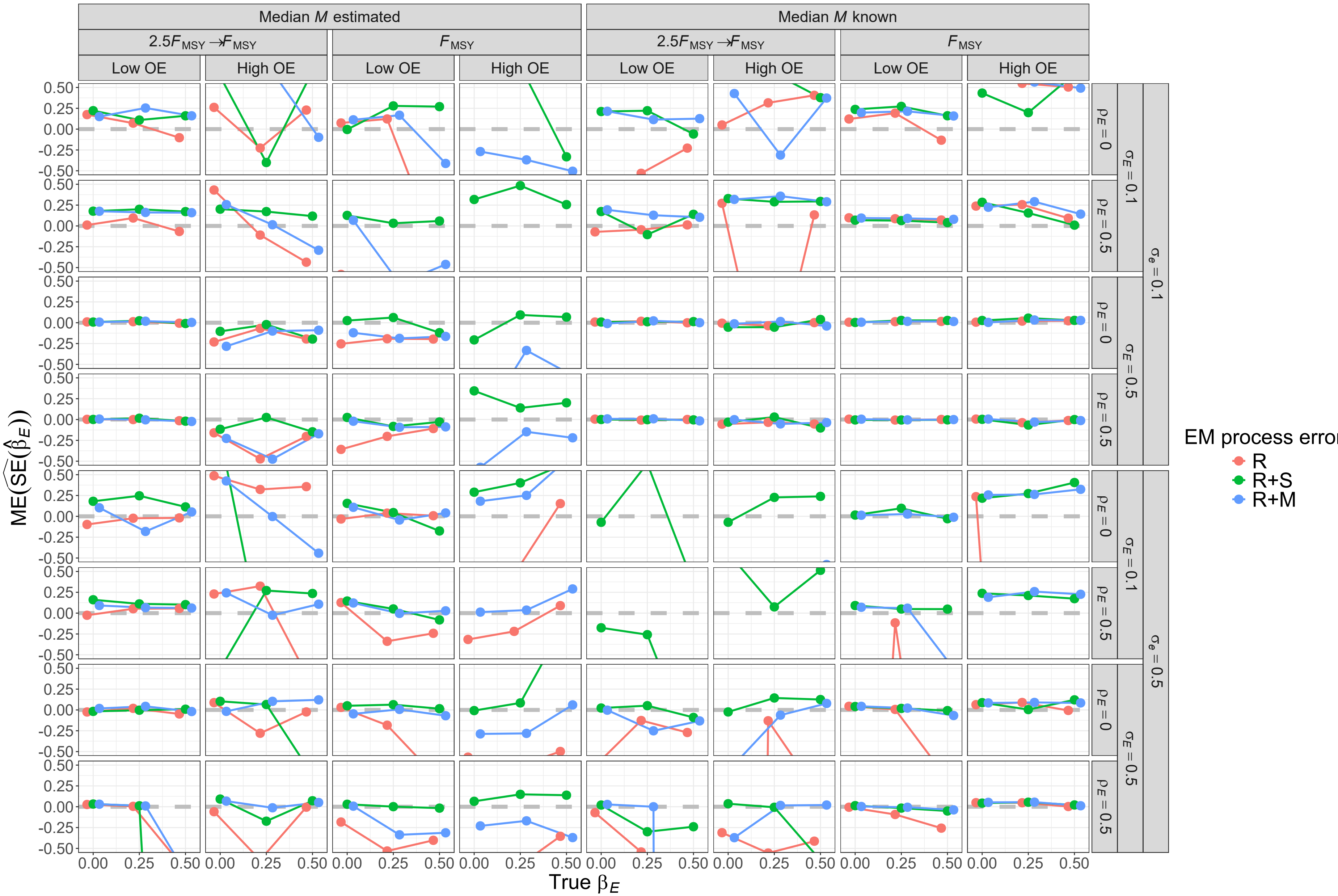

Fig. S18. For R OMs, median error (ME) of Hessian-based estimates of standard error for covariate effect on natural mortality β_E from fitting EMs with alternative process error assumptions and treatment of median natural mortality (e^{β}_{M} known or estimated). True standard error is defined as the mean of the standard error estimates accross converged fits to simulated data sets for a given OM scenario.

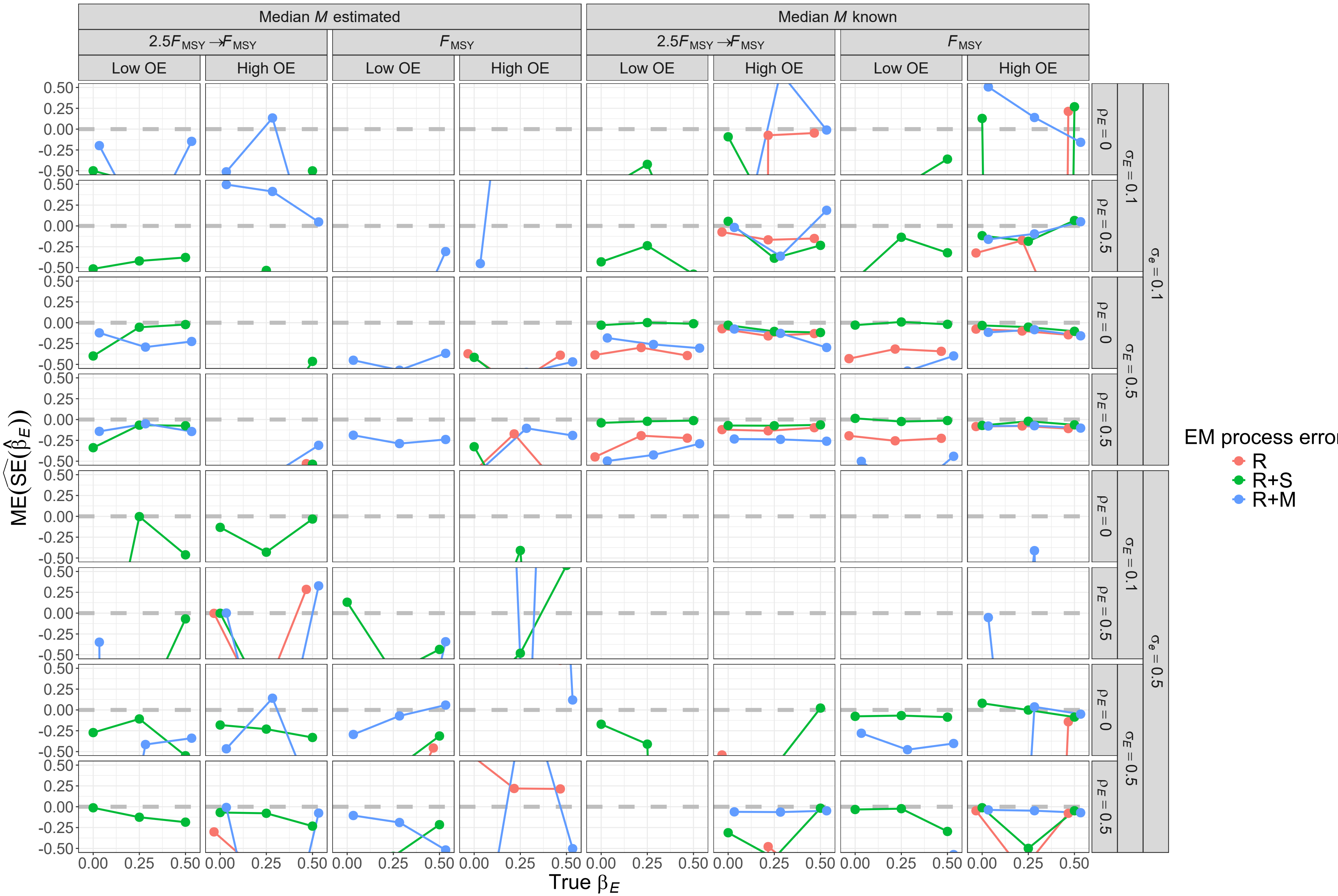

Fig. S19. For R+S OMs, median error (ME) of Hessian-based estimates of standard error for covariate effect on natural mortality β_E from fitting EMs with alternative process error assumptions and treatment of median natural mortality (e_M^{β} known or estimated). True standard error is defined as the mean of the standard error estimates accross converged fits to simulated data sets for a given OM scenario.

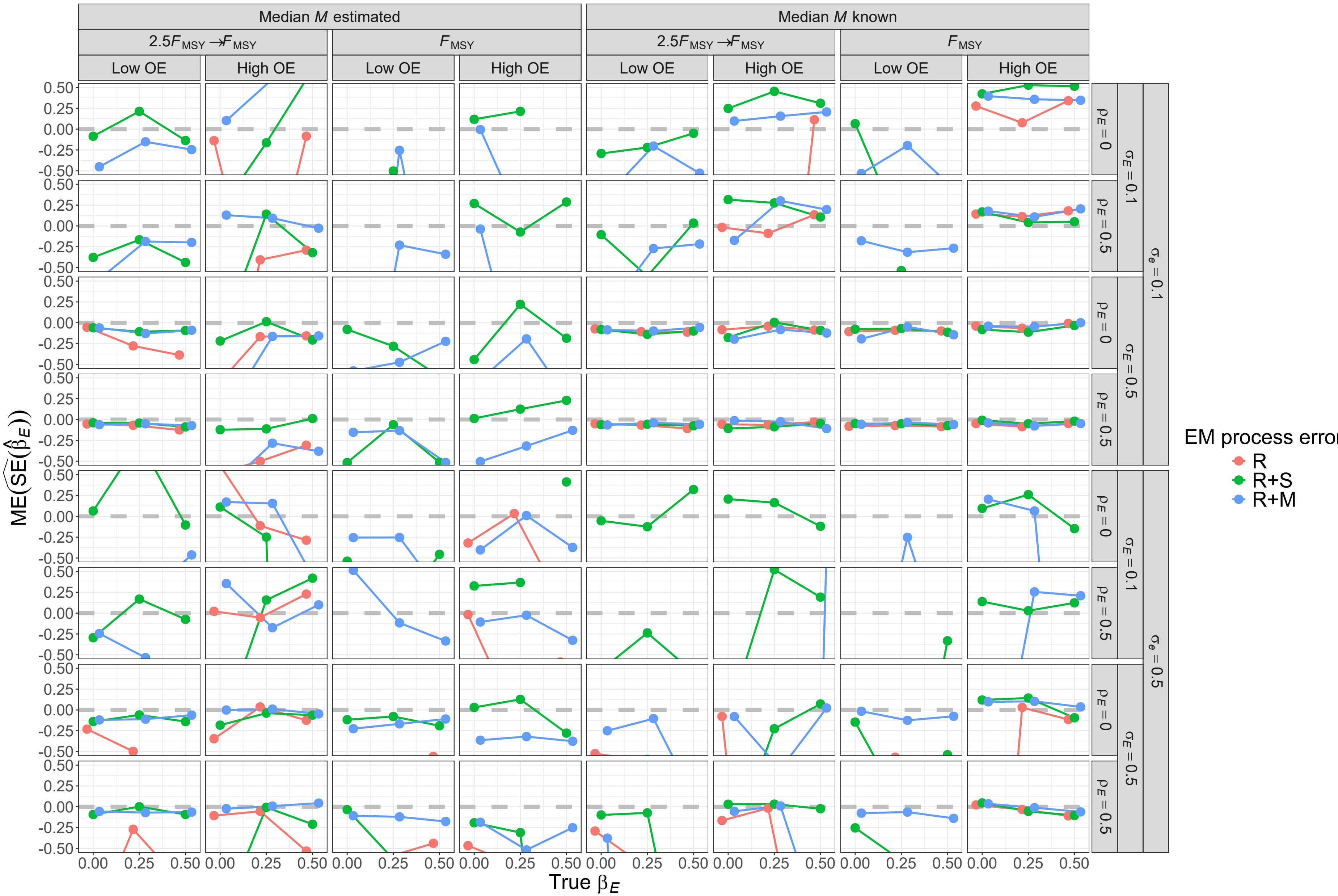

Fig. S20. For R+M OMs, median error (ME) of Hessian-based estimates of standard error for covariate effect on natural mortality β_E from fitting EMs with alternative process error assumptions and treatment of median natural mortality (e_M^{β} known or estimated). True standard error is defined as the mean of the standard error estimates accross converged fits to simulated data sets for a given OM scenario.

Covariate effect confidence interval coverage

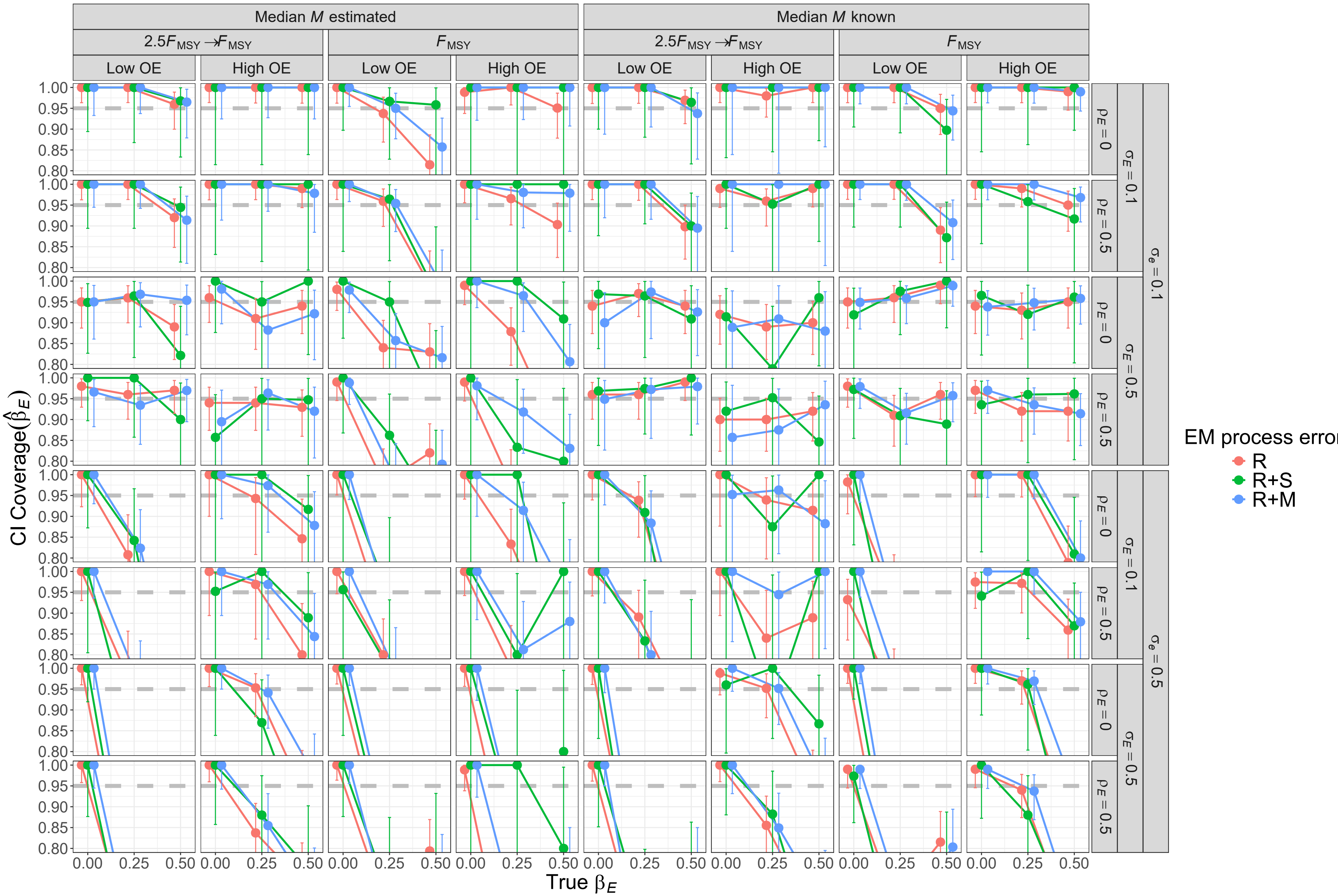

Fig. S21. For R OMs, probability of 95% confidence interval for β_E containing the true value for EMs with alternative process error assumptions and treatment of median natural mortality (e_M^{β} known or estimated). Vertical lines represent 95% confidence intervals.

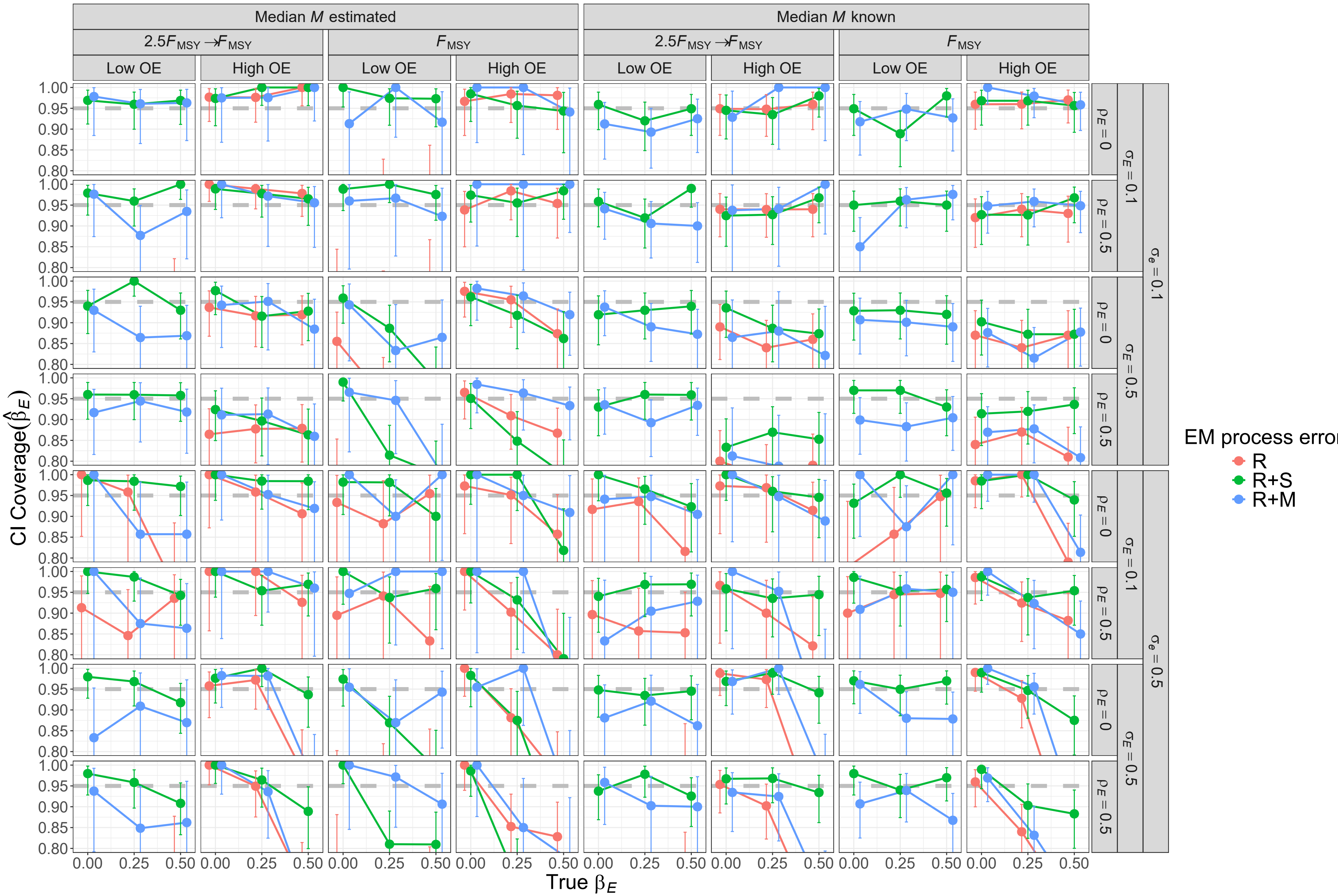

Fig. S22. For R+S OMs, probability of 95% confidence interval for β_E containing the true value for EMs with alternative process error assumptions and treatment of median natural mortality (e_M^{β} known or estimated). Vertical lines represent 95% confidence intervals.

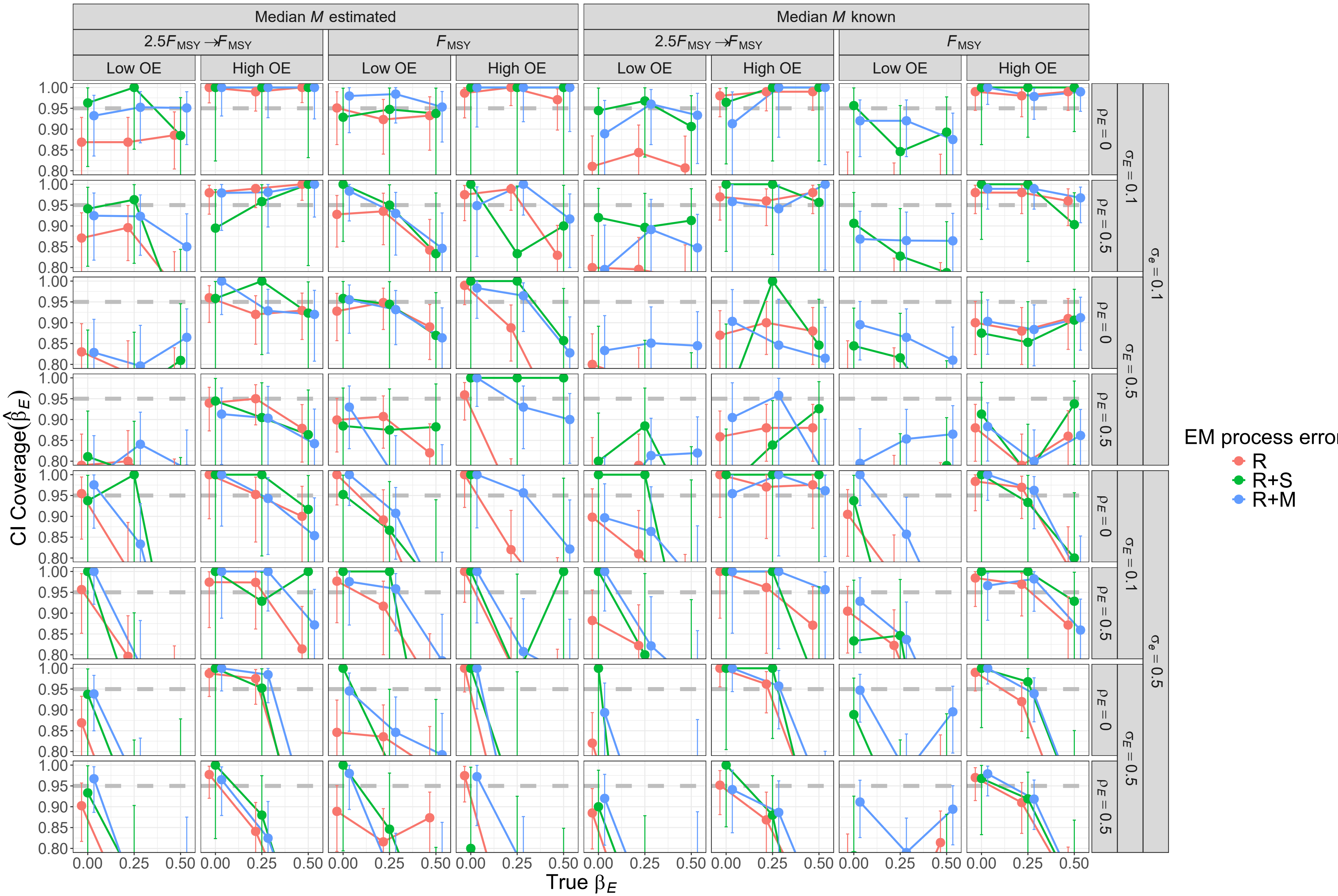

Fig. S23. For R+M OMs, probability of 95% confidence interval for β_E containing the true value for EMs with alternative process error assumptions and treatment of median natural mortality (e_M^{β} known or estimated). Vertical lines represent 95% confidence intervals.

Covariate effect RMSE

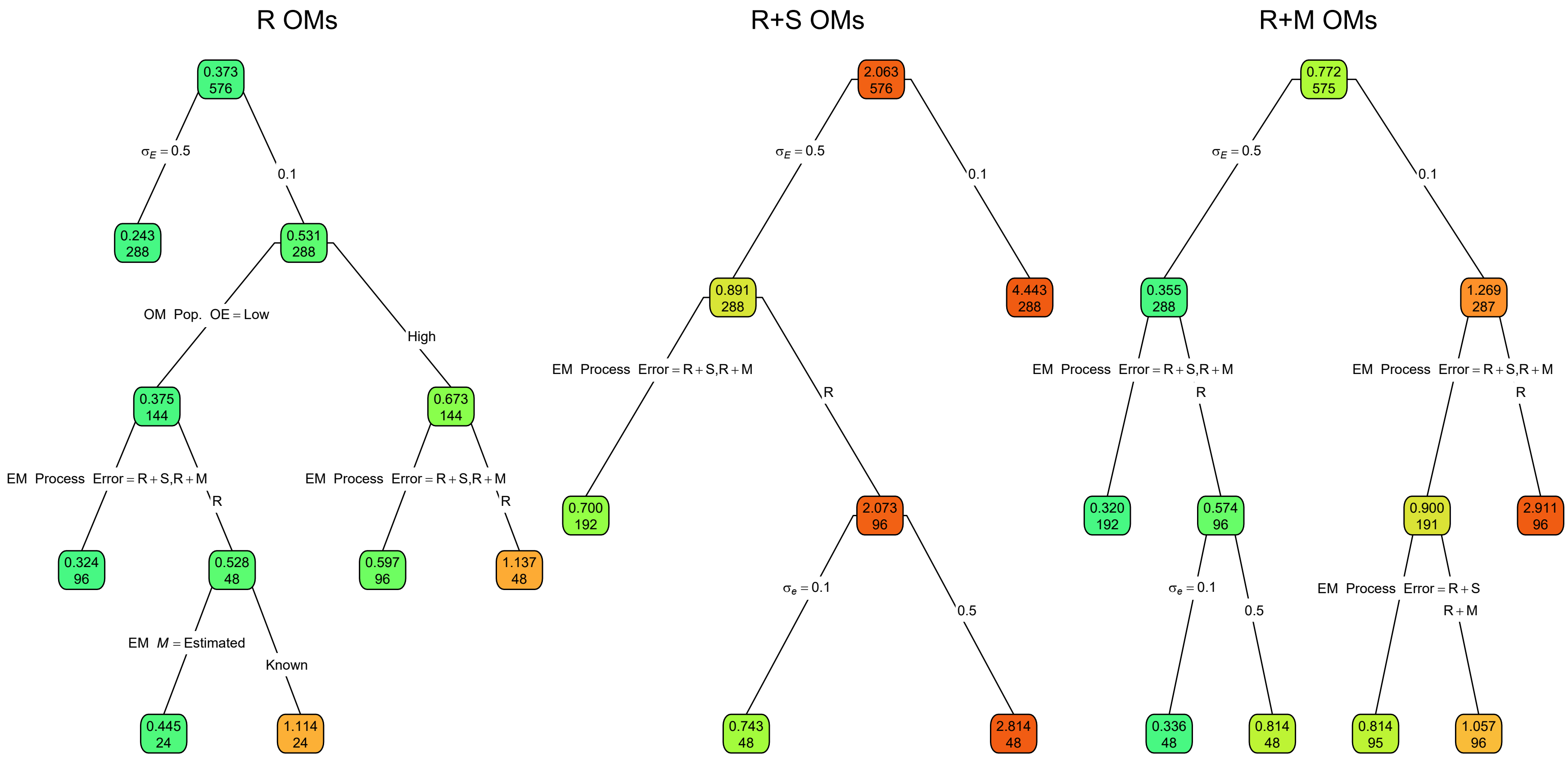

Fig. S24. Regression trees indicating primary factors determining reductions in sums of squares of errors in estimation measured by Eq. ?? for the RMSE of covariate effect on natural mortality (RMSE $\left(\widehat{\beta}_E\right)$) for R, R+S, and R+M OMs. Each node shows the median error (top) and number of observations (bottom) for the corresponding subset. Lower or higher median absolute errors of the process error assumption are indicated by more green or red polygons, respectively.

Table S7. For each OM process error source (columns), percent reduction in deviance for linear regression models fit to errors in estimation measured by Eq. **??** for the the RMSE for estimates of the covariate effect on natural mortality (RMSE $\left(\widehat{\beta}_E\right)$) with each OM and EM factor (rows) included individually, combined, and with second and third order interactions.

Factor	R	R+S	R+M
OM F history	0.50	0.11	0.12
OM Obs. Error	12.04	5.66	0.36
$OM \sigma_e$	0.87	5.78	2.60
$OM \sigma_E$	15.85	24.18	24.26
$OM \rho_E$	0.36	0.82	0.34
OM β_E	6.72	0.07	0.96
EM Process Error	6.49	11.31	11.98
EM M assumption	0.04	0.64	0.07
All factors	42.87	48.58	40.65
+ All Two Way	67.34	74.09	61.74
+ All Three Way	79.42	82.75	71.92

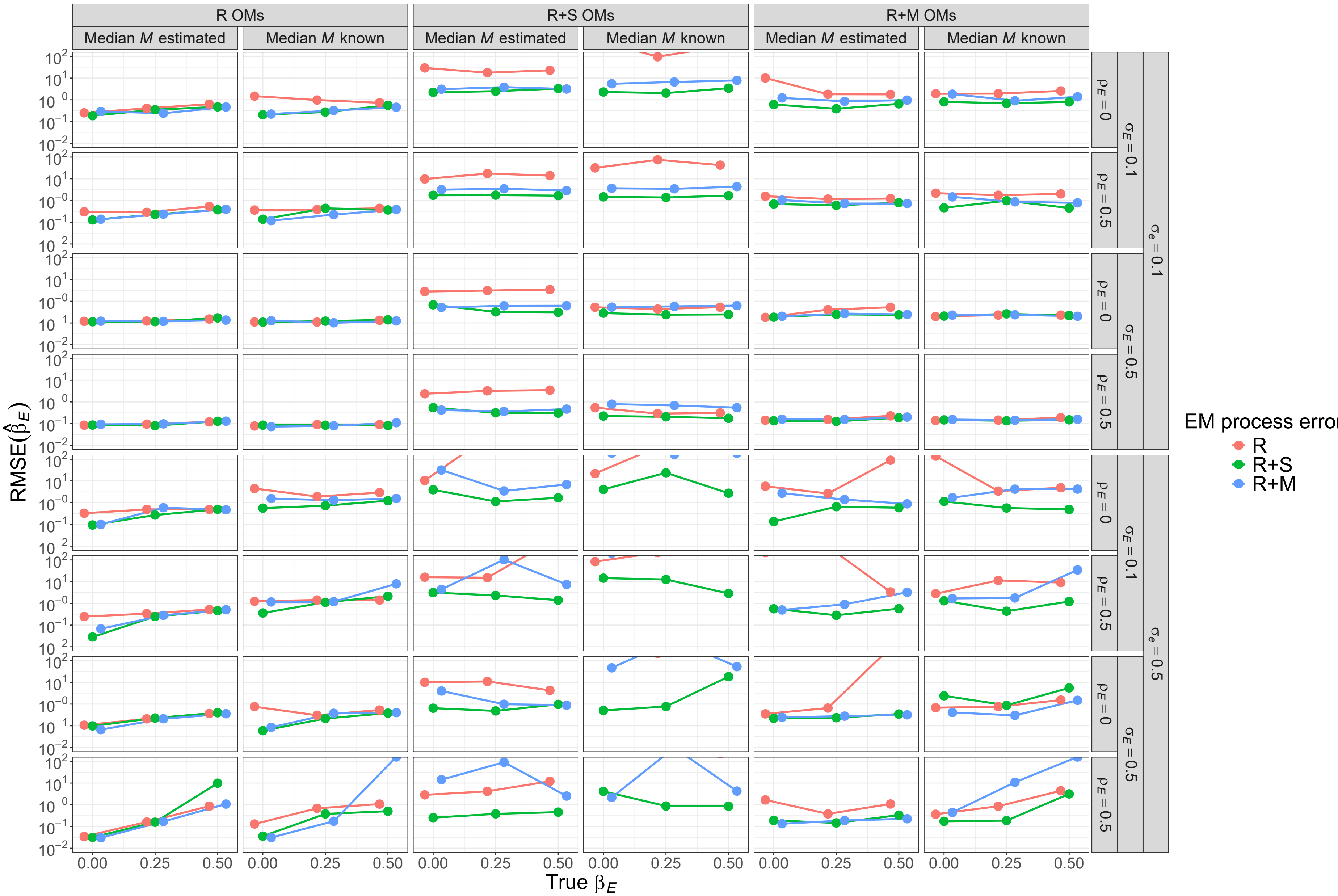

Fig. S25. Root mean square error (RMSE) of estimates of covariate effect on natural mortality β_E from fitting EMs with alternative process error assumptions and treatment of median natural mortality (e^{β}_{M} known or estimated). All OMs had low observation error and contrast in fishing mortality.

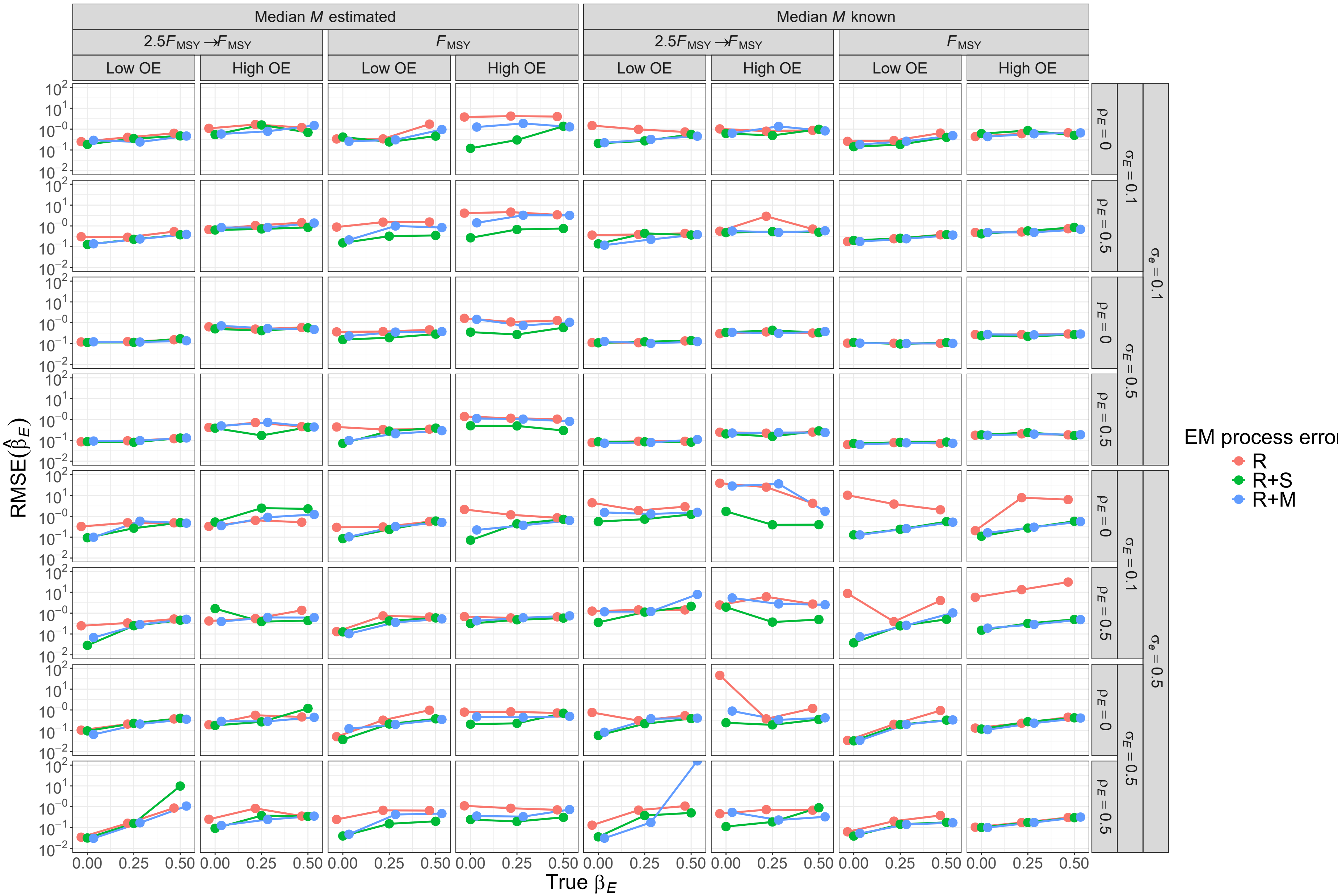

Fig. S26. For R OMs, root mean square error (RMSE) of estimates of covariate effect on natural mortality β_E from fitting EMs with alternative process error assumptions and treatment of median natural mortality (e^{β}_{M} known or estimated).

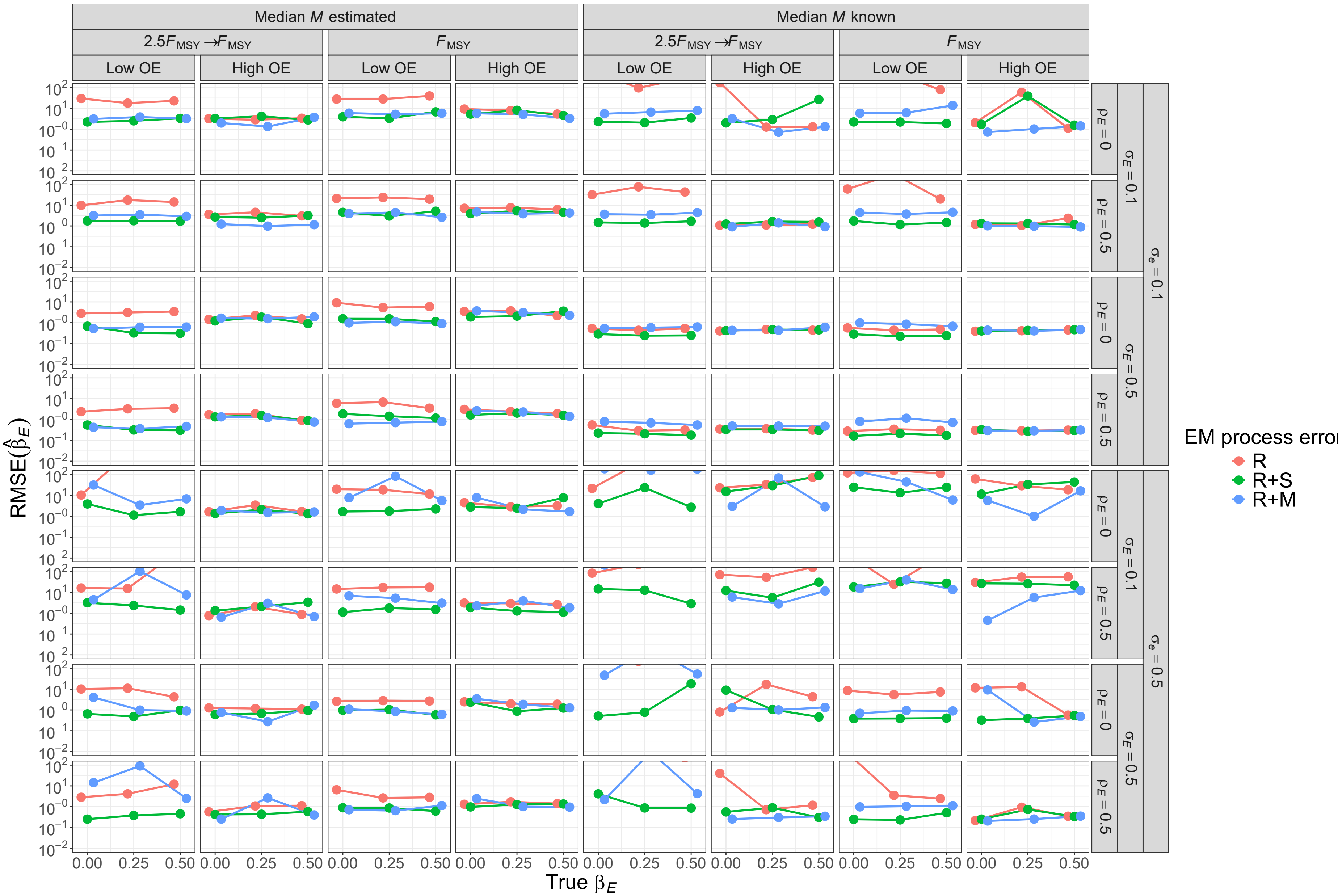

Fig. S27. For R+S OMs, root mean square error (RMSE) of estimates of covariate effect on natural mortality β_E from fitting EMs with alternative process error assumptions and treatment of median natural mortality (e^{β}_{M} known or estimated).

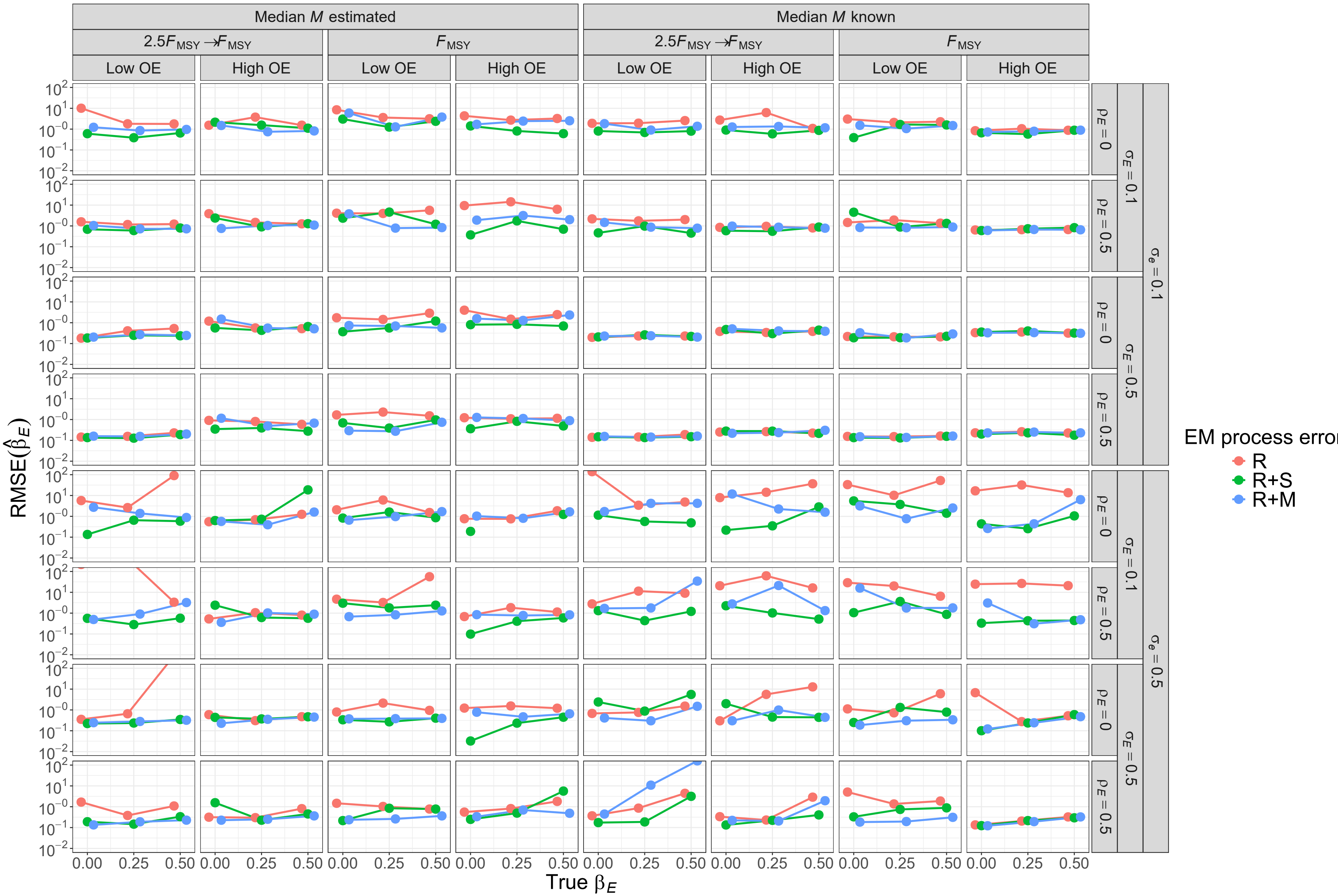

Fig. S28. For R+M OMs, root mean square error (RMSE) of estimates of covariate effect on natural mortality β_E from fitting EMs with alternative process error assumptions and treatment of median natural mortality (e^{β}_{M} known or estimated).

Covariate effect estimate and standard error example

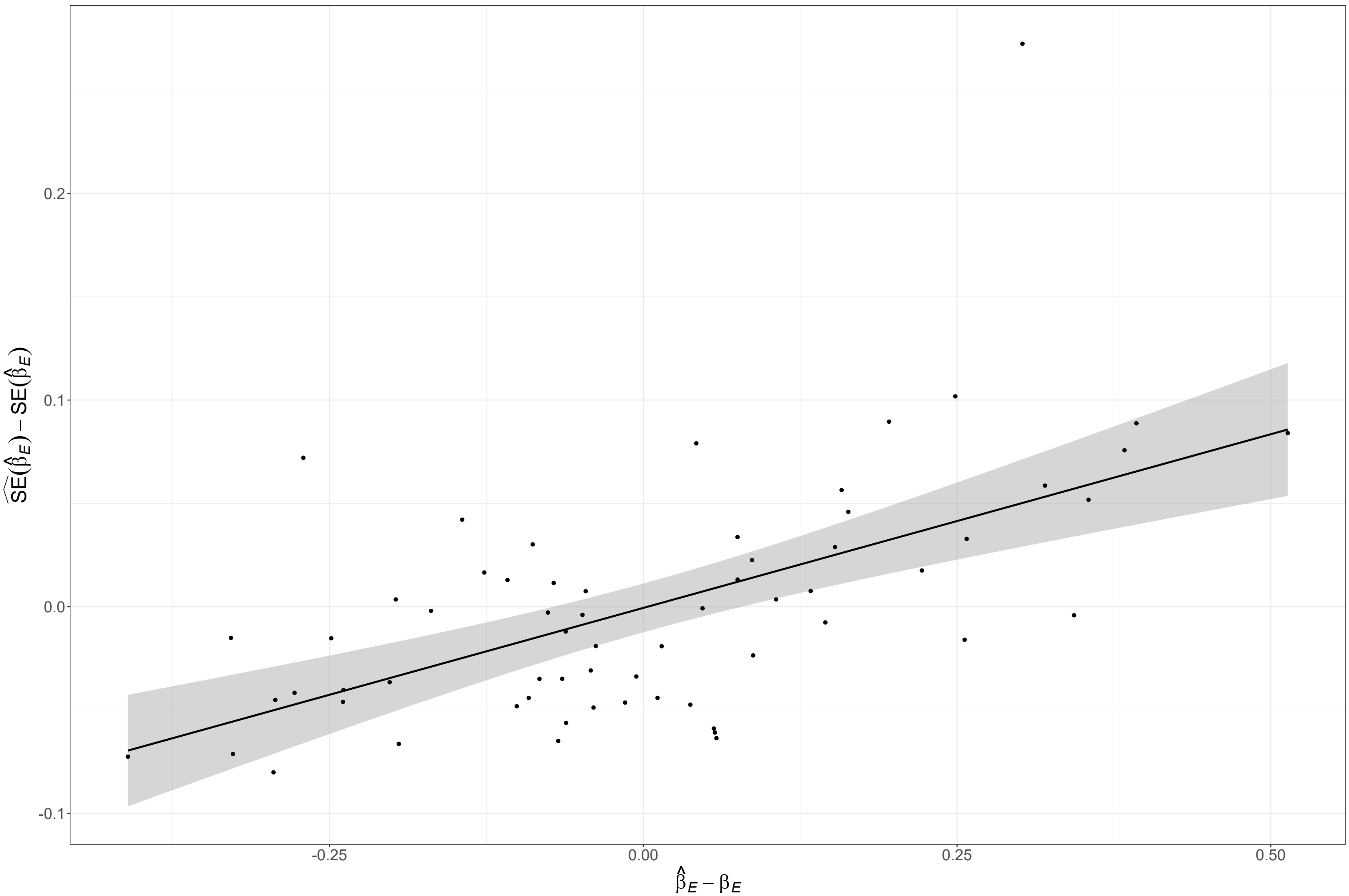

Fig. S29. Positive correlation of covariate effect estimates and Hessian-based standard error estimates for EM that also estimates the median natural mortality parameter and has correct R+M process error assumption fitted to simulated data from the OM with R+M process errors, temporal contrast in fishing pressure, low observation uncertainty for both population ($LowOE$) and covariate observations ($\sigma_e = 0.1$), high and uncorrelated temporal variability in the true covariate ($\sigma_E = 0.5$ and $\rho_E = 0$), and the strongest covariate effect on natural mortality ($\beta_E = 0.5$).

Median Natural mortality parameter bias

Table S8. For each OM process error source (columns), percent reduction in deviance for linear regression models fit to errors in estimation measured by Eq. **??** for the median natural mortality rate parameter with each OM and EM factor (rows) included individually, combined, and with second and third order interactions. Includes results from all unconverged and converged models.

Factor	R	R+S	R+M
OM F history	<0.01	<0.01	<0.01
OM Obs. Error	<0.01	<0.01	<0.01
OM σ_e	<0.01	<0.01	<0.01
OM σ_E	<0.01	<0.01	<0.01
OM ρ_E	<0.01	<0.01	<0.01
OM β_E	<0.01	<0.01	0.01
EM Convergence	<0.01	0.01	0.01
EM Process Error	<0.01	0.01	0.01
EM β_E assumption	<0.01	<0.01	<0.01
All factors	0.02	0.02	0.05
+ All Two Way	0.15	0.15	0.24
+ All Three Way	0.57	0.54	0.73

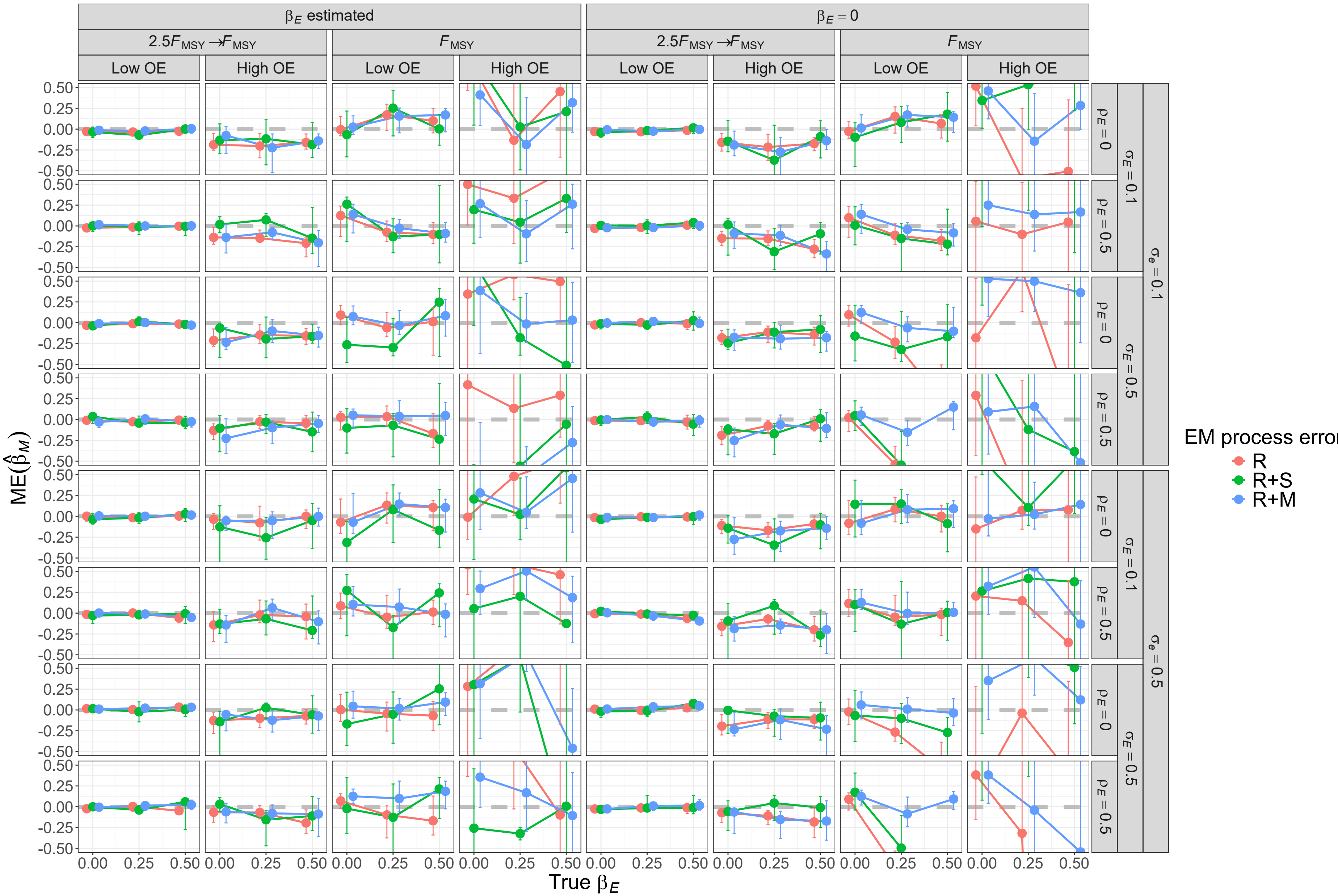

Fig. S30. For R OMs, median error (ME) of estimates of β_M from fitting EMs with alternative process error assumptions and treatment of covariate effect ($\beta_E = 0$ or estimated). Vertical lines represent 95% confidence intervals.

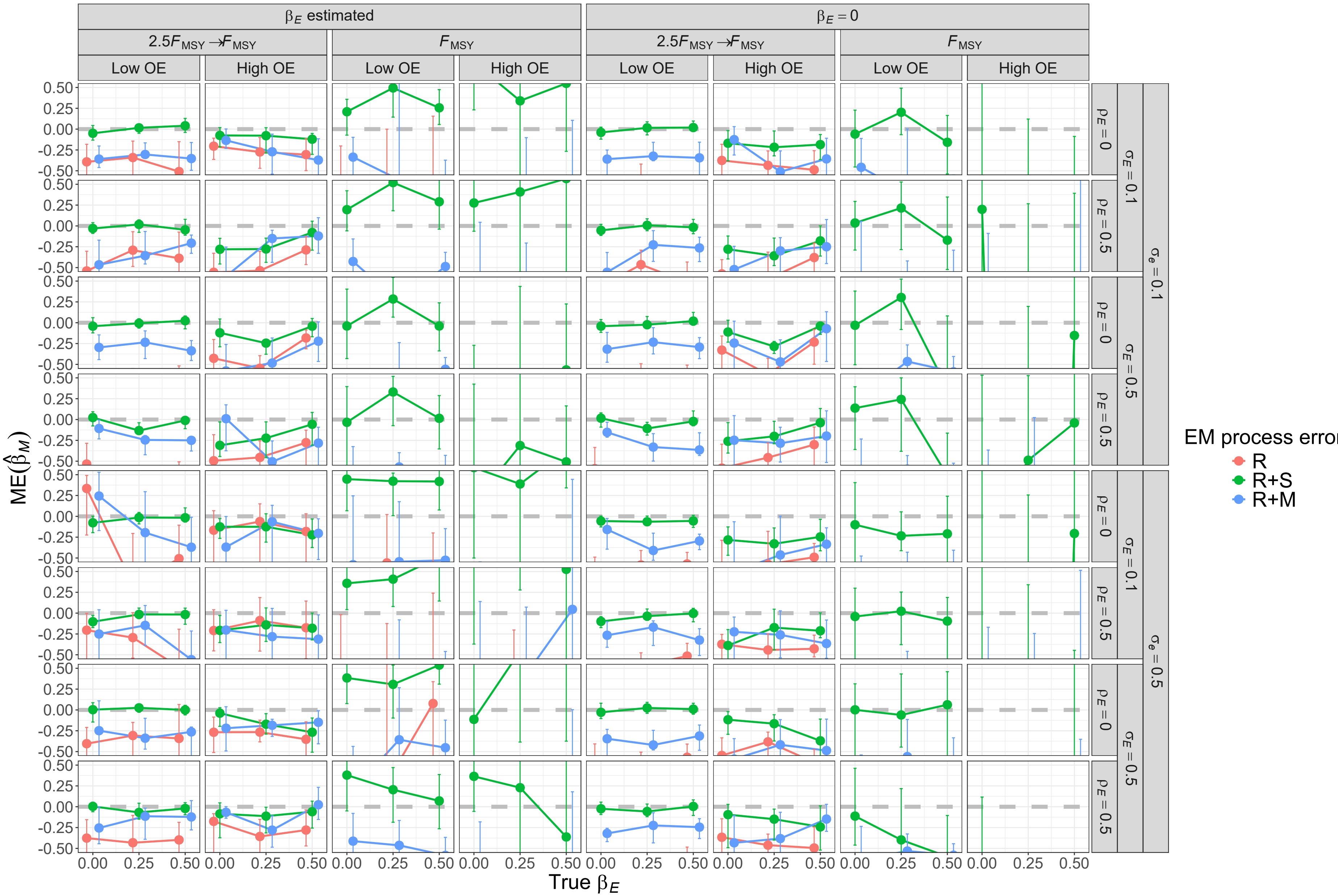

Fig. S31. For R+S OMs, median error (ME) of estimates of β_M from fitting EMs with alternative process error assumptions and treatment of covariate effect ($\beta_E = 0$ or estimated). Vertical lines represent 95% confidence intervals.

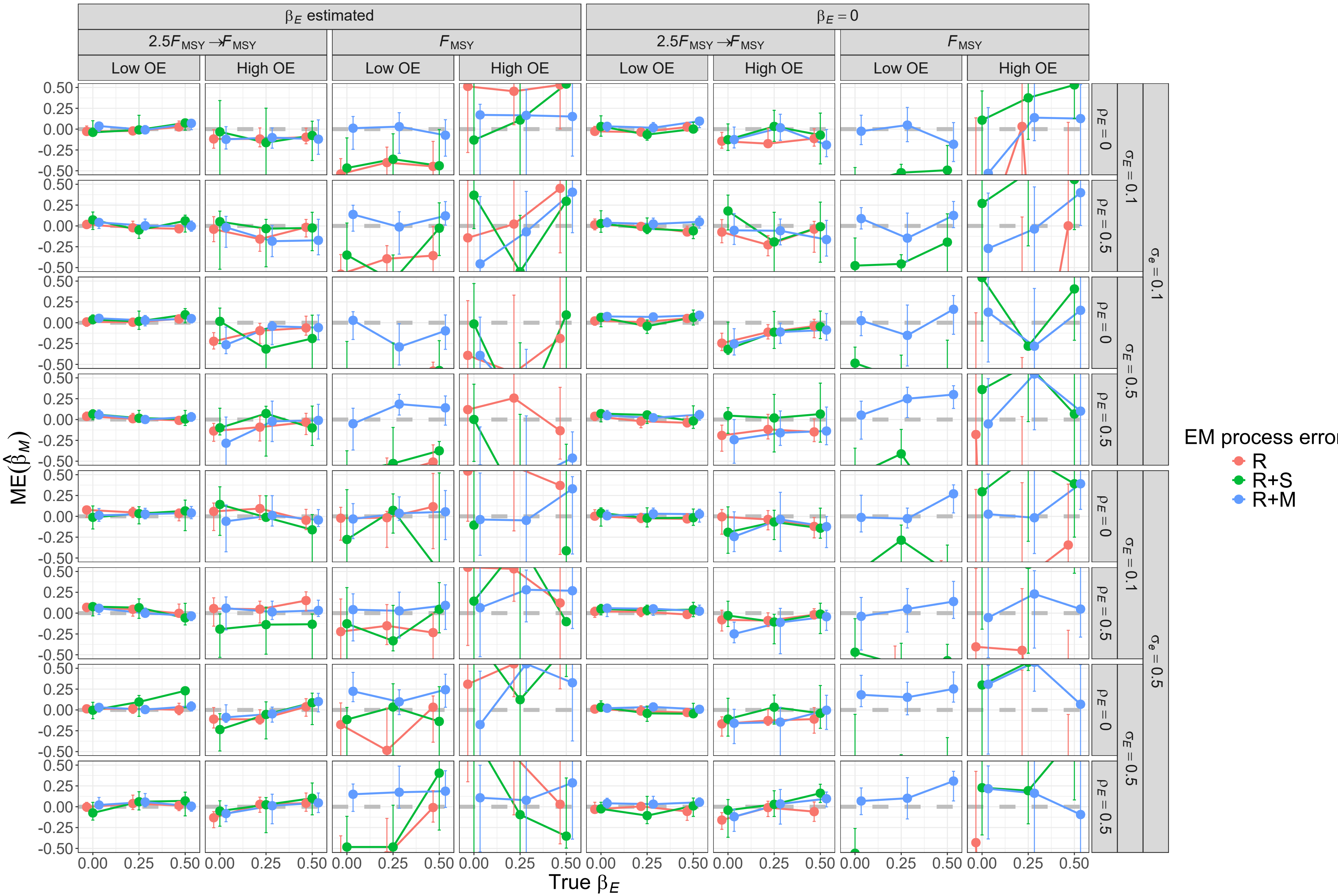

Fig. S32. For R+M OMs, median error (ME) of estimates of β_M from fitting EMs with alternative process error assumptions and treatment of covariate effect ($\beta_E = 0$ or estimated). Vertical lines represent 95% confidence intervals.

Median natural mortality parameter standard error estimation bias

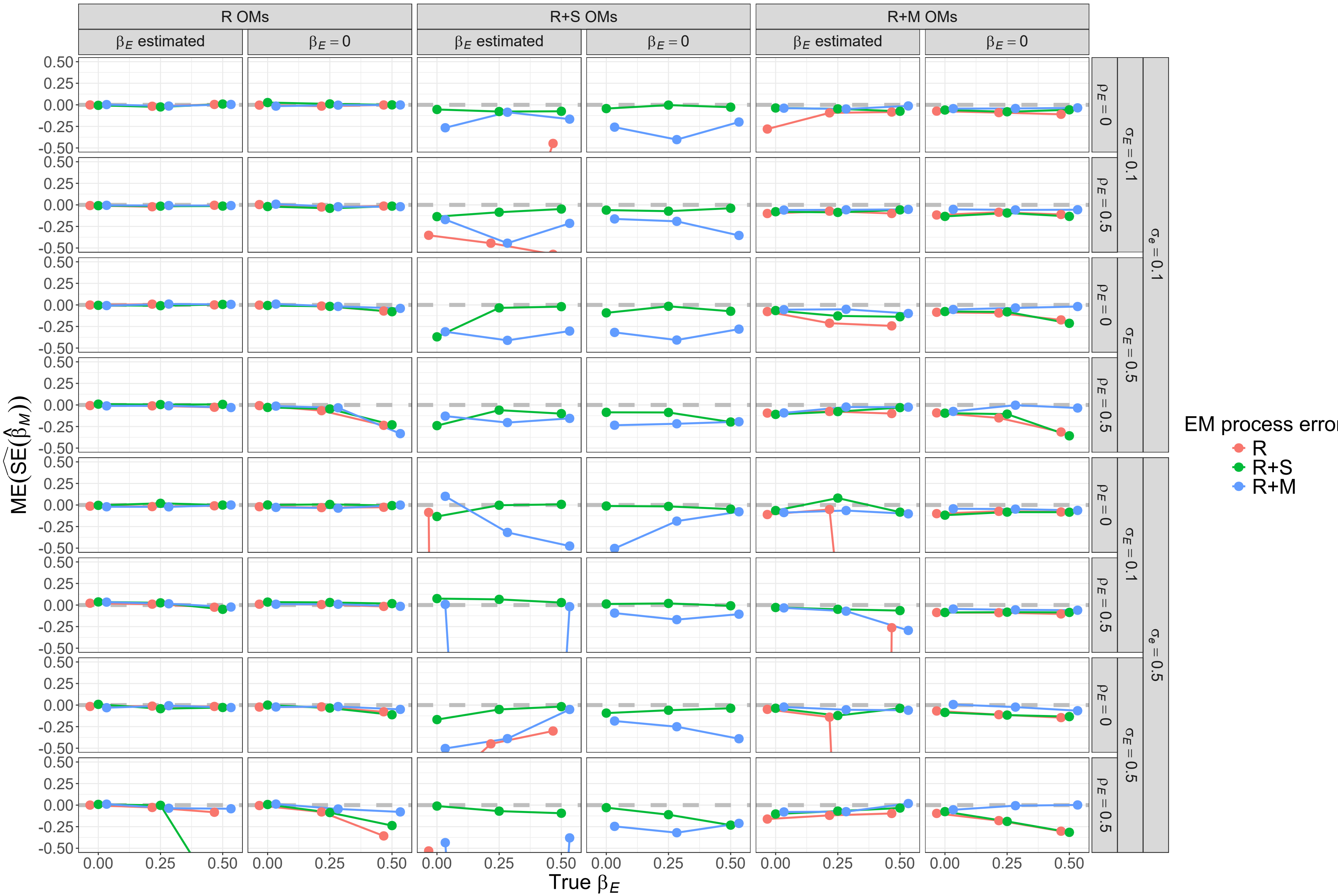

Fig. S33. Median error (ME) of Hessian-based estimates of standard error for median natural mortality parameter β_M from fitting EMs with alternative process error assumptions and treatment of the covariate effect ($\beta_E = 0$ or estimated). All OMs had low observation error and contrast in fishing mortality. True standard error is defined as the mean of the standard error estimates accross converged fits to simulated data sets for a given OM scenario.

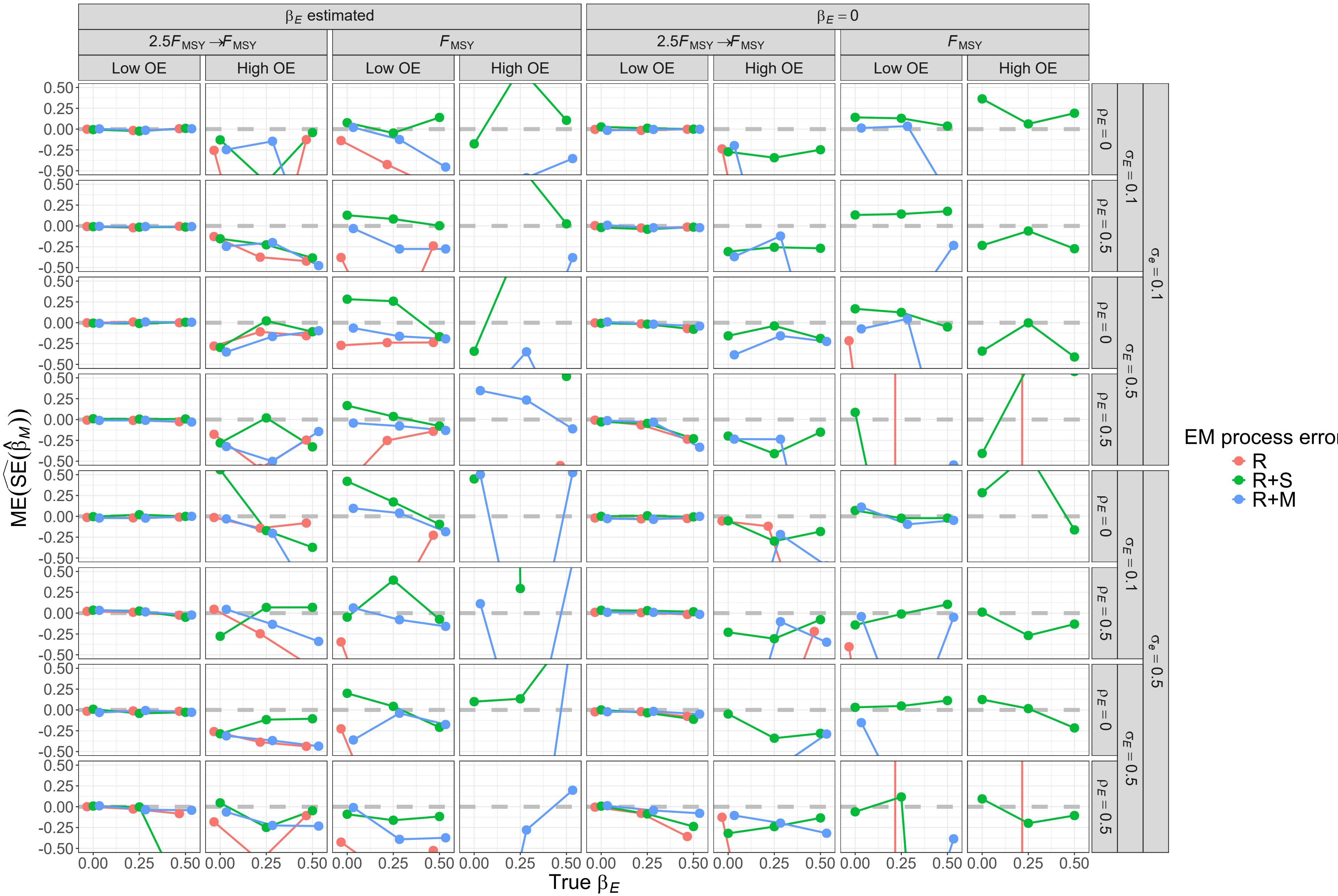

Fig. S34. For R OMs, median error (ME) of Hessian-based estimates of standard error for median natural mortality parameter β_M from fitting EMs with alternative process error assumptions and treatment of the covariate effect ($\beta_E = 0$ or estimated). True standard error is defined as the mean of the standard error estimates accross converged fits to simulated data sets for a given OM scenario.

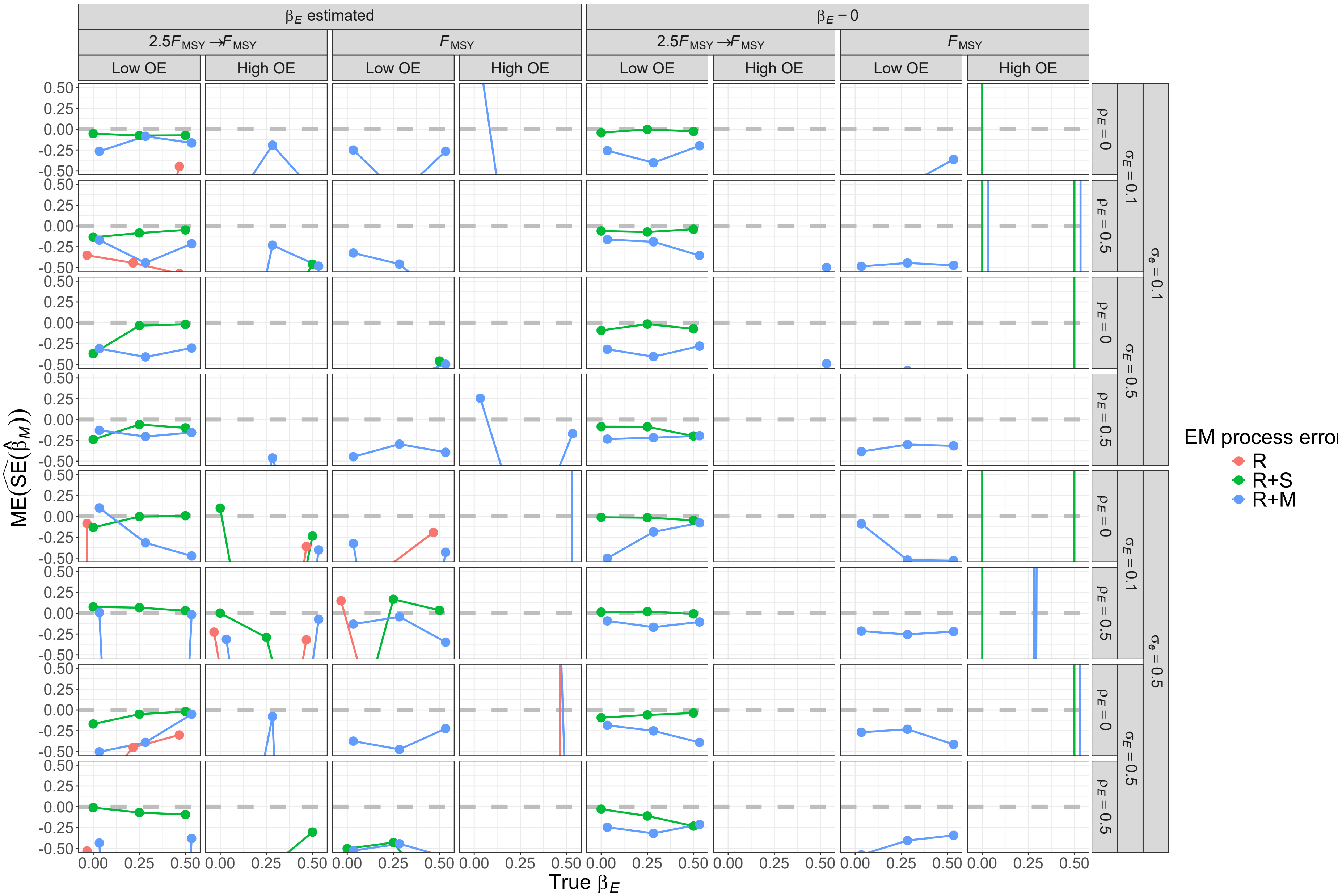

Fig. S35. For R+S OMs, median error (ME) of Hessian-based estimates of standard error for median natural mortality parameter β_M from fitting EMs with alternative process error assumptions and treatment of the covariate effect ($\beta_E = 0$ or estimated). True standard error is defined as the mean of the standard error estimates accross converged fits to simulated data sets for a given OM scenario.

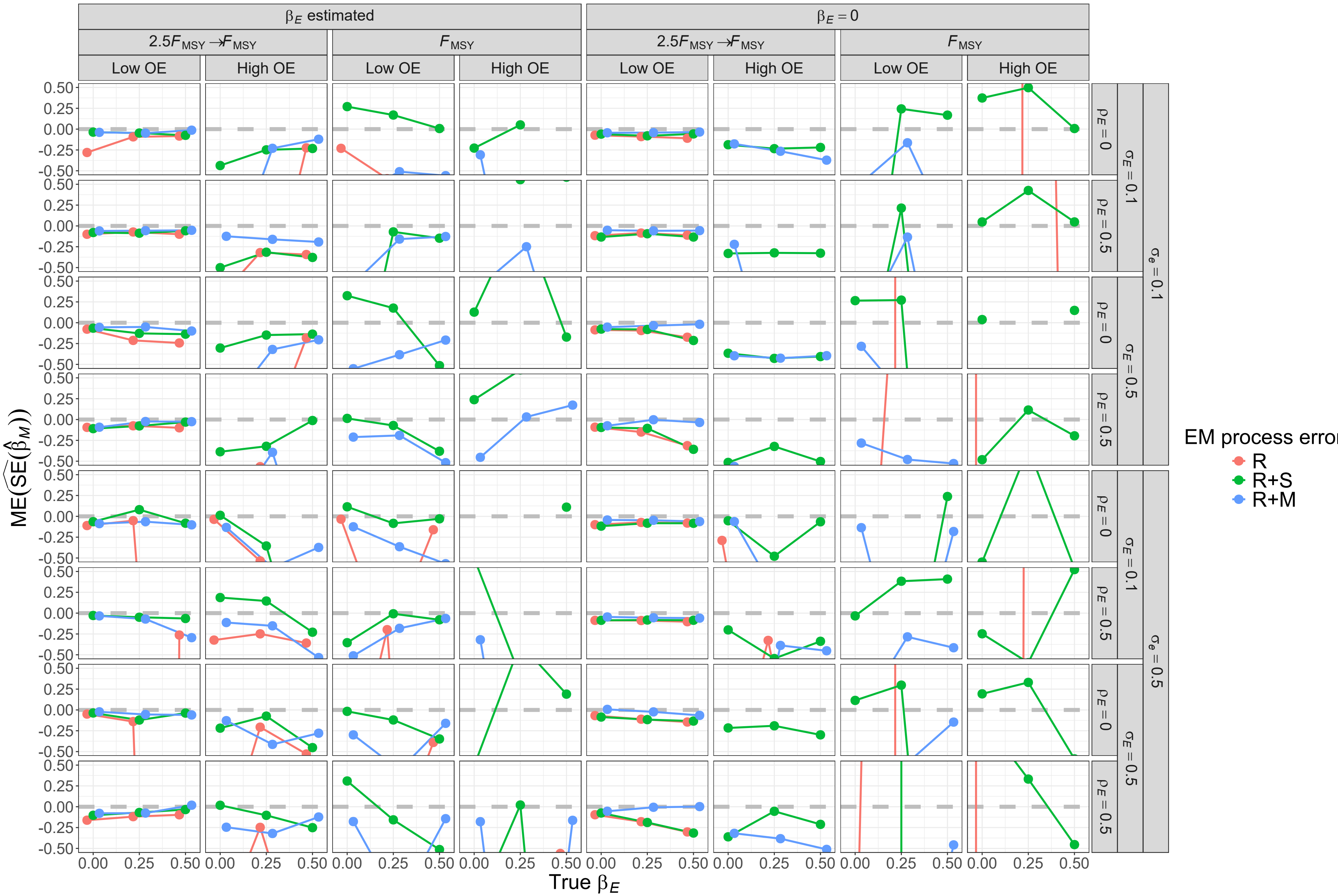

Fig. S36. For R+M OMs, median error (ME) of Hessian-based estimates of standard error for median natural mortality parameter β_M from fitting EMs with alternative process error assumptions and treatment of the covariate effect ($\beta_E = 0$ or estimated). True standard error is defined as the mean of the standard error estimates accross converged fits to simulated data sets for a given OM scenario.

Median Natural mortality parameter confidence interval coverage

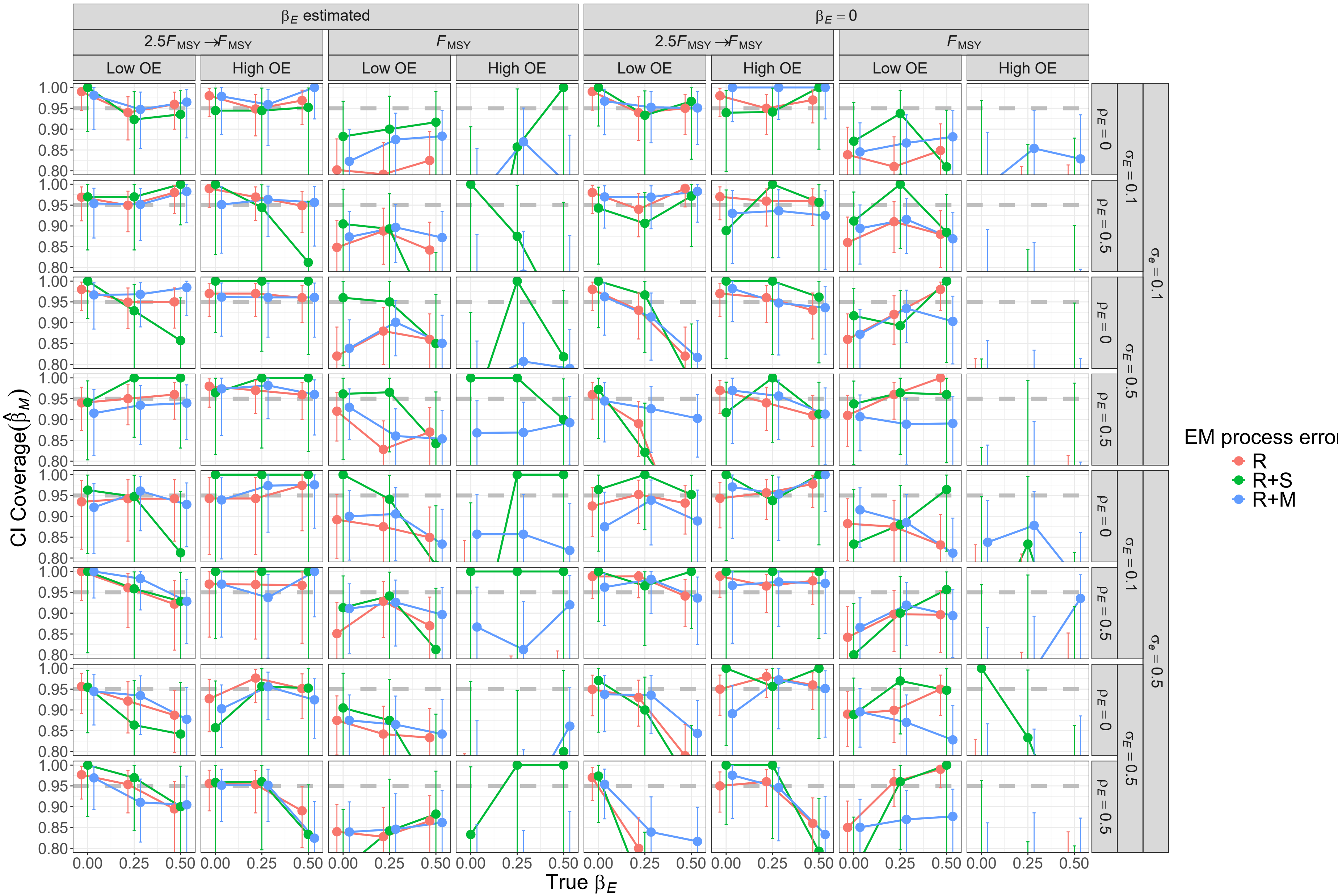

Fig. S37. For R OMs, probability of 95% confidence interval for β_M containing the true value for EMs with alternative process error assumptions and treatment of covariate effect ($\beta_E = 0$ or estimated). Vertical lines represent 95% confidence intervals.

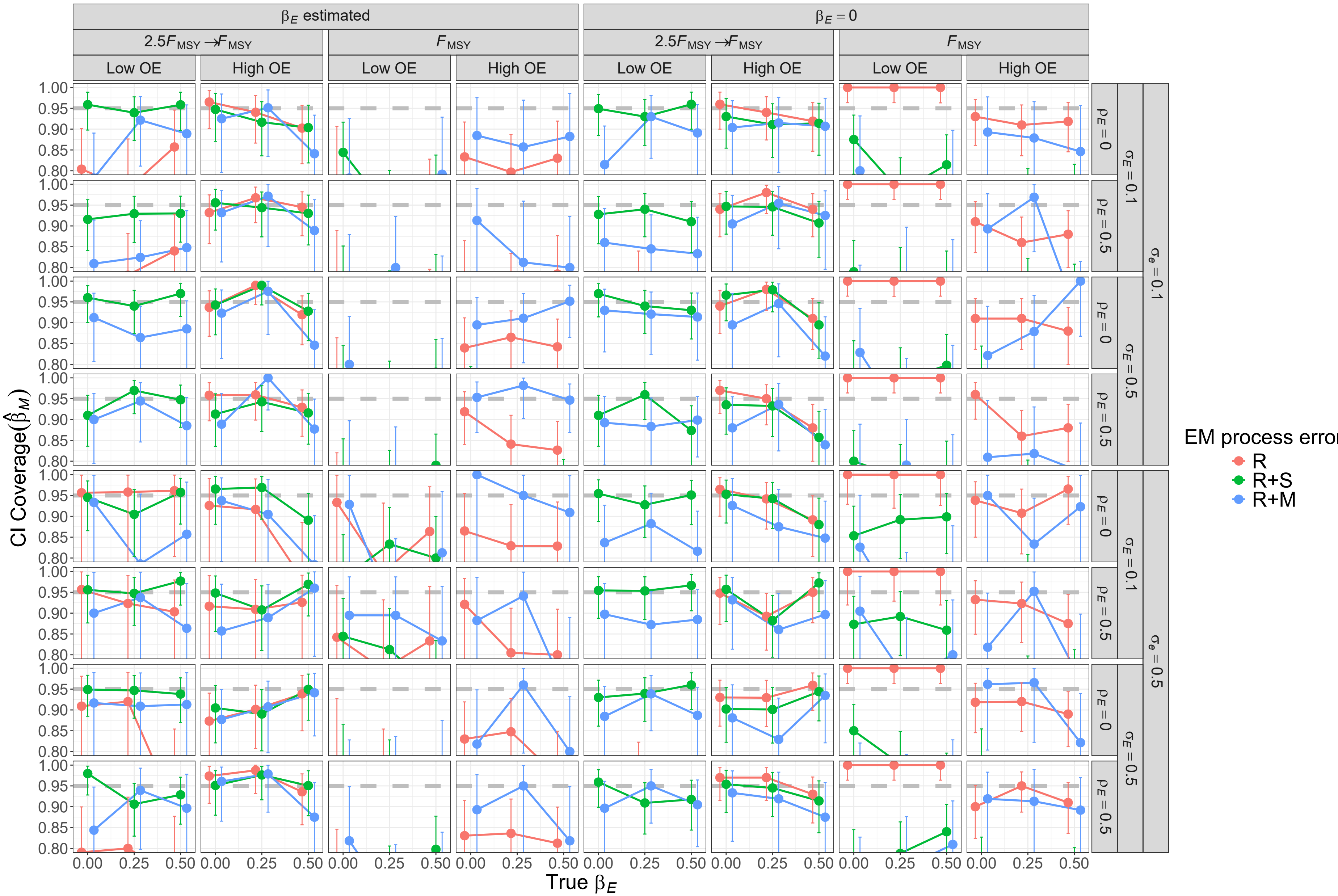

Fig. S38. For R+S OMs, probability of 95% confidence interval for β_M containing the true value for EMs with alternative process error assumptions and treatment of covariate effect ($\beta_E = 0$ or estimated). Vertical lines represent 95% confidence intervals.

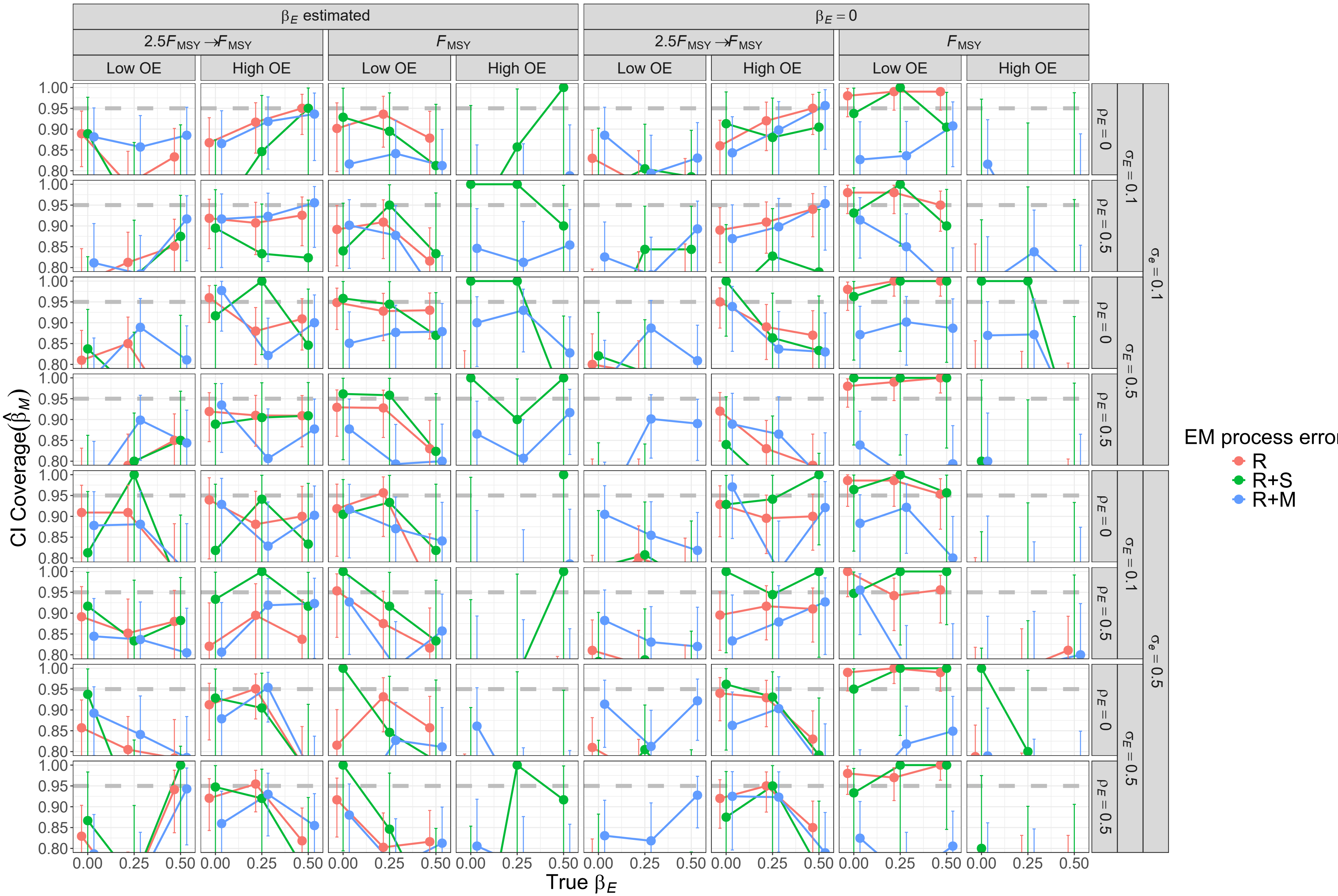

Fig. S39. For R+M OMs, probability of 95% confidence interval for β_M containing the true value for EMs with alternative process error assumptions and treatment of covariate effect ($\beta_E = 0$ or estimated). Vertical lines represent 95% confidence intervals.

Median natural mortality parameter estimate and standard error example

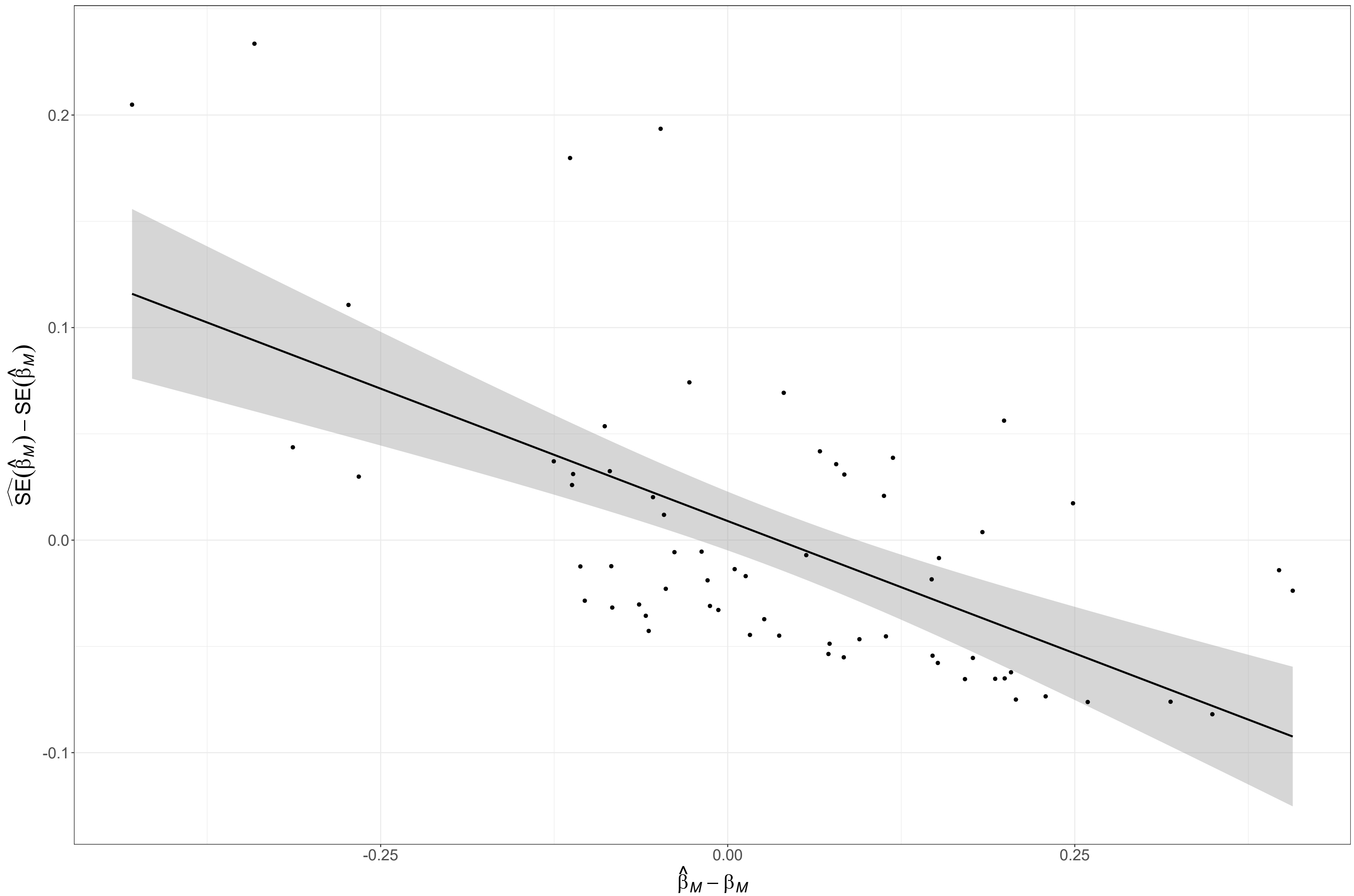

Fig. S40. Negative correlation of β_M estimates and Hessian-based standard error estimates for EM that also estimates the covariate effect and has correct R+M process error assumption fitted to simulated data from the OM with R+M process errors, temporal contrast in fishing pressure, low observation uncertainty for both population ($LowOE$) and covariate observations ($\sigma_e = 0.1$), high and uncorrelated temporal variability in the true covariate ($\sigma_E = 0.5$ and $\rho_E = 0$), and the strongest covariate effect on natural mortality ($\beta_E = 0.5$).

Median Natural mortality parameter RMSE

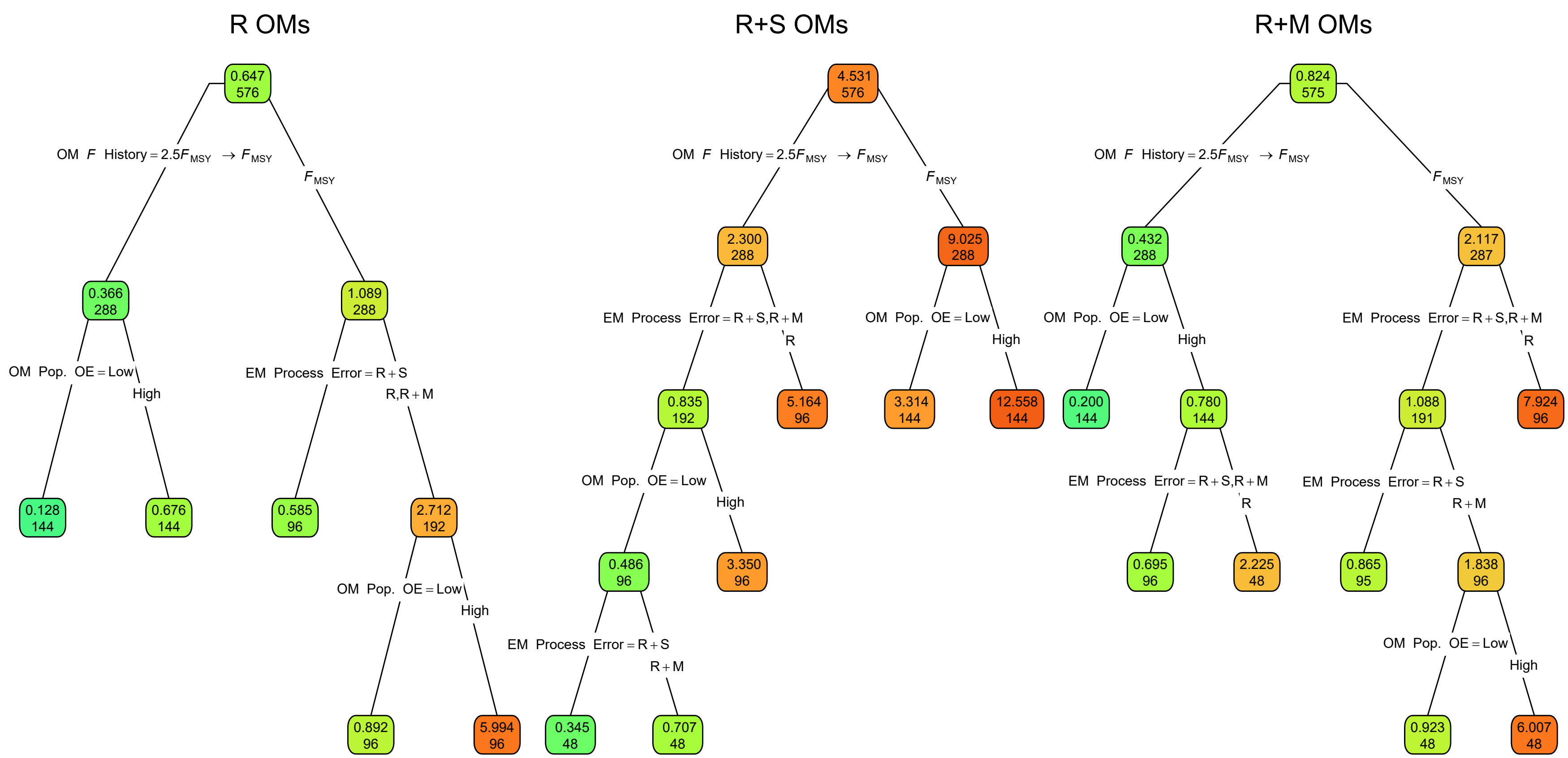

Fig. S41. Regression trees indicating primary factors determining reductions in sums of squares of errors in estimation measured by Eq. **??** for the RMSE of estimates for the median natural mortality rate parameter (RMSE $\left(\widehat{\beta}_M\right)$) in EMs fitted to R, R+S, and R+M OMs. Each node shows the median error (top) and number of observations (bottom) for the corresponding subset. Lower or higher median absolute errors of the process error assumption are indicated by more green or red polygons, respectively.

Table S9. For each OM process error source (columns), percent reduction in deviance for linear regression models fit to errors in estimation measured by Eq. **??** for the the RMSE for estimates of the median natural mortality rate parameter (RMSE $\left(\widehat{\beta}_M\right)$) with each OM and EM factor (rows) included individually, combined, and with second and third order interactions.

Factor	R	R+S	R+M
OM F history	34.70	24.62	34.91
OM Obs. Error	25.55	17.82	12.77
$OM\sigma_e$	0.01	<0.01	<0.01
$OM\sigma_E$	0.19	0.02	0.07
$OM\rho_E$	0.05	0.26	0.03
OM β_E	0.58	0.23	0.16
EM Process Error	10.75	18.12	14.50
$EM\beta_E$ assumption	3.96	6.39	2.93
All factors	75.79	67.47	65.26
+ All Two Way	90.21	80.65	77.59
+ All Three Way	93.72	88.68	86.44

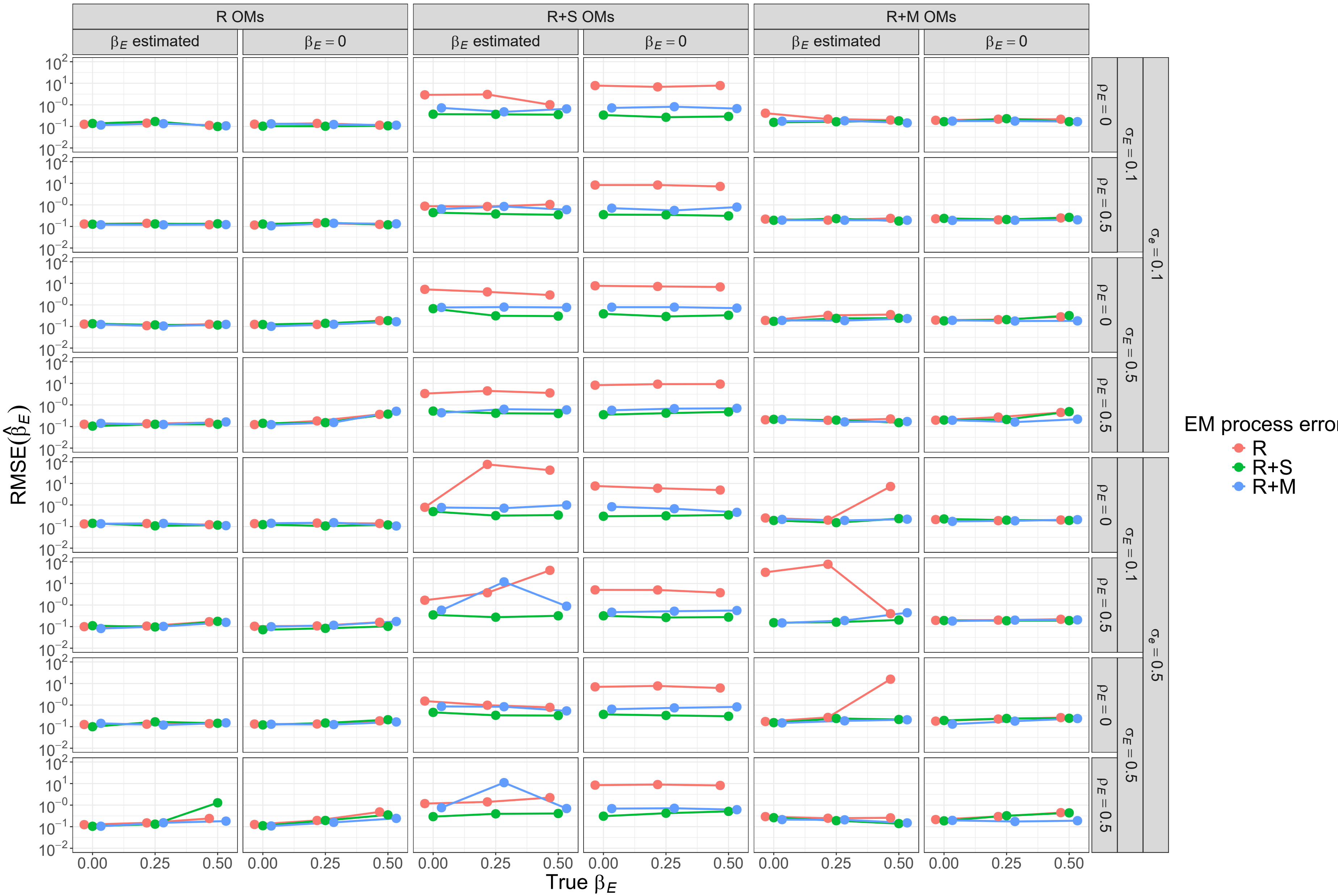

Fig. S42. Root mean square error (RMSE) of estimates of β_M from fitting EMs with alternative process error assumptions and treatment of covariate effect ($\beta_E = 0$ or estimated). All OMs had low observation error for population observations and contrast in fishing mortality.

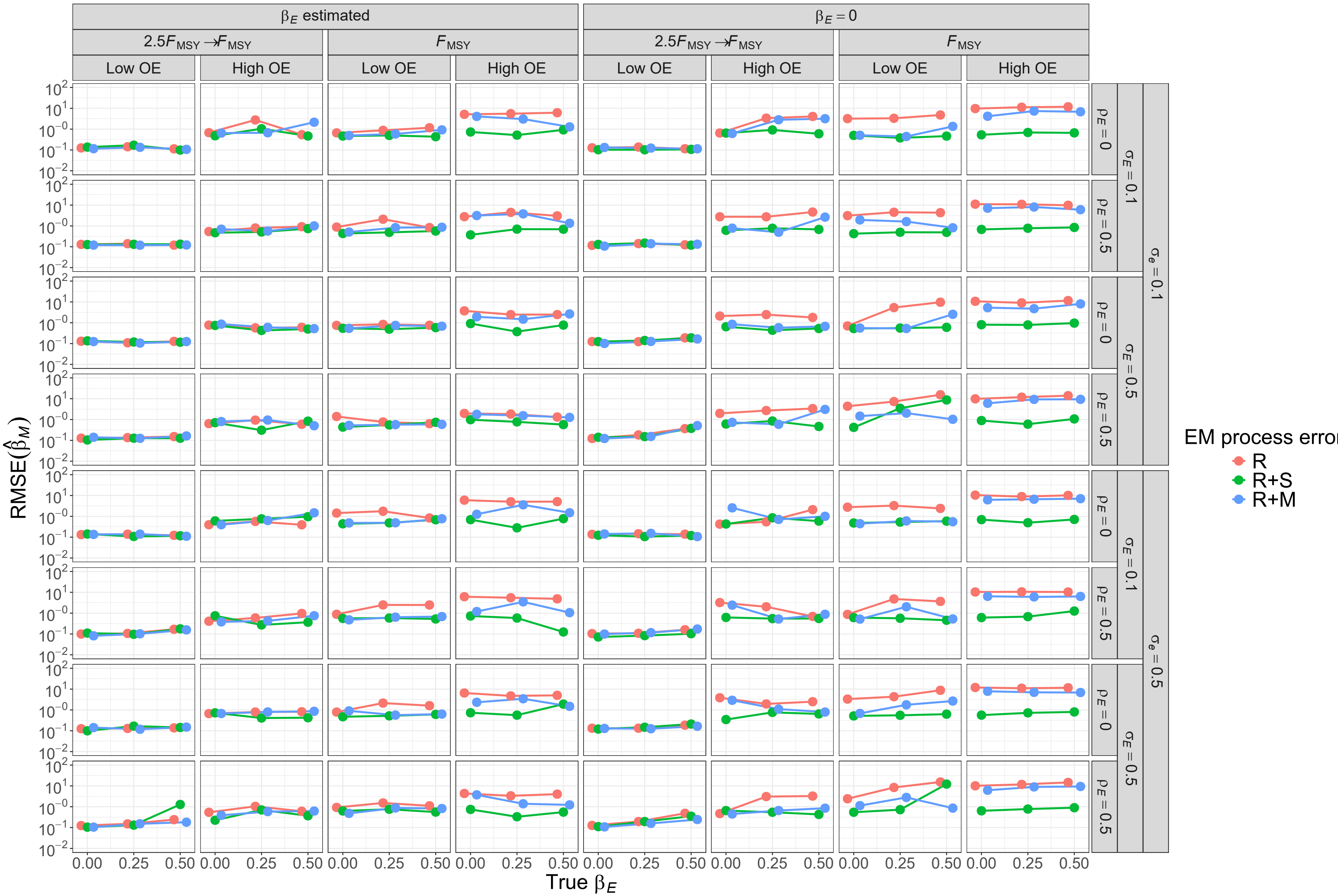

Fig. S43. For R OMs, root mean square error (RMSE) of estimates of β_M from fitting EMs with alternative process error assumptions and treatment of covariate effect ($\beta_E = 0$ or estimated).

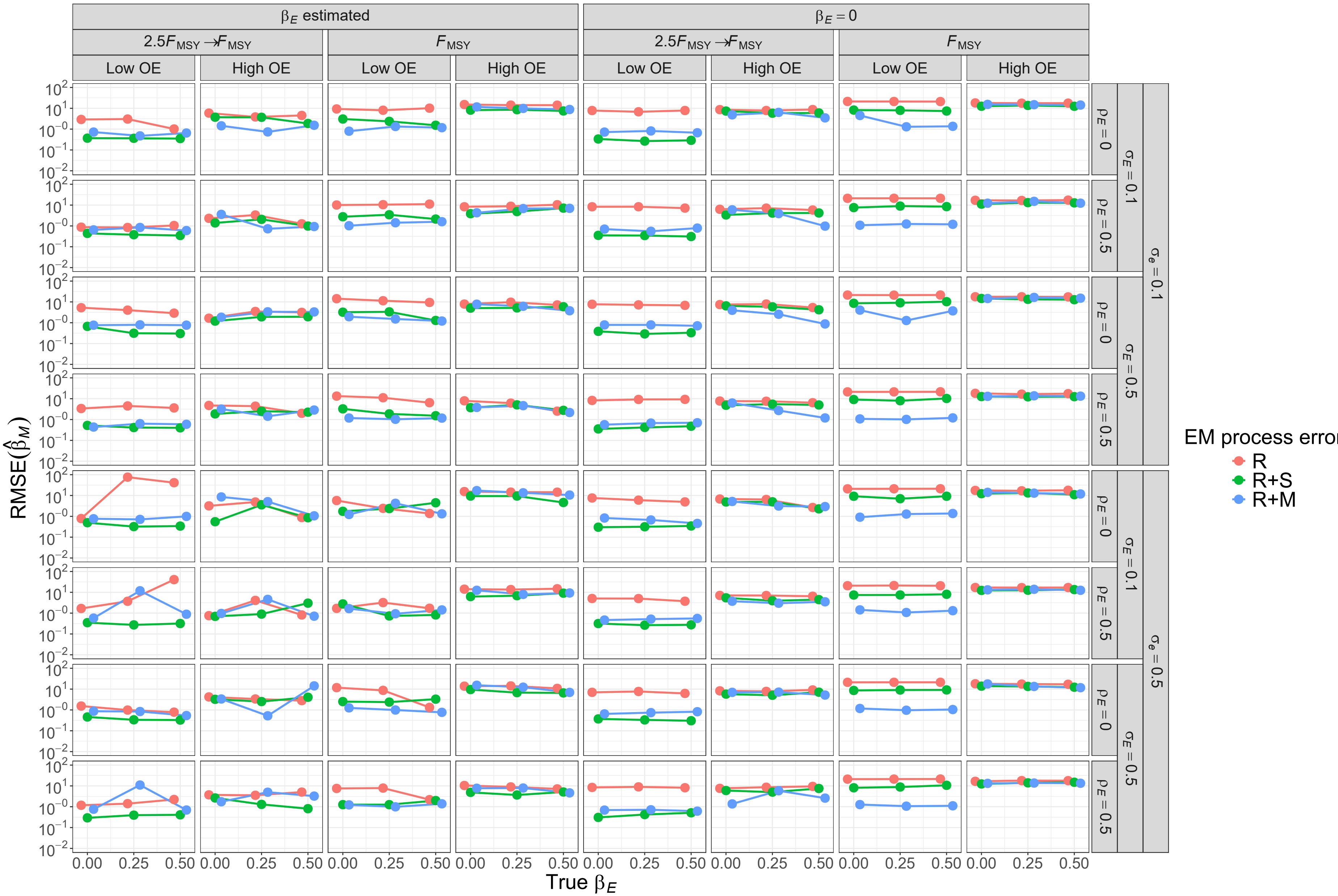

Fig. S44. For R+S OMs, root mean square error (RMSE) of estimates of β_M from fitting EMs with alternative process error assumptions and treatment of covariate effect ($\beta_E = 0$ or estimated).

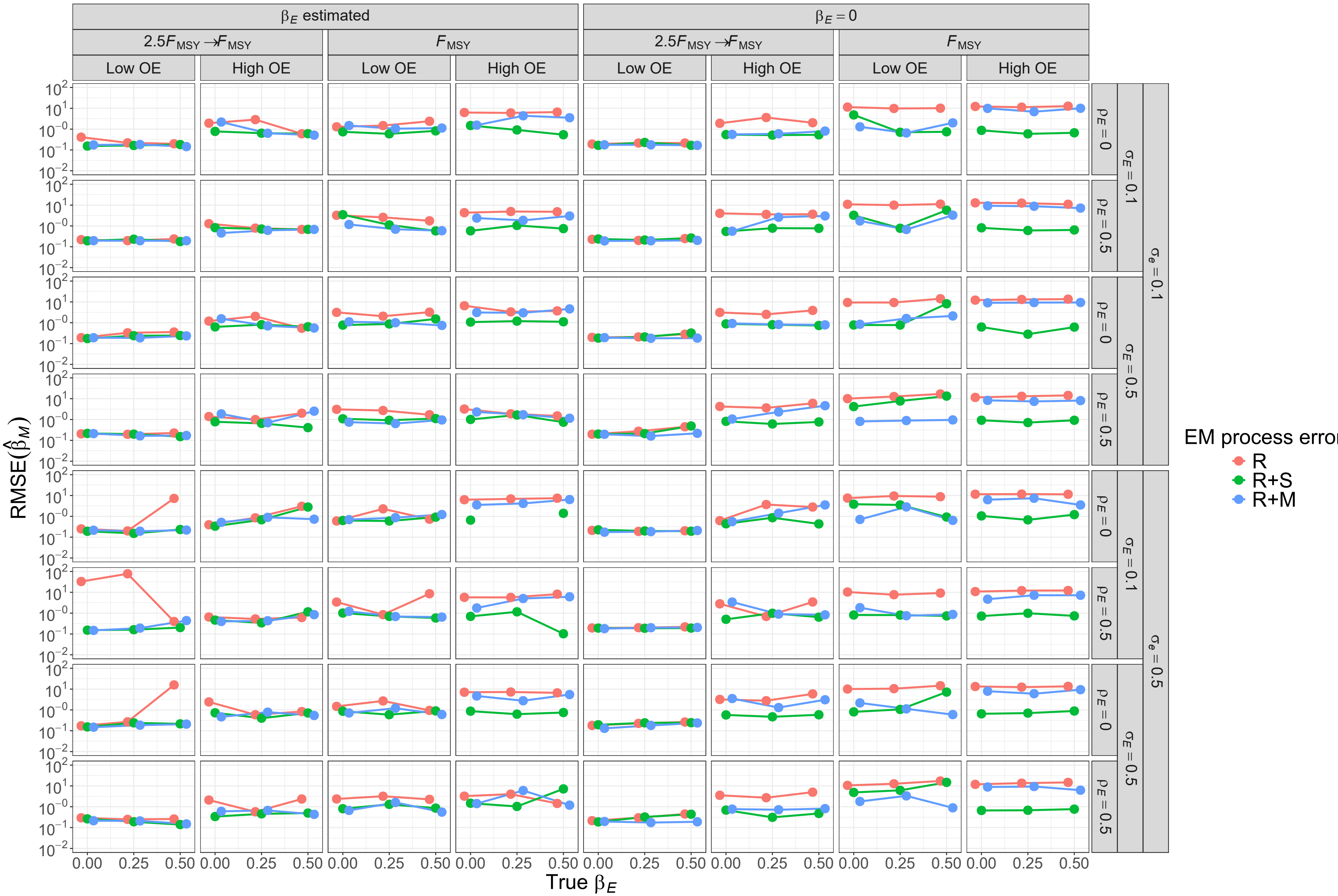

Fig. S45. For R+M OMs, root mean square error (RMSE) of estimates of β_M from fitting EMs with alternative process error assumptions and treatment of covariate effect ($\beta_E = 0$ or estimated).

Terminal year natural mortality bias

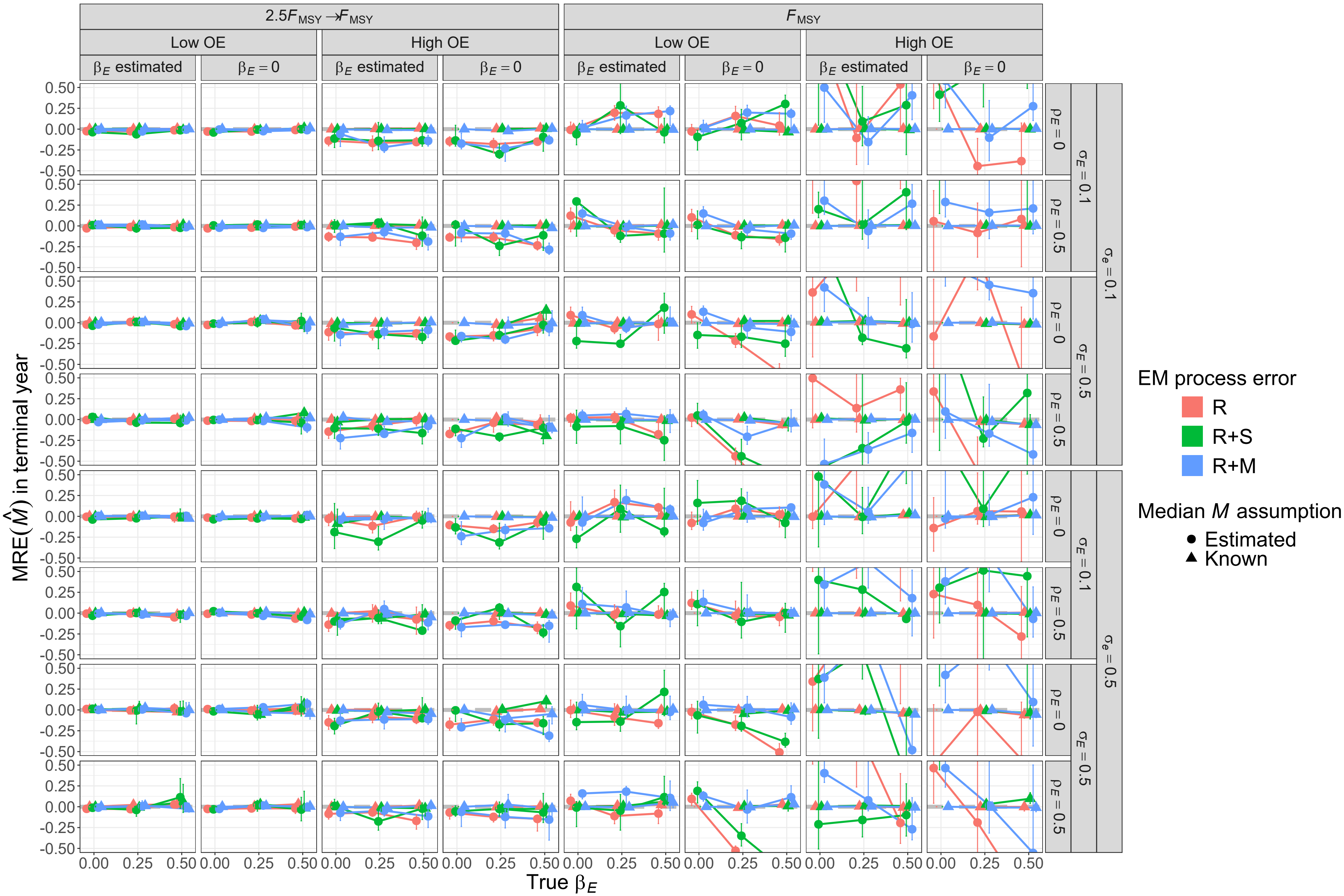

Fig. S46. For R OMs, median relative error (MRE) of estimates of natural mortality rate in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known).

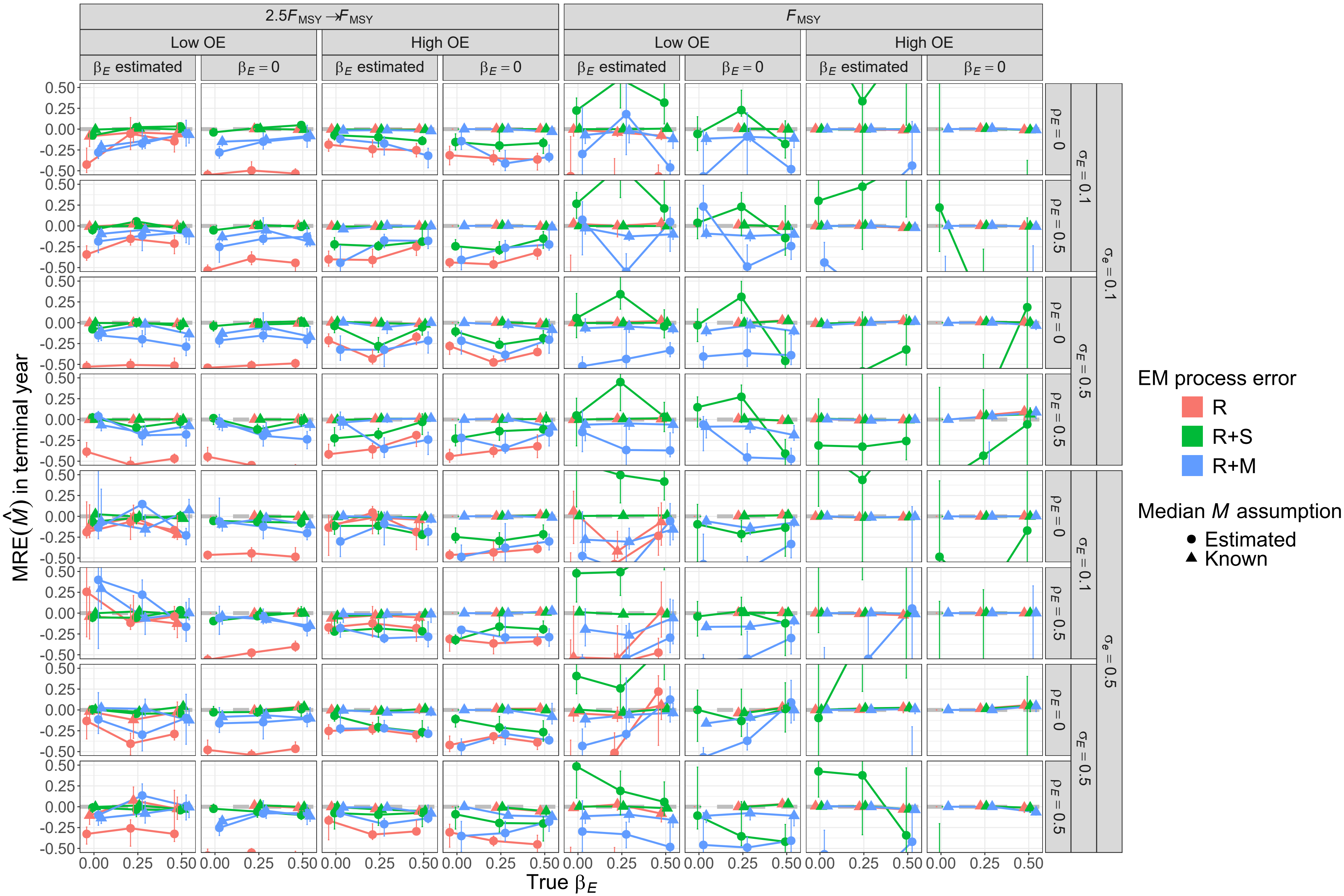

Fig. S47. For R+S OMs, median relative error (MRE) of estimates of natural mortality rate in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known).

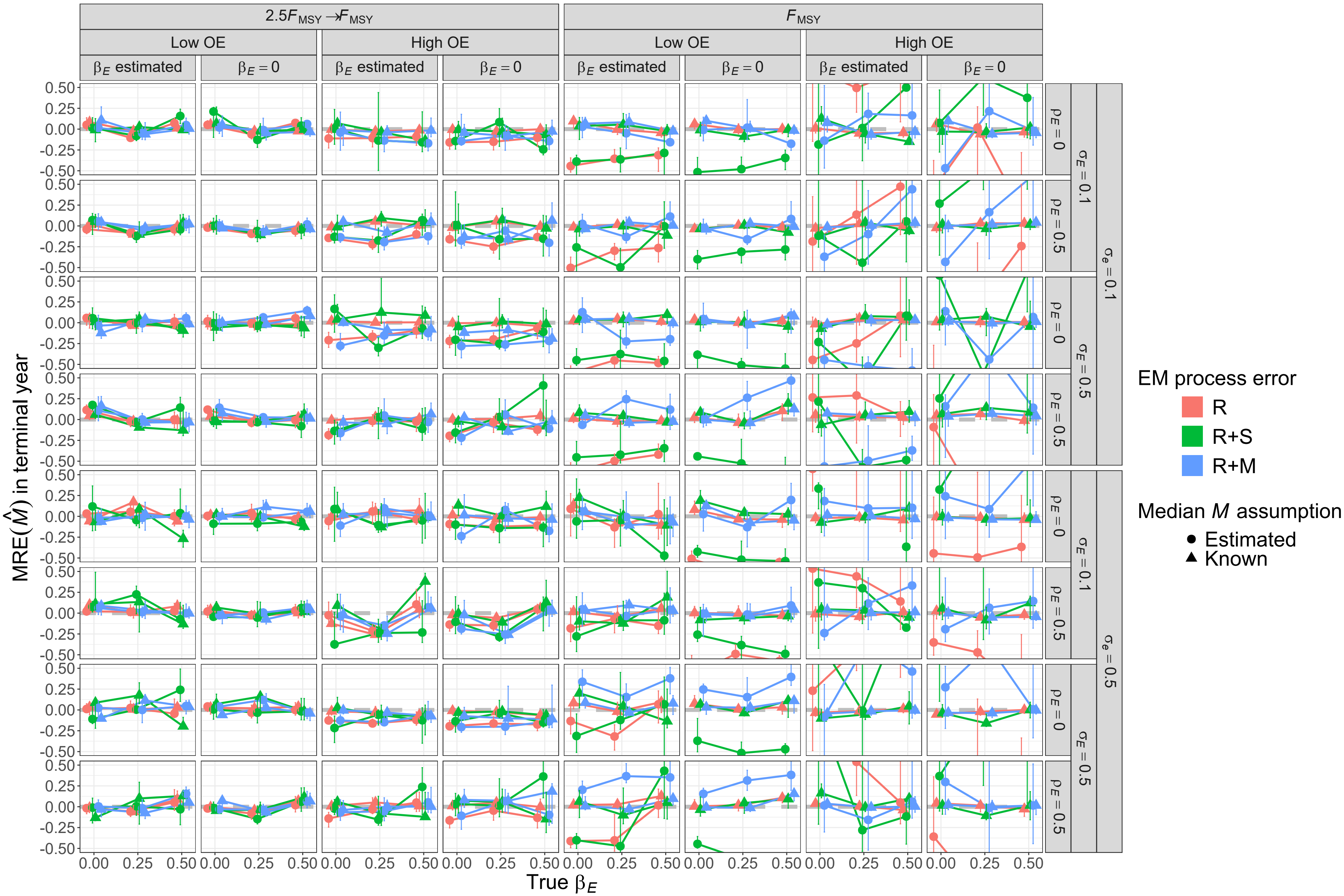

Fig. S48. For R+M OMs, median relative error (MRE) of estimates of natural mortality rate in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known).

Terminal year natural mortality RMSE

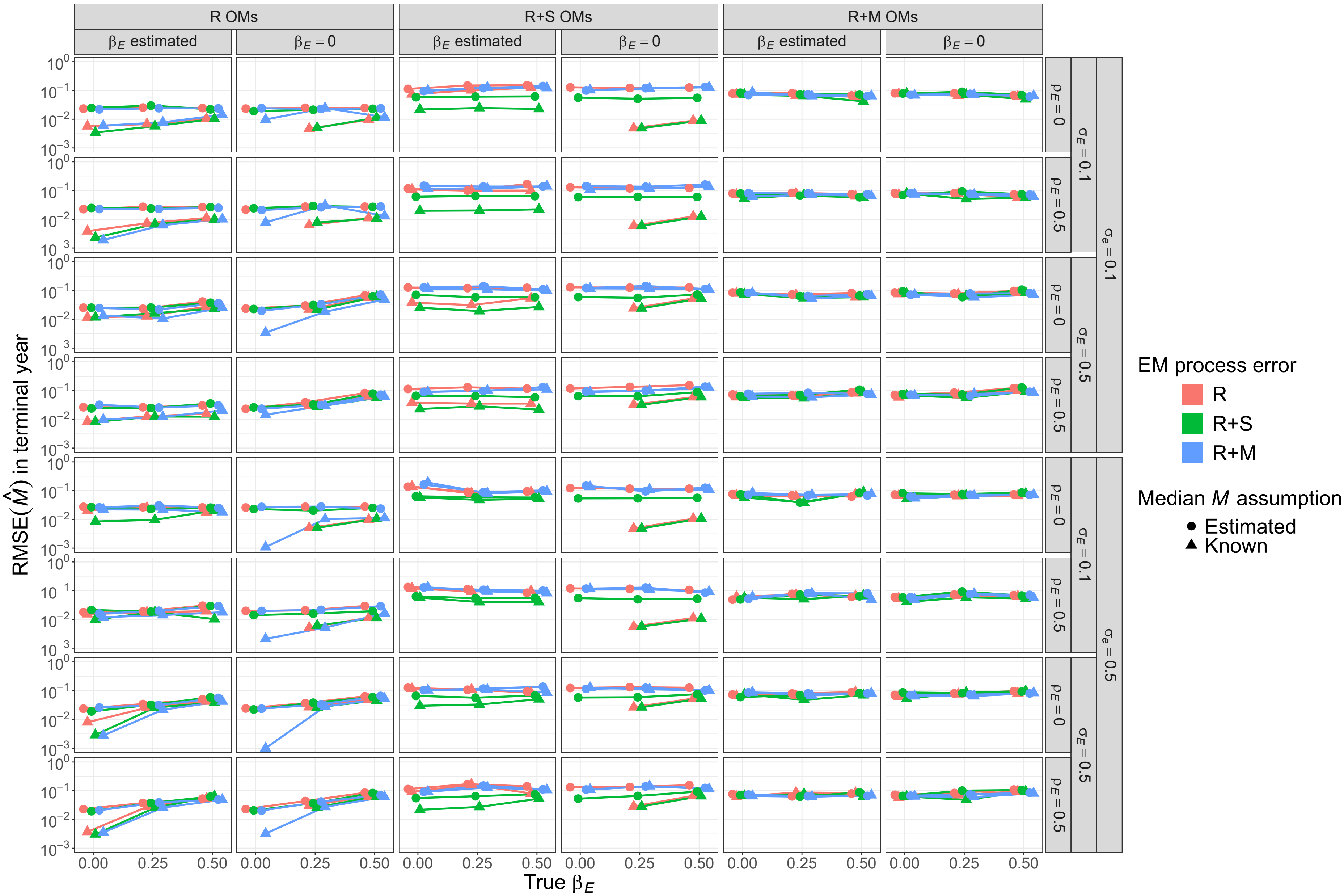

Fig. S49. Root mean square error (RMSE) of estimates of natural mortality rate in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known). All OMs had low population observation error and contrast in fishing mortality.

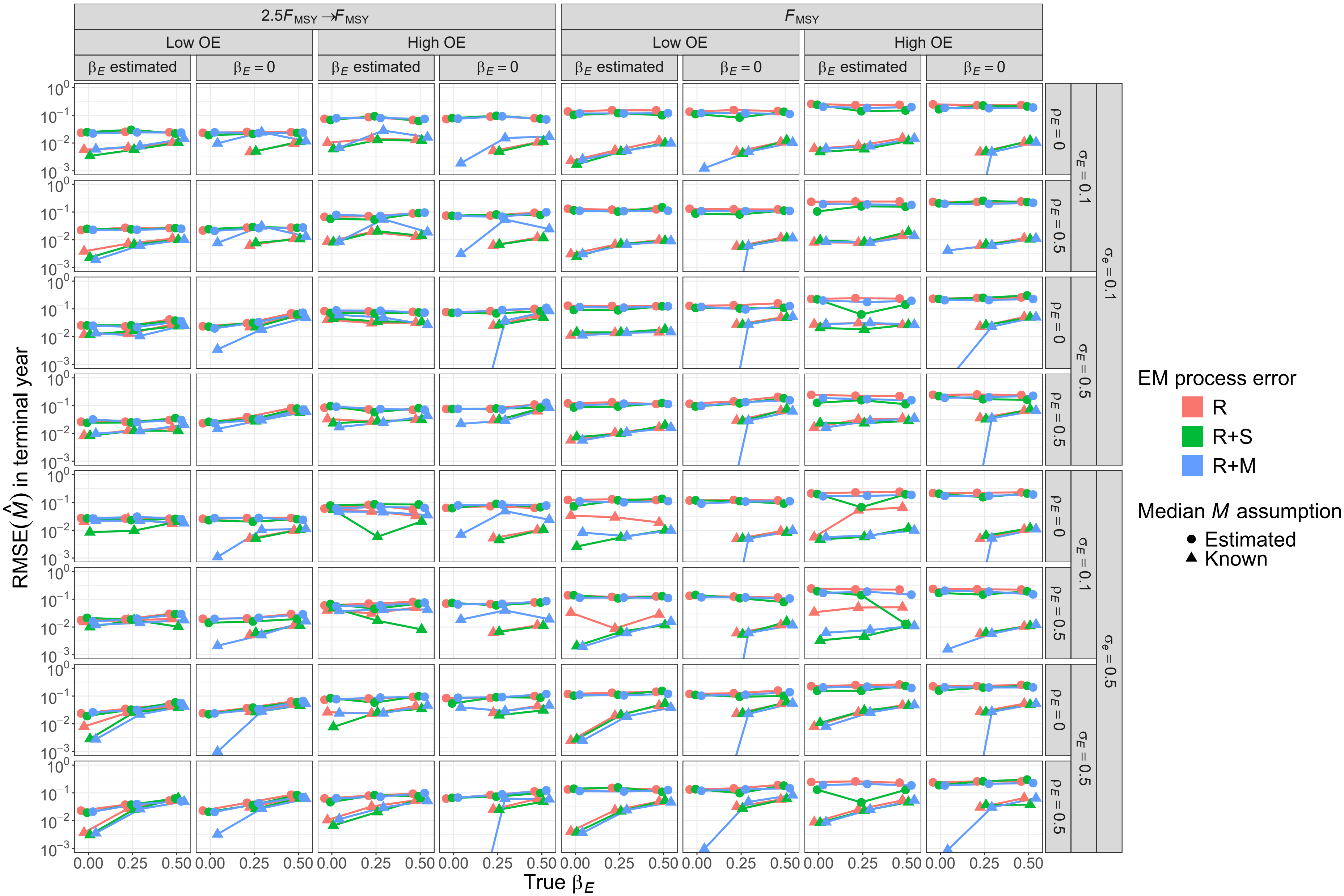

Fig. S50. For R OMs, root mean square error (RMSE) of estimates of natural mortality rate in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known).

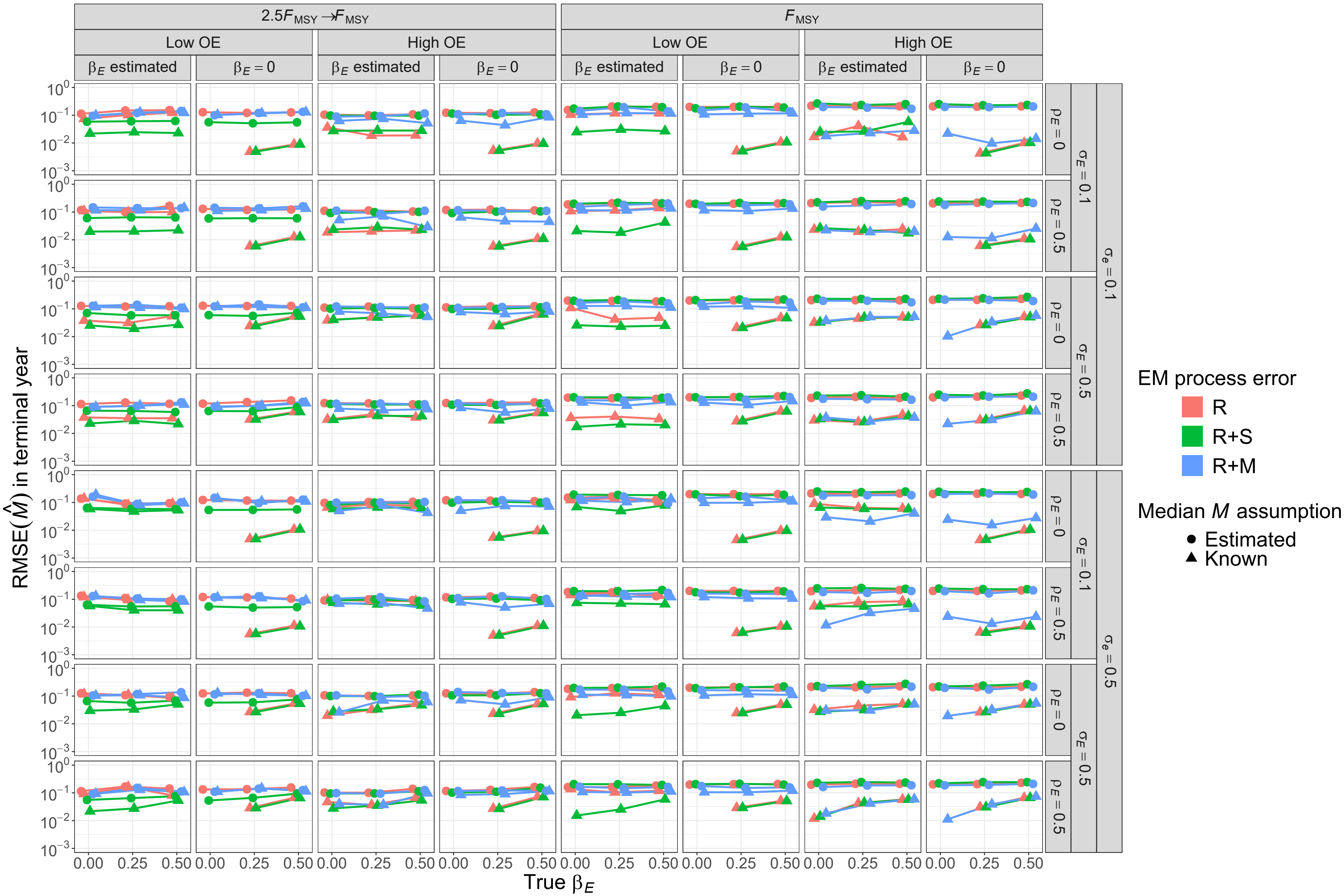

Fig. S51. For R+S OMs, root mean square error (RMSE) of estimates of natural mortality rate in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known).

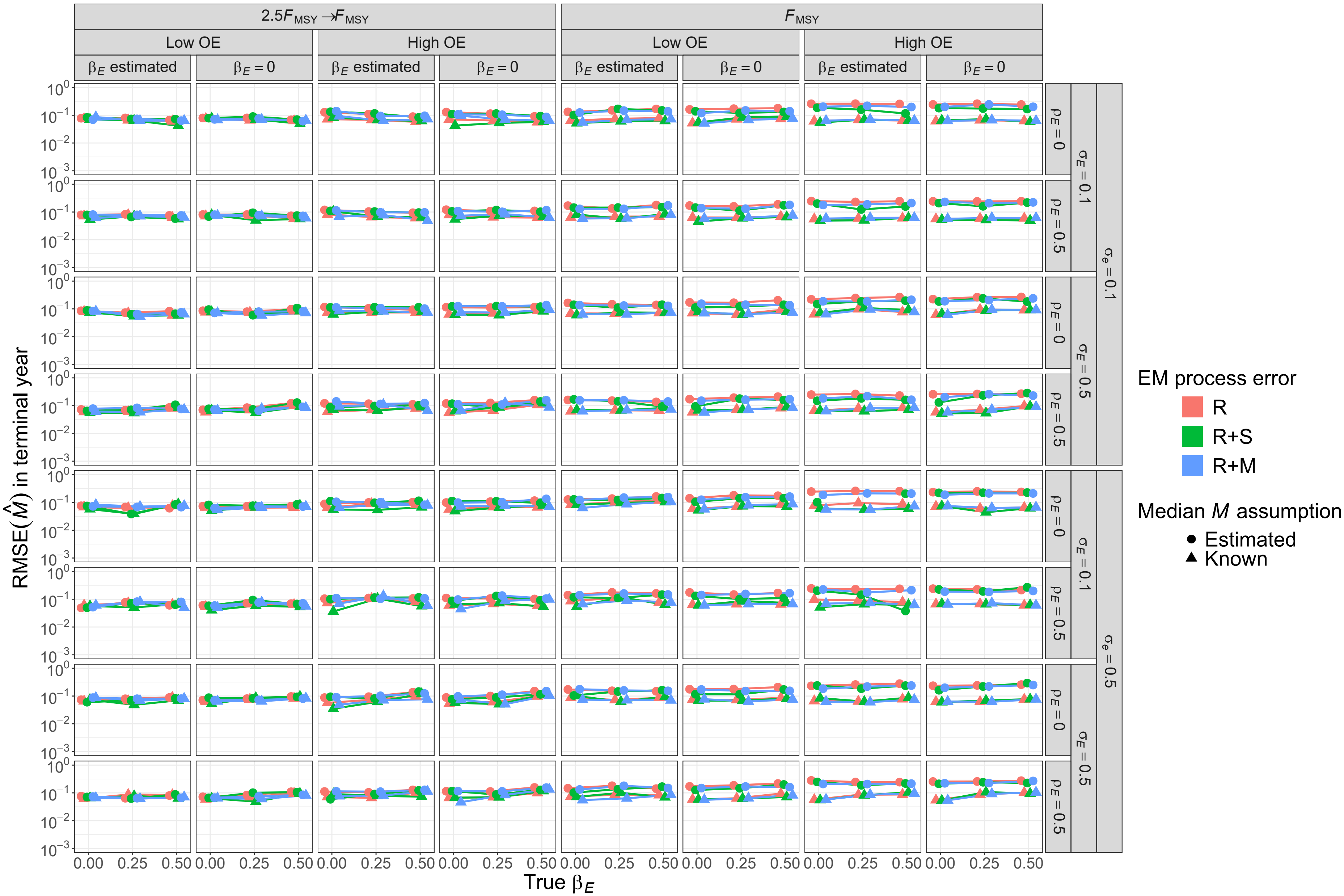

Fig. S52. For R+M OMs, root mean square error (RMSE) of estimates of natural mortality rate in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known).

Terminal year spawning stock biomass bias

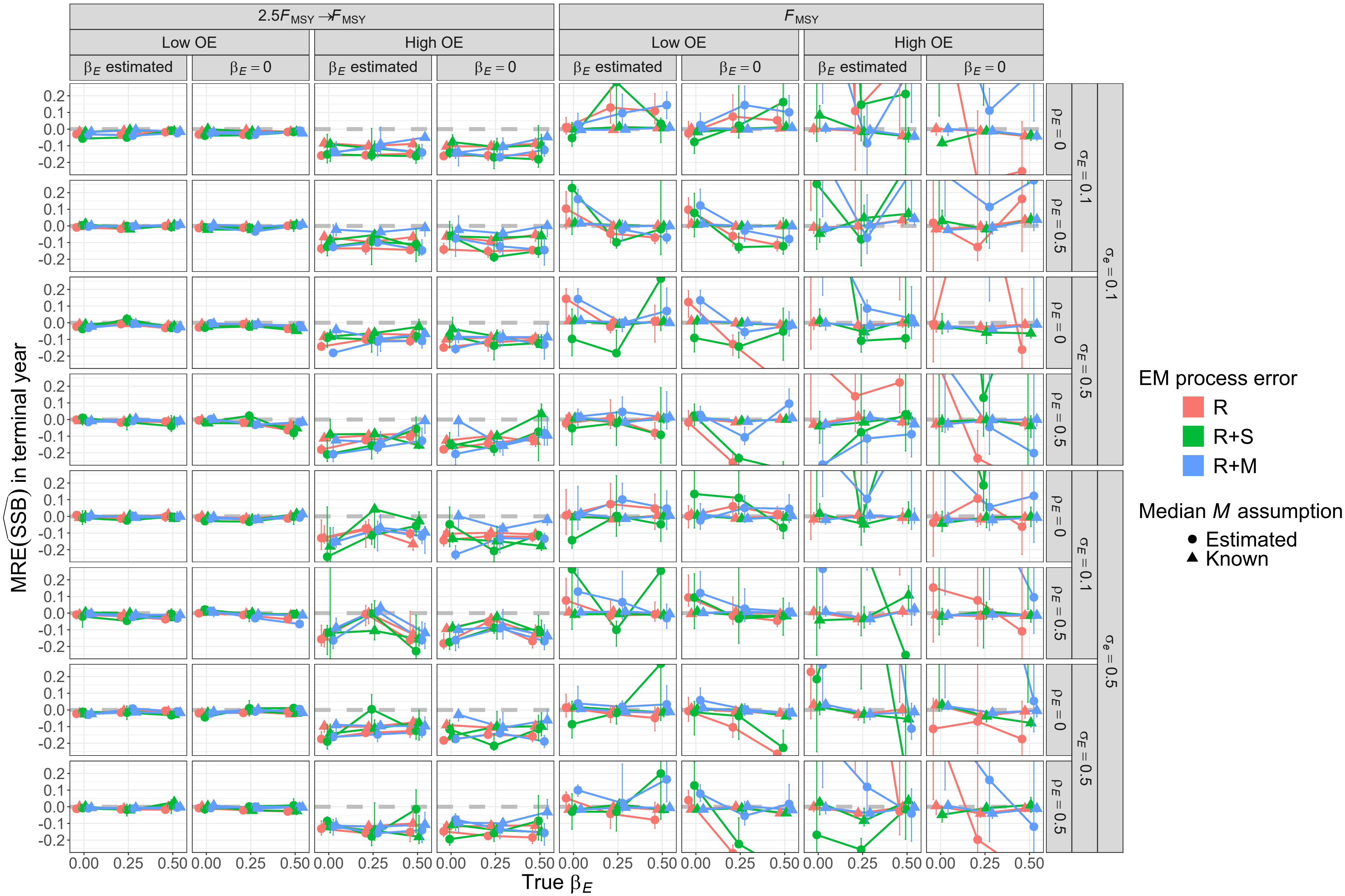

Fig. S53. For R OMs, median relative error (MRE) of estimates of spawning stock biomass (SSB) in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known).

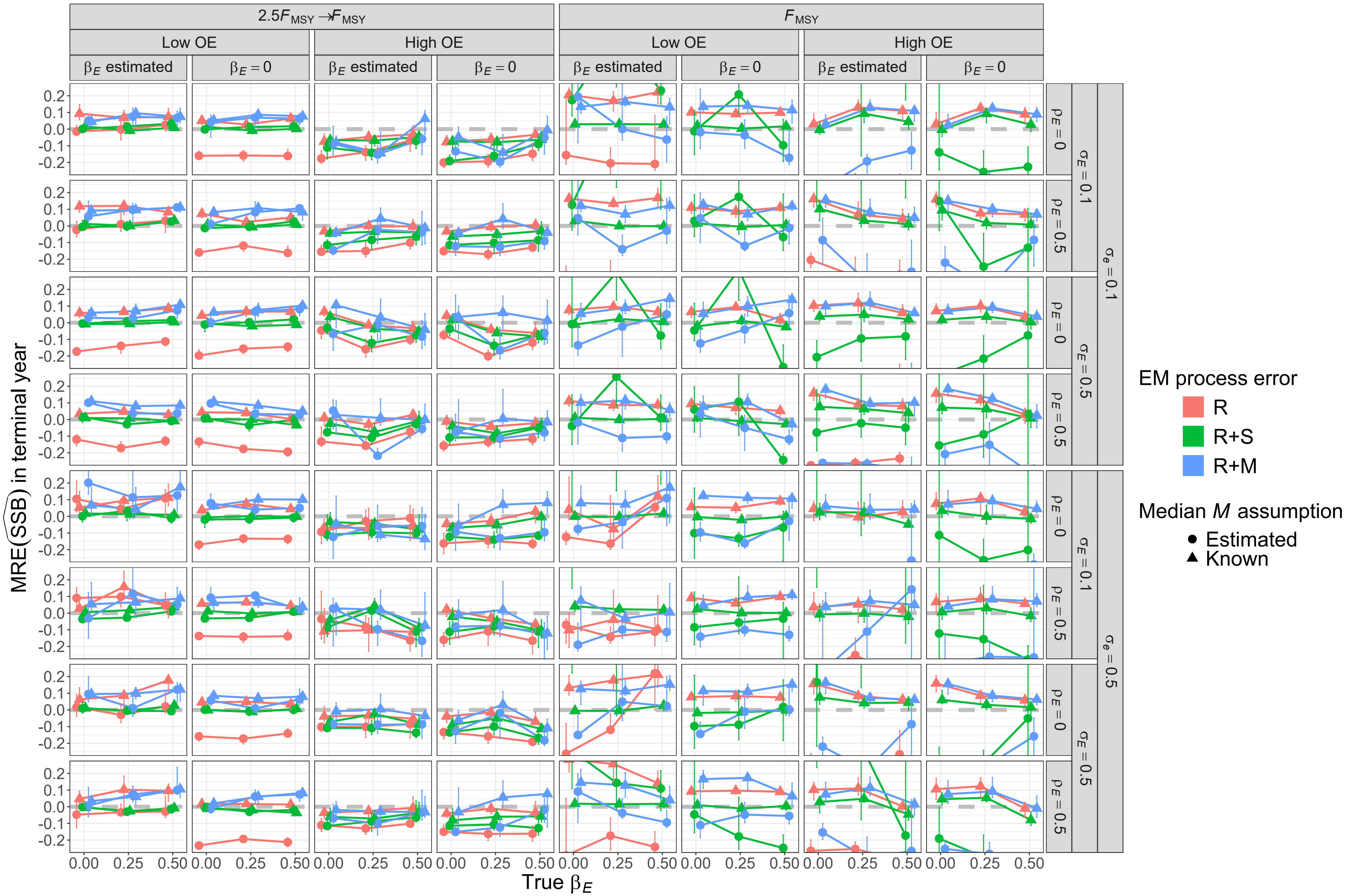

Fig. S54. For R+S OMs, median relative error (MRE) of estimates of spawning stock biomass (SSB) in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known).

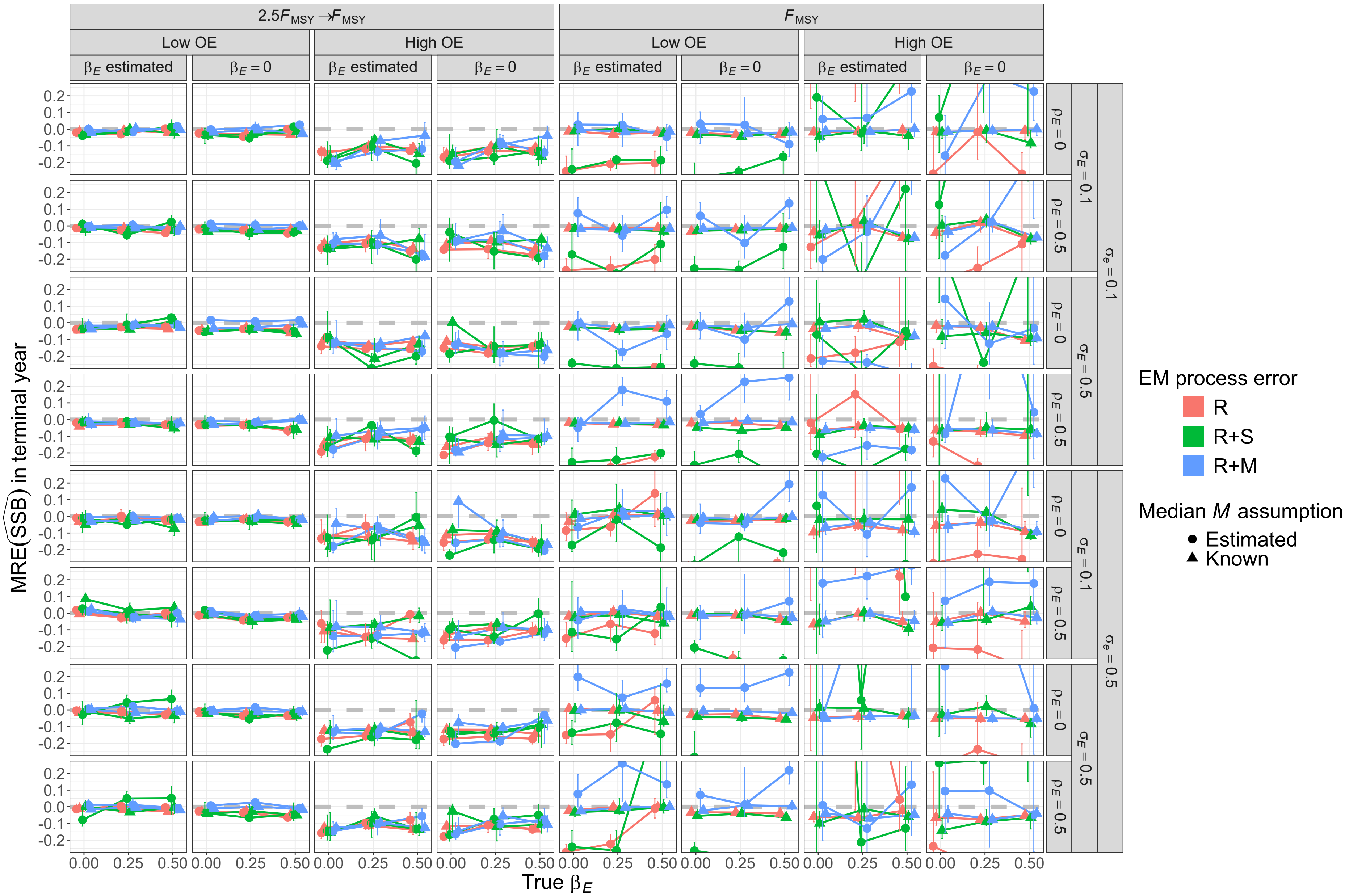

Fig. S55. For R+M OMs, median relative error (MRE) of estimates of spawning stock biomass (SSB) in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known).

Terminal year spawning stock biomass RMSE

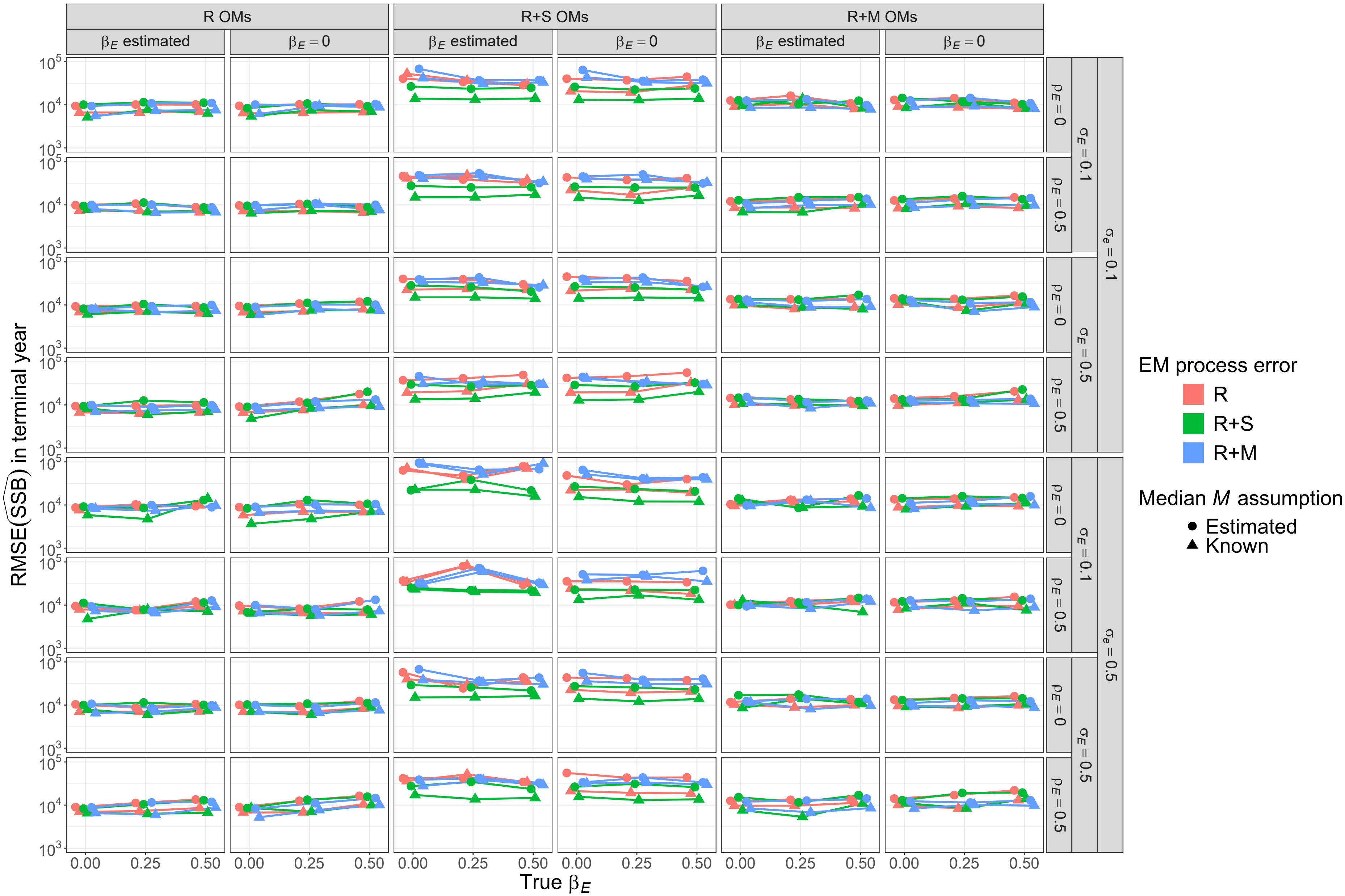

Fig. S56. Root mean square error (RMSE) of estimates of spawning stock biomass in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known). All OMs had low population observation error and contrast in fishing mortality.

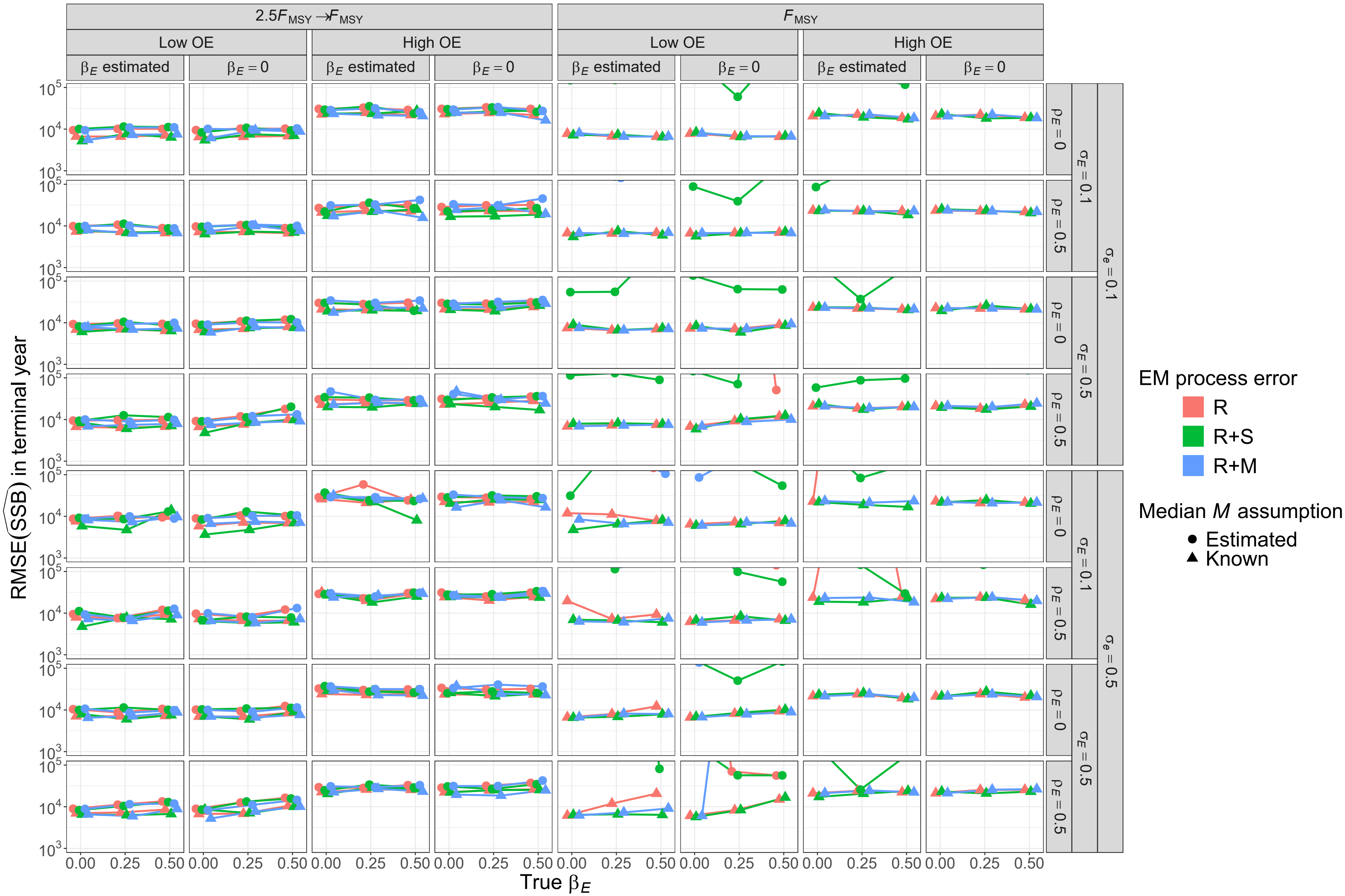

Fig. S57. For R OMs, root mean square error (RMSE) of estimates of spawning stock biomass in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known).

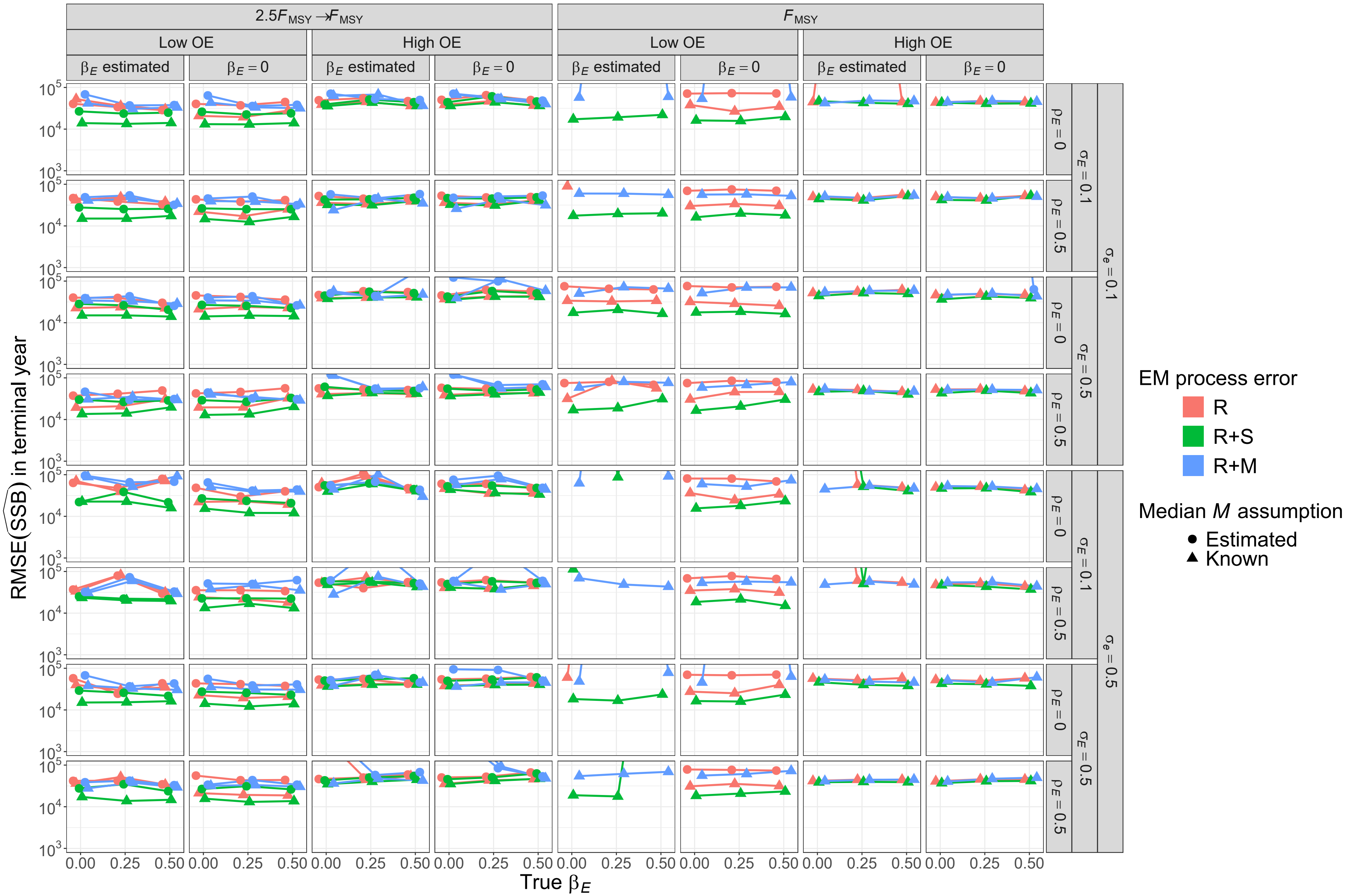

Fig. S58. For R+S OMs, root mean square error (RMSE) of estimates of spawning stock biomass in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known).

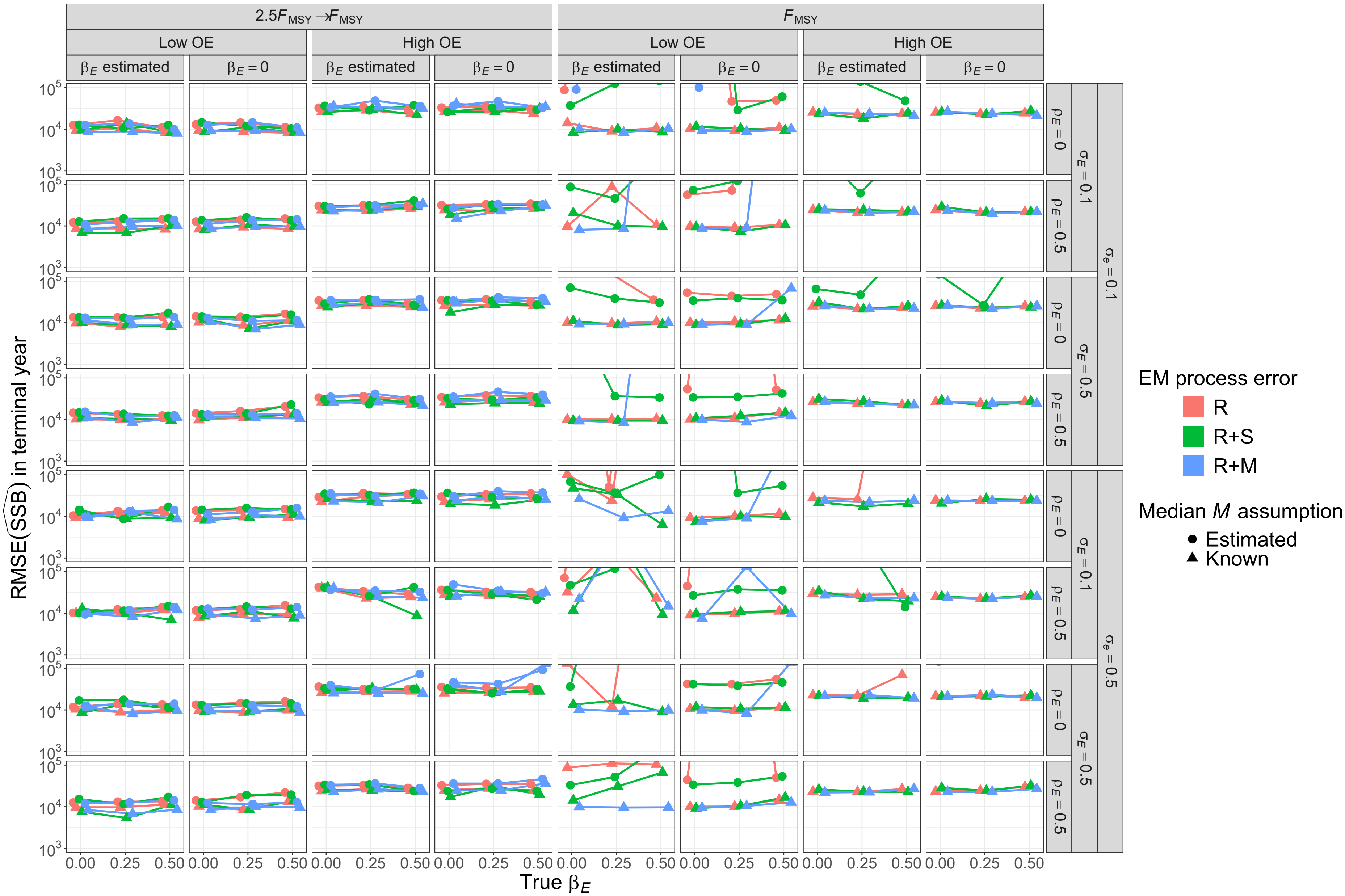

Fig. S59. For R+M OMs, root mean square error (RMSE) of estimates of spawning stock biomass in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known).

Terminal year fishing mortality bias

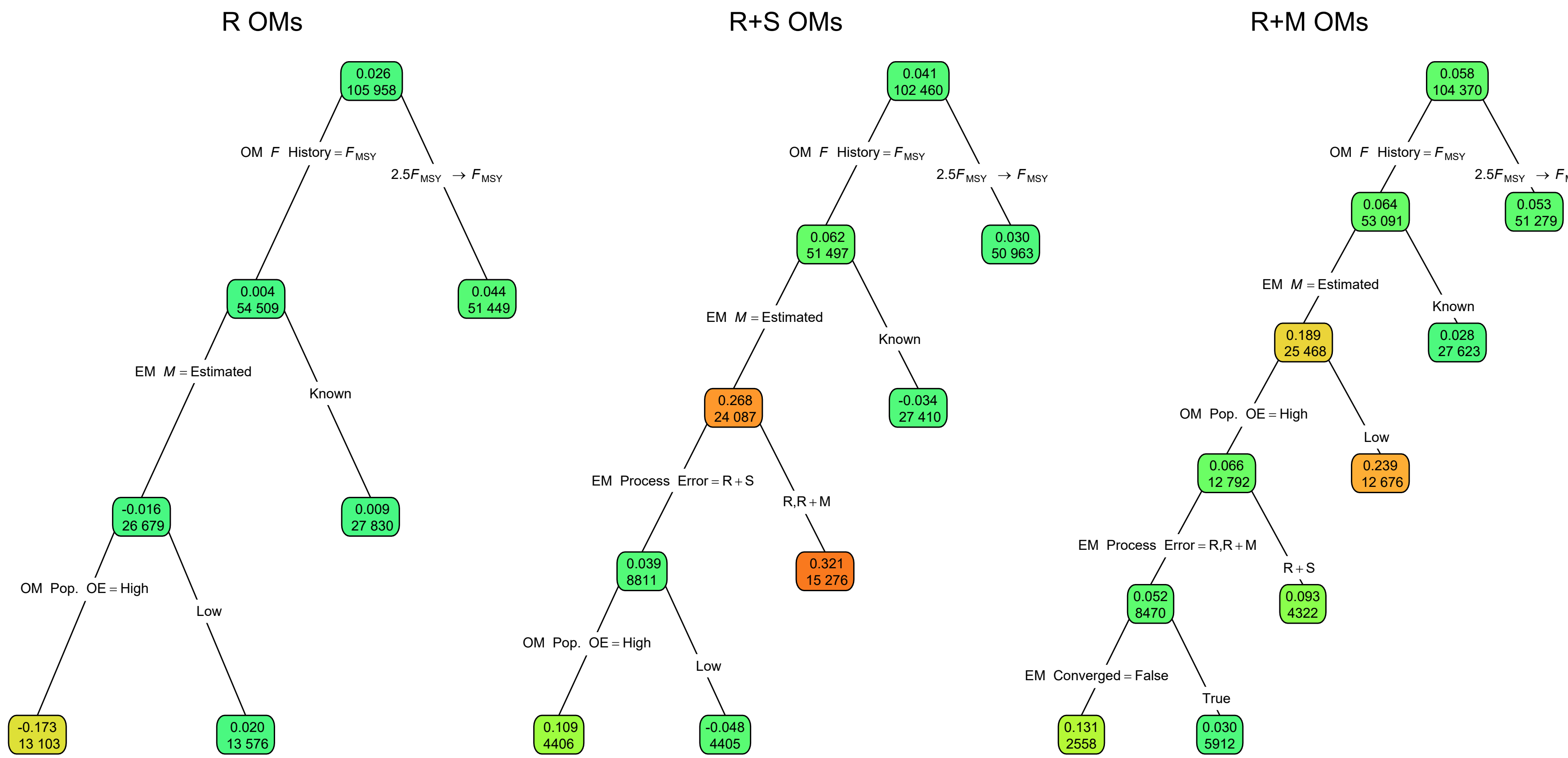

Fig. S60. Regression trees indicating primary factors determining reductions in sums of squares of errors in estimation measured by Eq. **??** for the terminal year fishing mortality rate in EMs fitted to R, R+S, and R+M OMs. Each node shows the median error (top) and number of observations (bottom) for the corresponding subset. Lower or higher median absolute errors of the process error assumption are indicated by more green or red polygons, respectively.

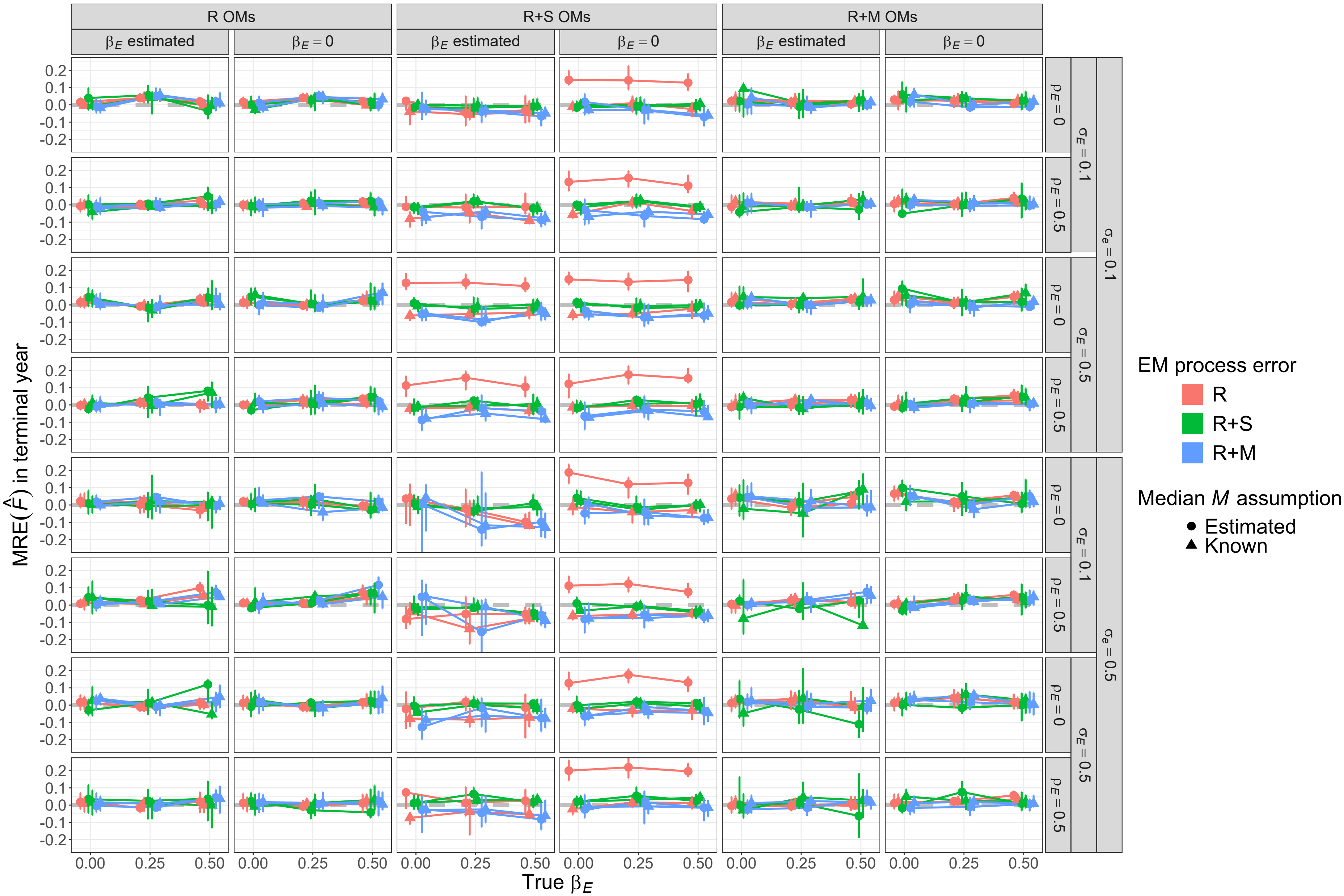

Fig. S61. Median relative error (MRE) of estimates of fully-selected fishing mortality (F) in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known). All OMs had low population observation error and contrast in fishing mortality.

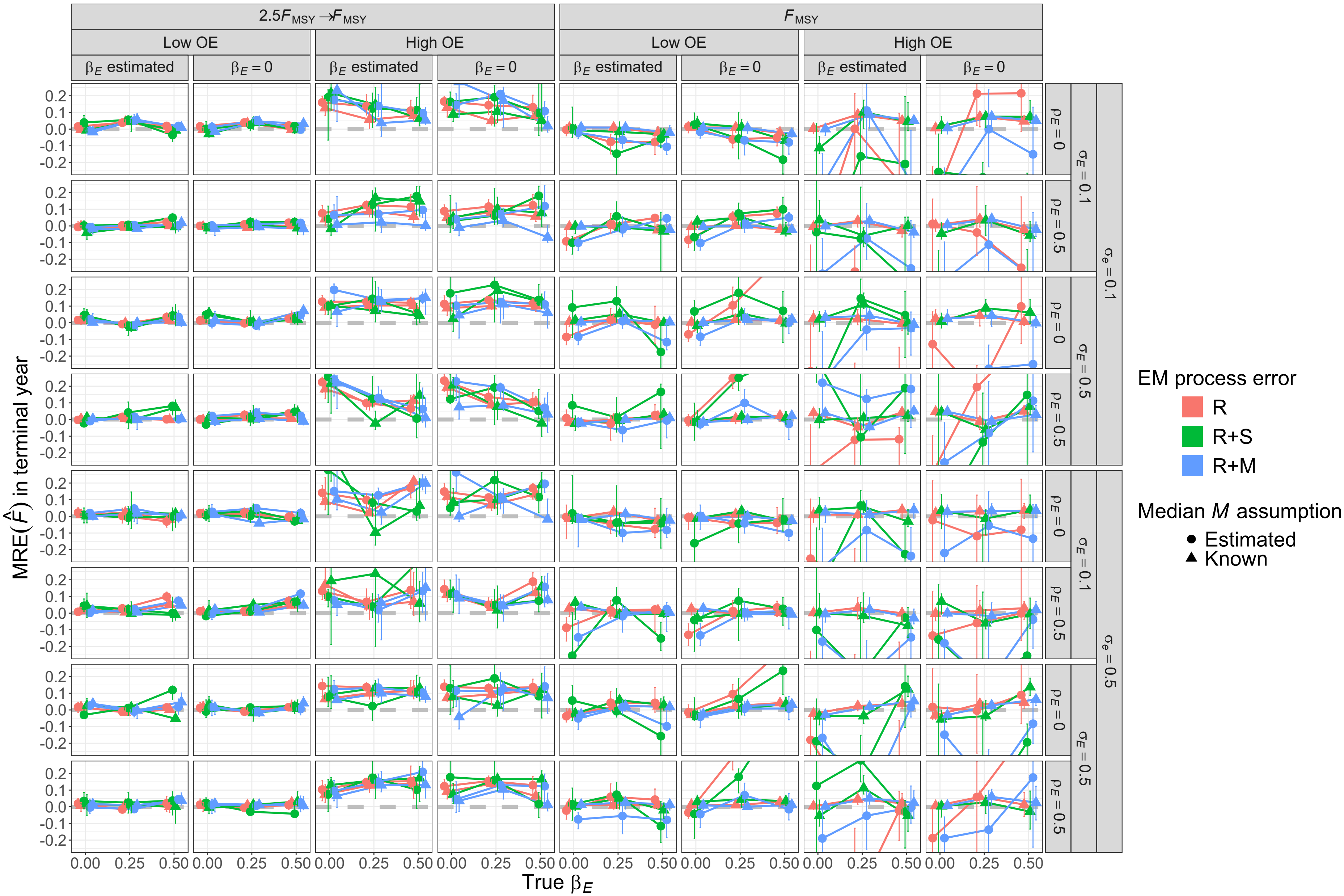

Fig. S62. For R OMs, median relative error (MRE) of estimates of fully-selected fishing mortality (F) in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known). All OMs had low observation error and contrast in fishing mortality.

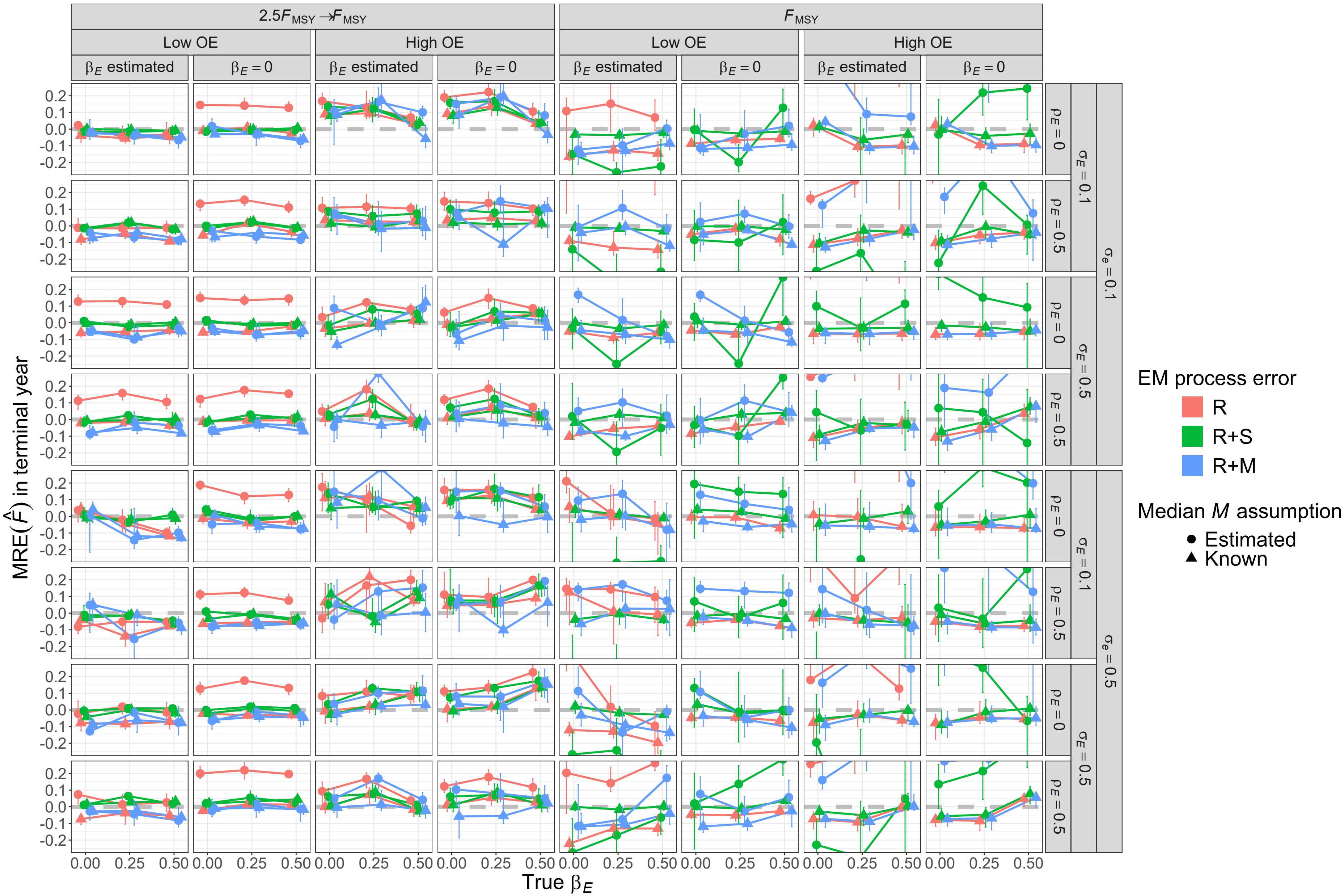

Fig. S63. For R+S OMs, median relative error (MRE) of estimates of fully-selected fishing mortality (F) in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known). All OMs had low observation error and contrast in fishing mortality.

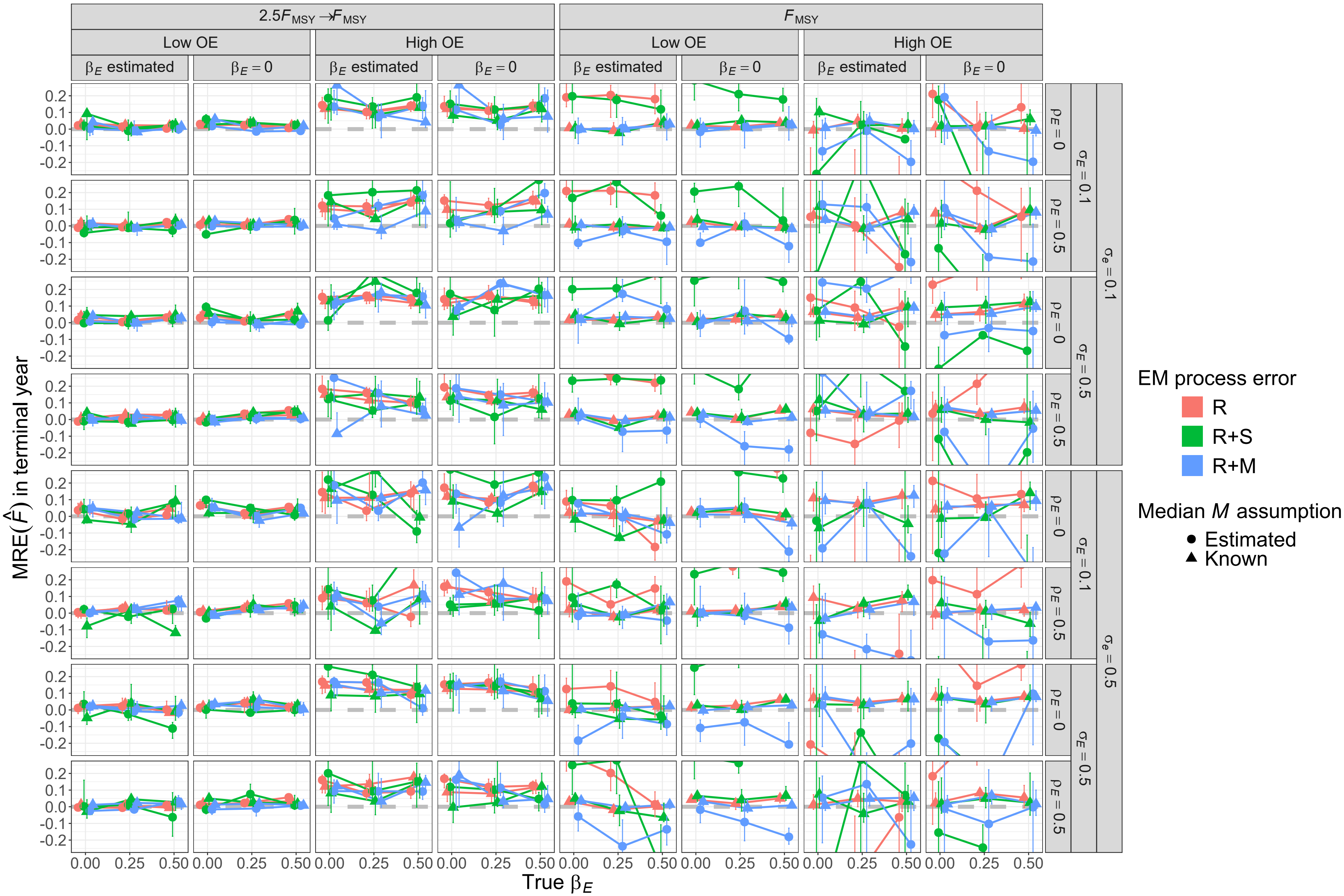

Fig. S64. For R+M OMs, median relative error (MRE) of estimates of fully-selected fishing mortality (F) in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known). All OMs had low observation error and contrast in fishing mortality.

Terminal year fishing mortality RMSE

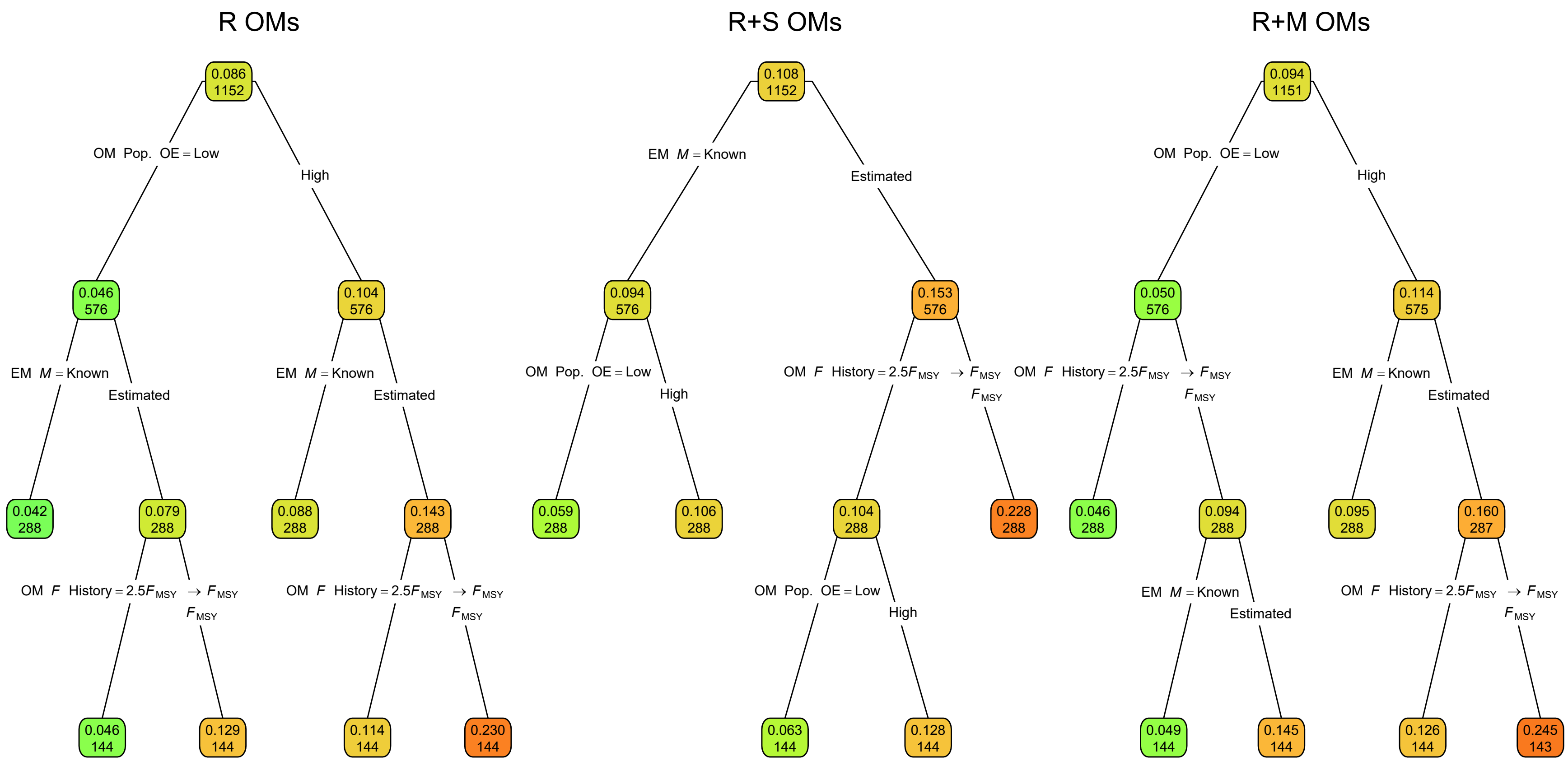

Fig. S65. Regression trees indicating primary factors determining reductions in sums of squares of errors in estimation measured by Eq. ?? for the RMSE of terminal year fishing mortality rate in EMs fitted to R, R+S, and R+M OMs. Each node shows the median error (top) and number of observations (bottom) for the corresponding subset. Lower or higher median absolute errors of the process error assumption are indicated by more green or red polygons, respectively.

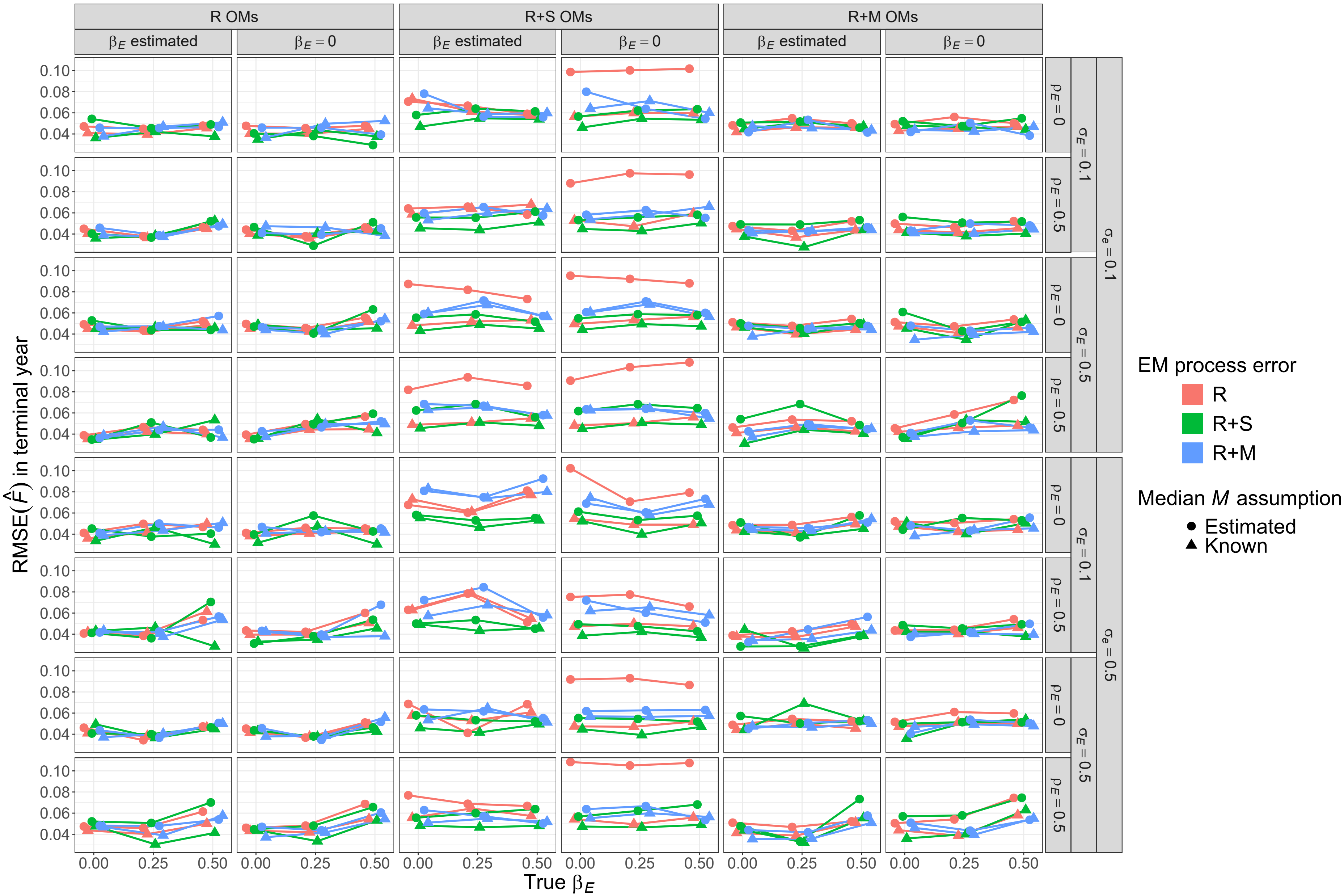

Fig. S66. Root mean square error (RMSE) of estimates of fully-selected fishing mortality (F) in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known). All OMs had low population observation error and contrast in fishing mortality.

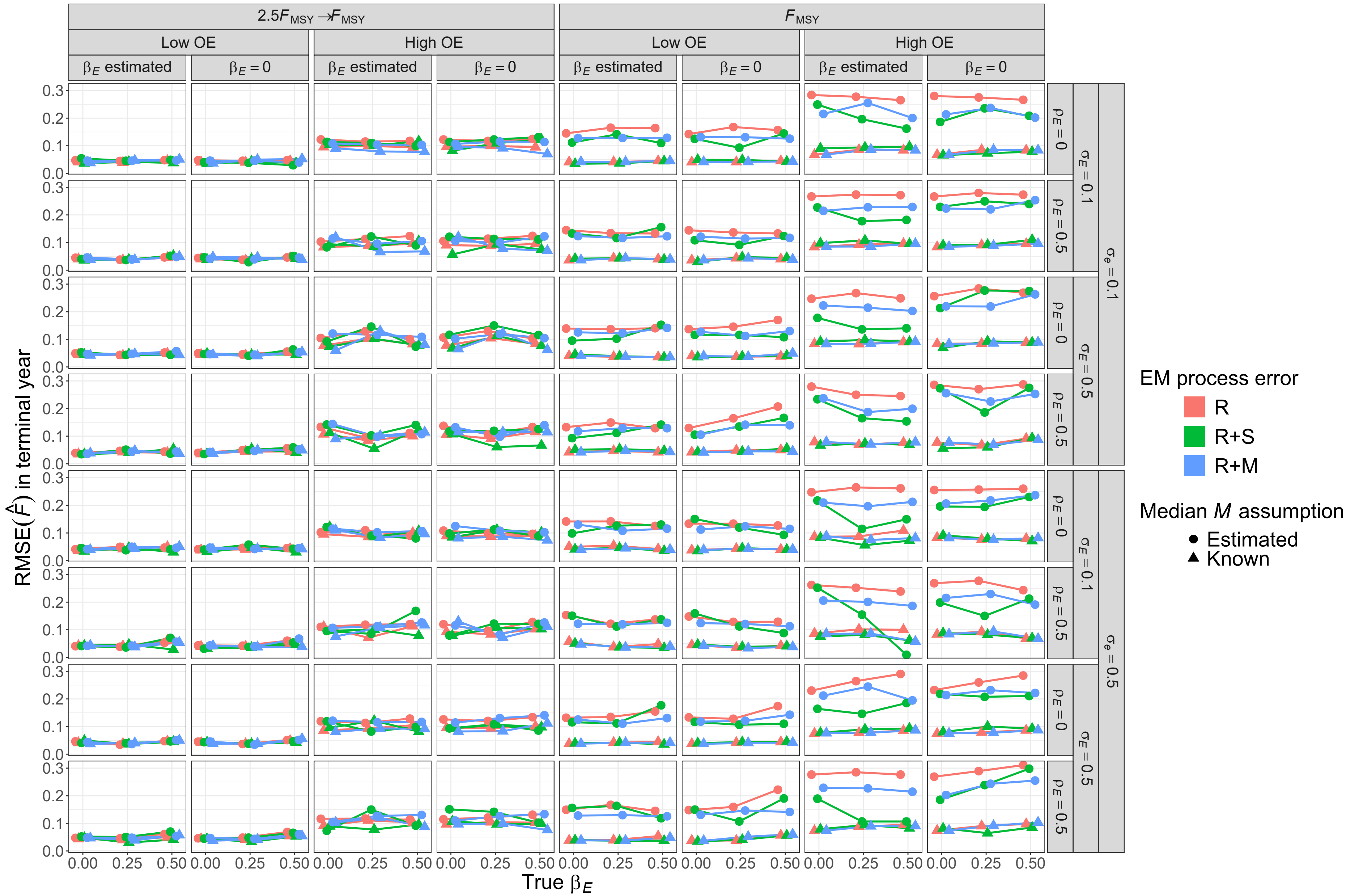

Fig. S67. For R OMs, root mean square error (RMSE) of estimates of fully-selected fishing mortality (F) in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known).

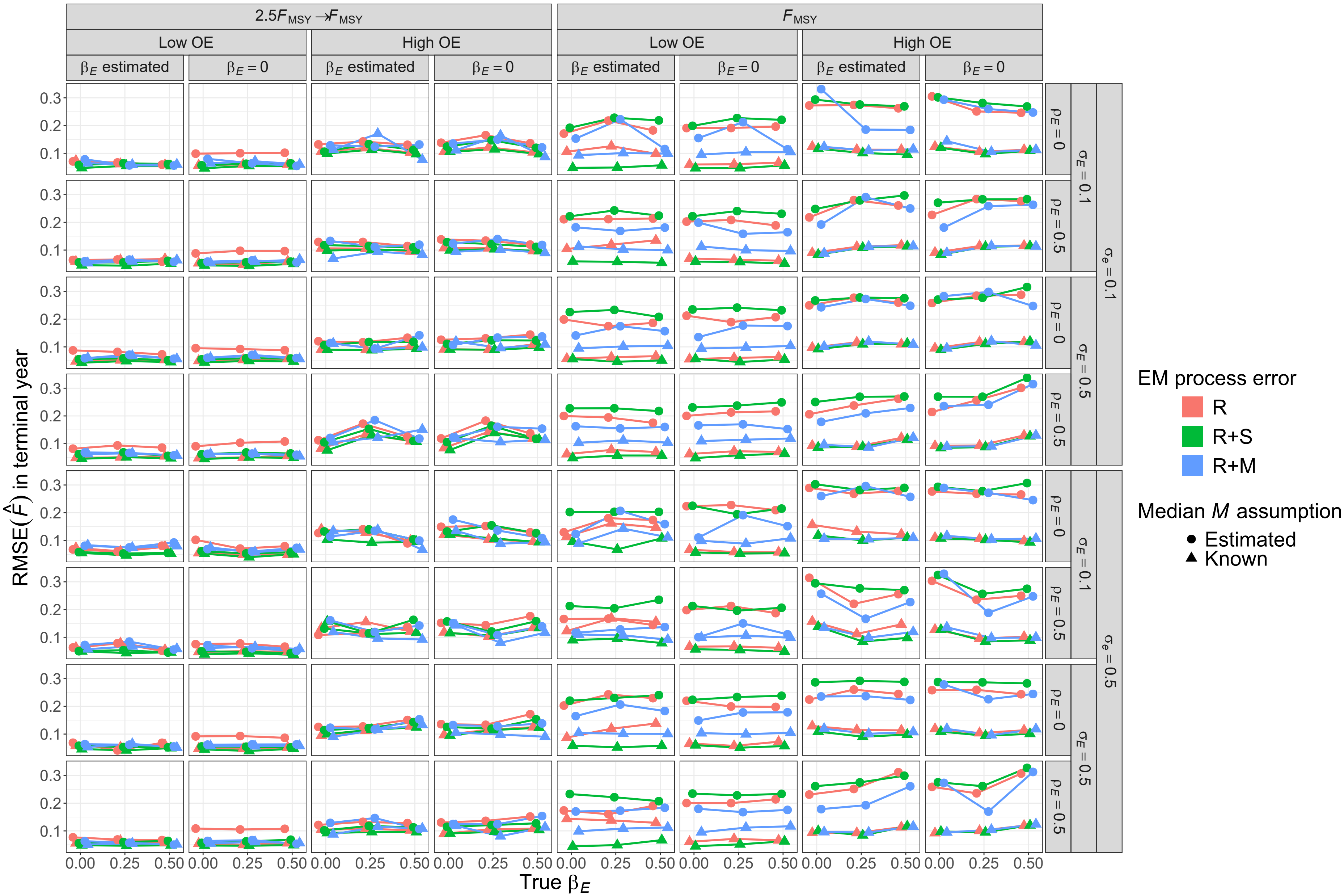

Fig. S68. For R+S OMs, root mean square error (RMSE) of estimates of fully-selected fishing mortality (F) in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known).

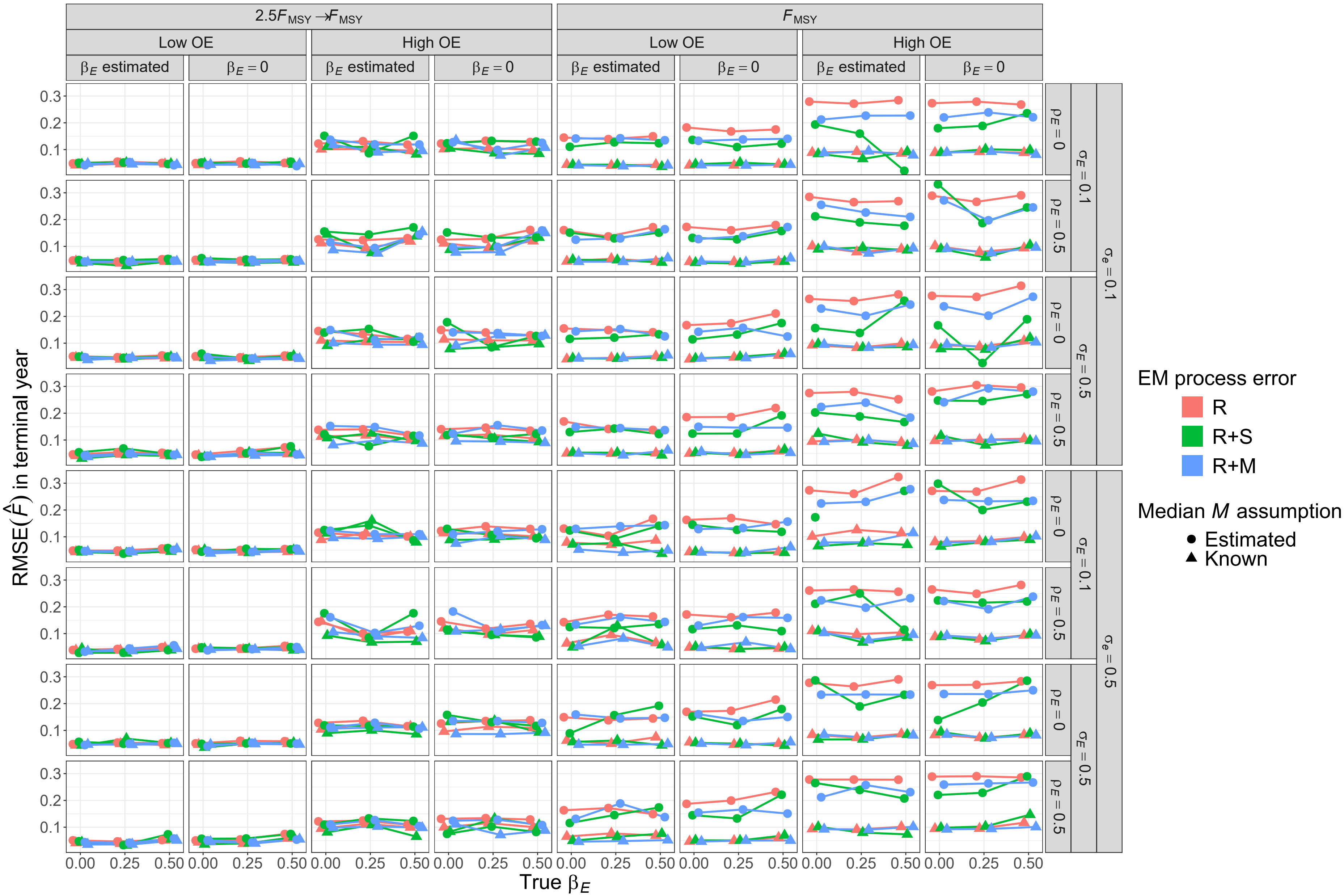

Fig. S69. For R+M OMs, root mean square error (RMSE) of estimates of fully-selected fishing mortality (F) in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known).

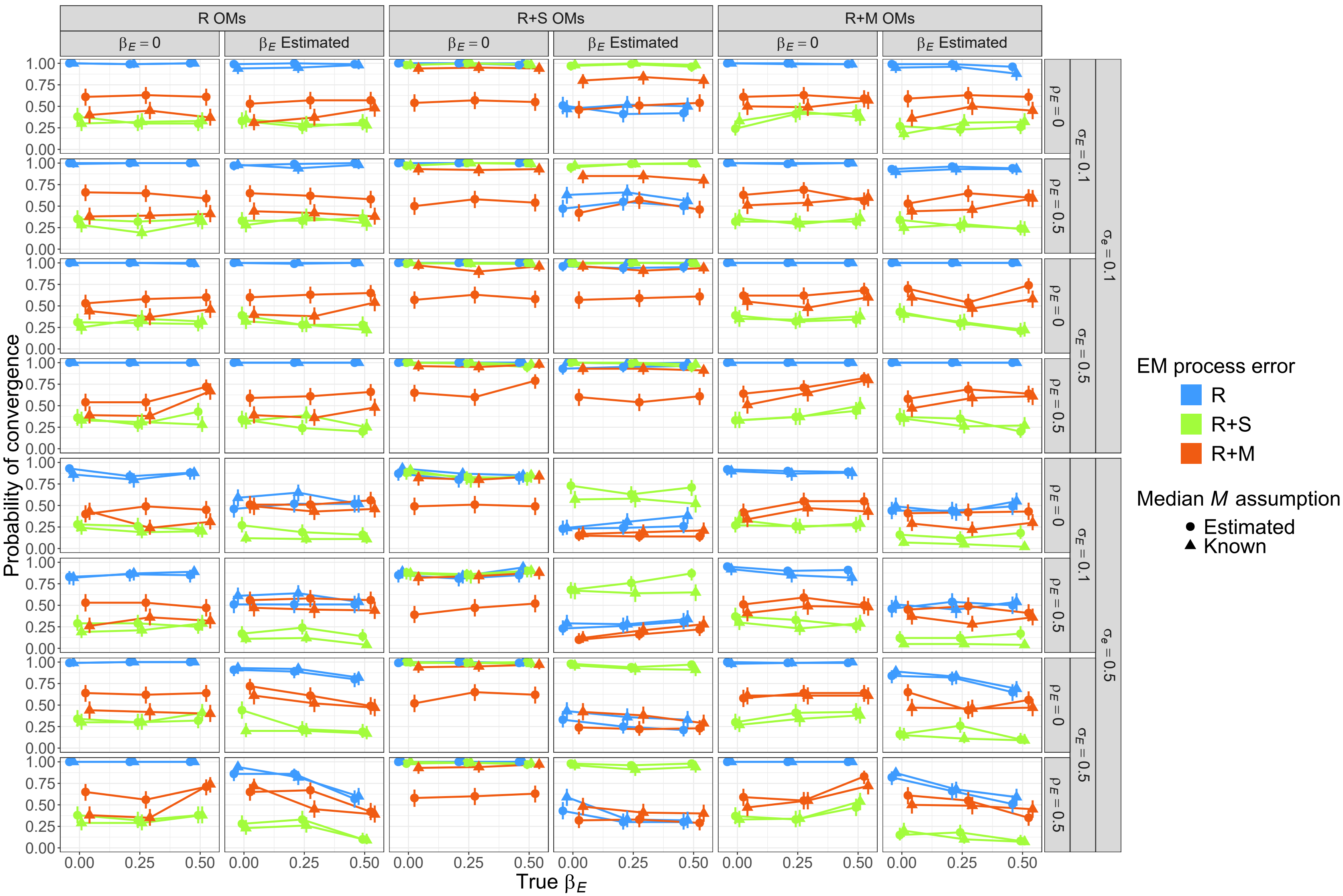

Fig. S70. Estimated probability of fits providing Hessian-based standard errors for EMs with alternative process error assumptions, treatment of median natural mortality (β_M known or estimated), and treatment of covariate effect ($\beta_E = 0$ or estimated). The OMs have R (left) and R+S (middle), or R+M (right) process error structures, alternative configurations of covariate time series structure and levels of observation uncertainty (rows), and three levels of true covariate effect on median natural mortality (x axis). All OMs had low observation error for fish population observations and temporal contrast in fishing pressure. Vertical lines represent 95% confidence intervals.

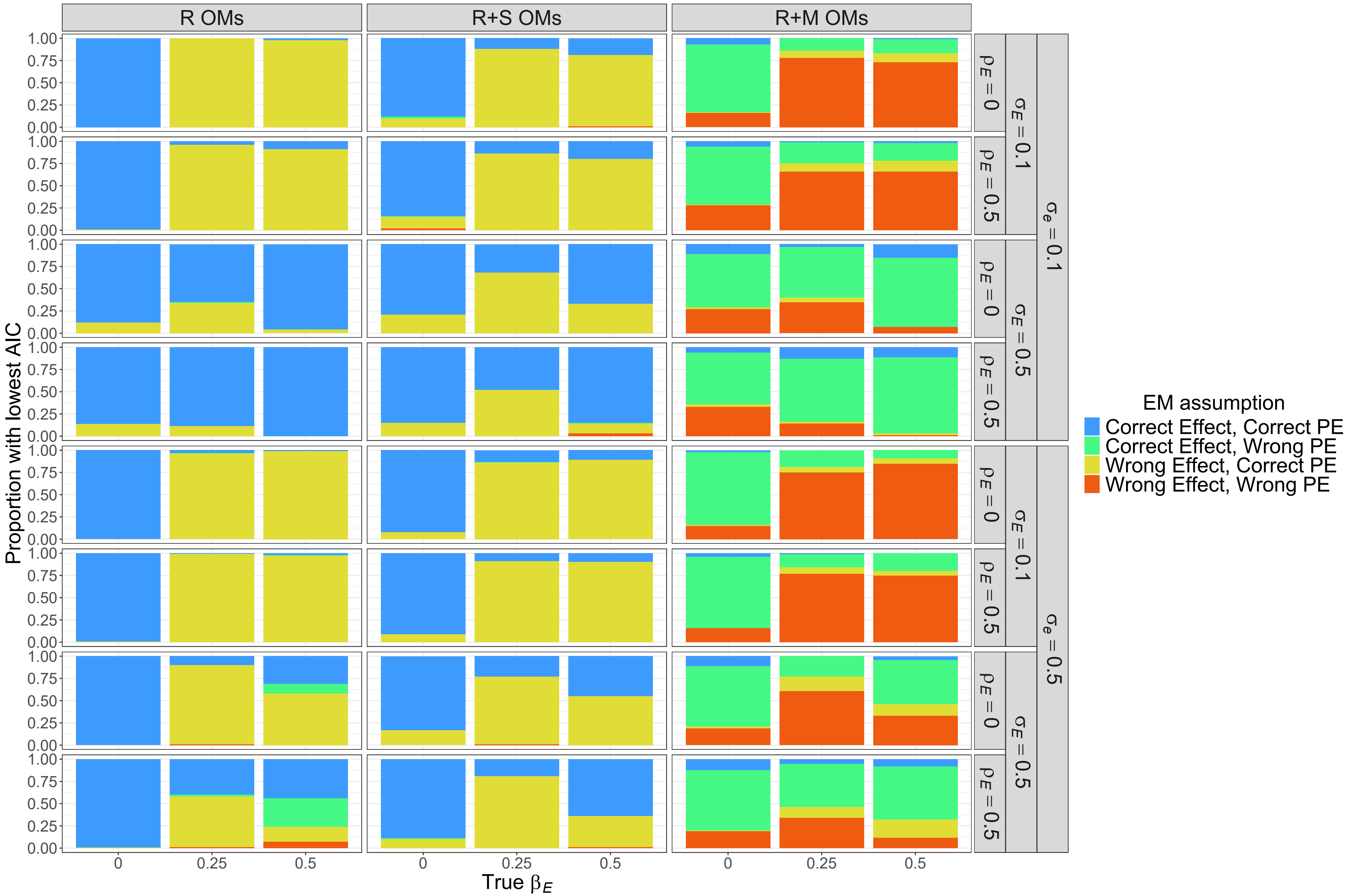

Fig. S71. For each OM, the proportion of simulated data sets where the EM type (treatment of environmental covariate effect and assumed process error type) had the lowest AIC. All OMs had low observation error for fish population observations and temporal contrast in fishing pressure. All EMs estimated median natural mortality rate.

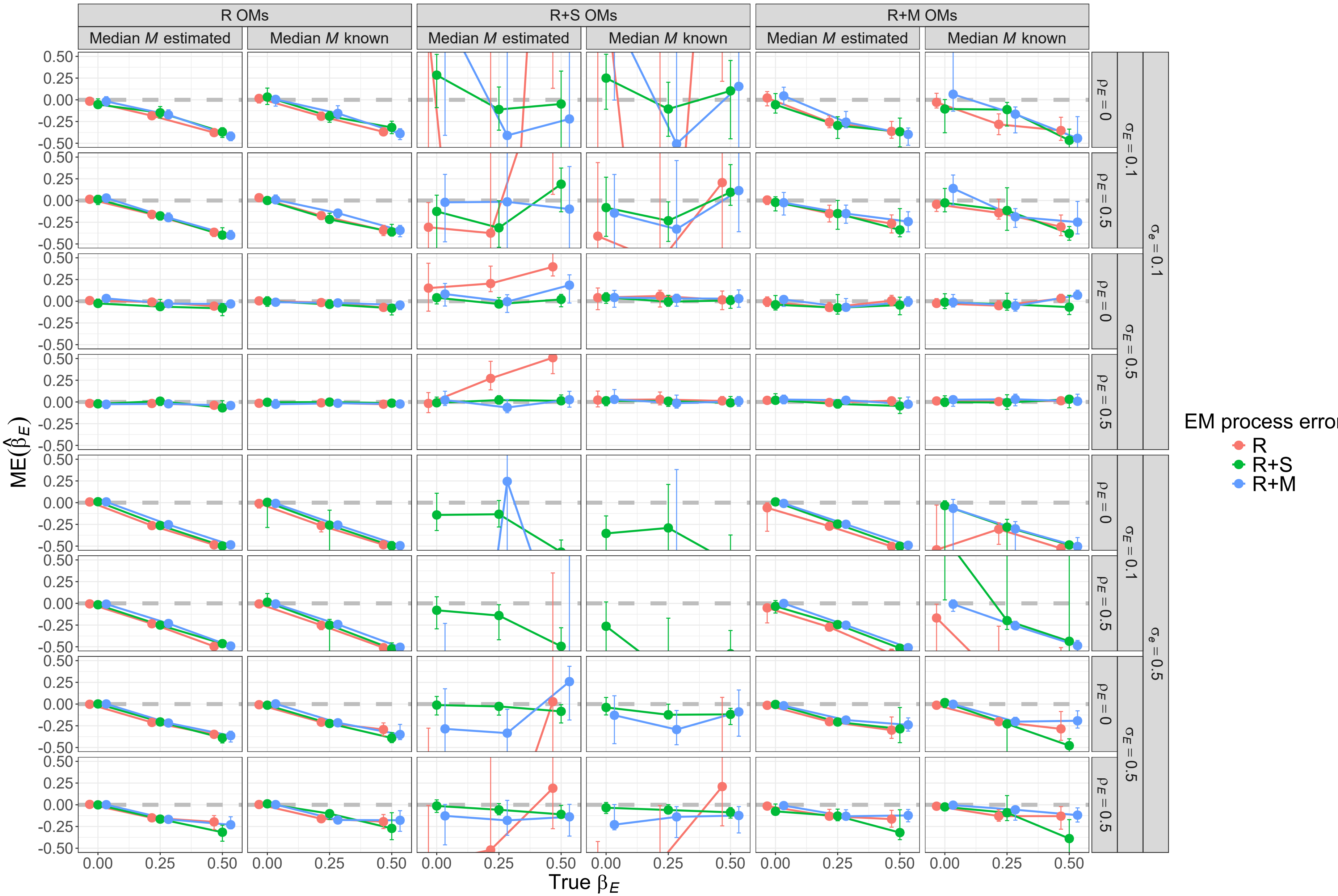

Fig. S72. Median error (ME) of estimates of environmental effect on natural mortality β_E from fitting EMs with alternative process error assumptions and treatment of median natural mortality (β_M known or estimated). All OMs had low observation error and contrast in fishing mortality. Vertical lines represent 95% confidence intervals.

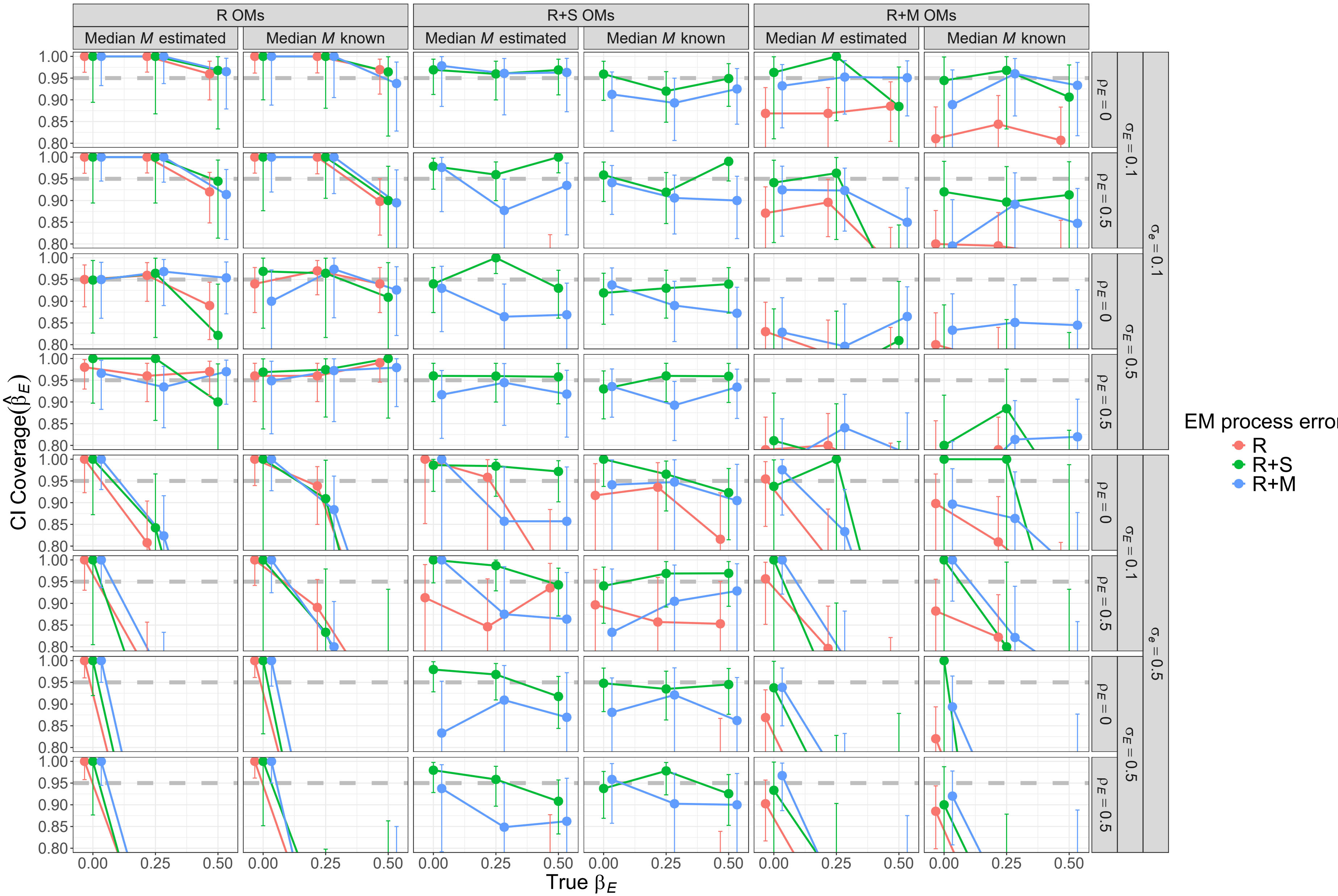

Fig. S73. Probability of 95% confidence interval for β_E containing the true value for EMs with alternative process error assumptions and treatment of median natural mortality (β_M known or estimated). All OMs had low observation error and contrast in fishing mortality. Vertical lines represent 95% confidence intervals.

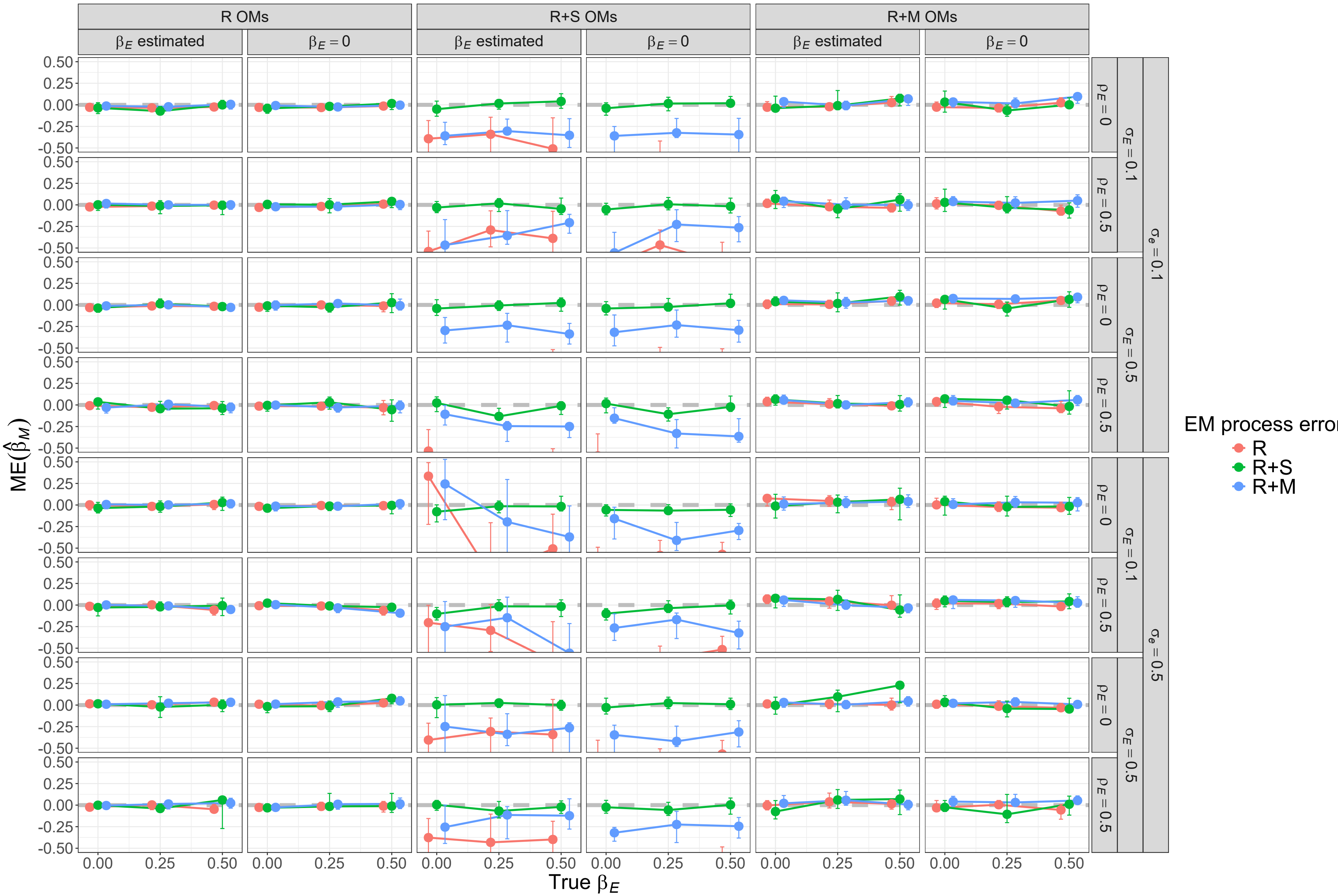

Fig. S74. Median error (ME) of estimates of β_M from fitting EMs with alternative process error assumptions and treatment of covariate effect ($\beta_E = 0$ or estimated). All OMs had low observation error and contrast in fishing mortality. Vertical lines represent 95% confidence intervals.

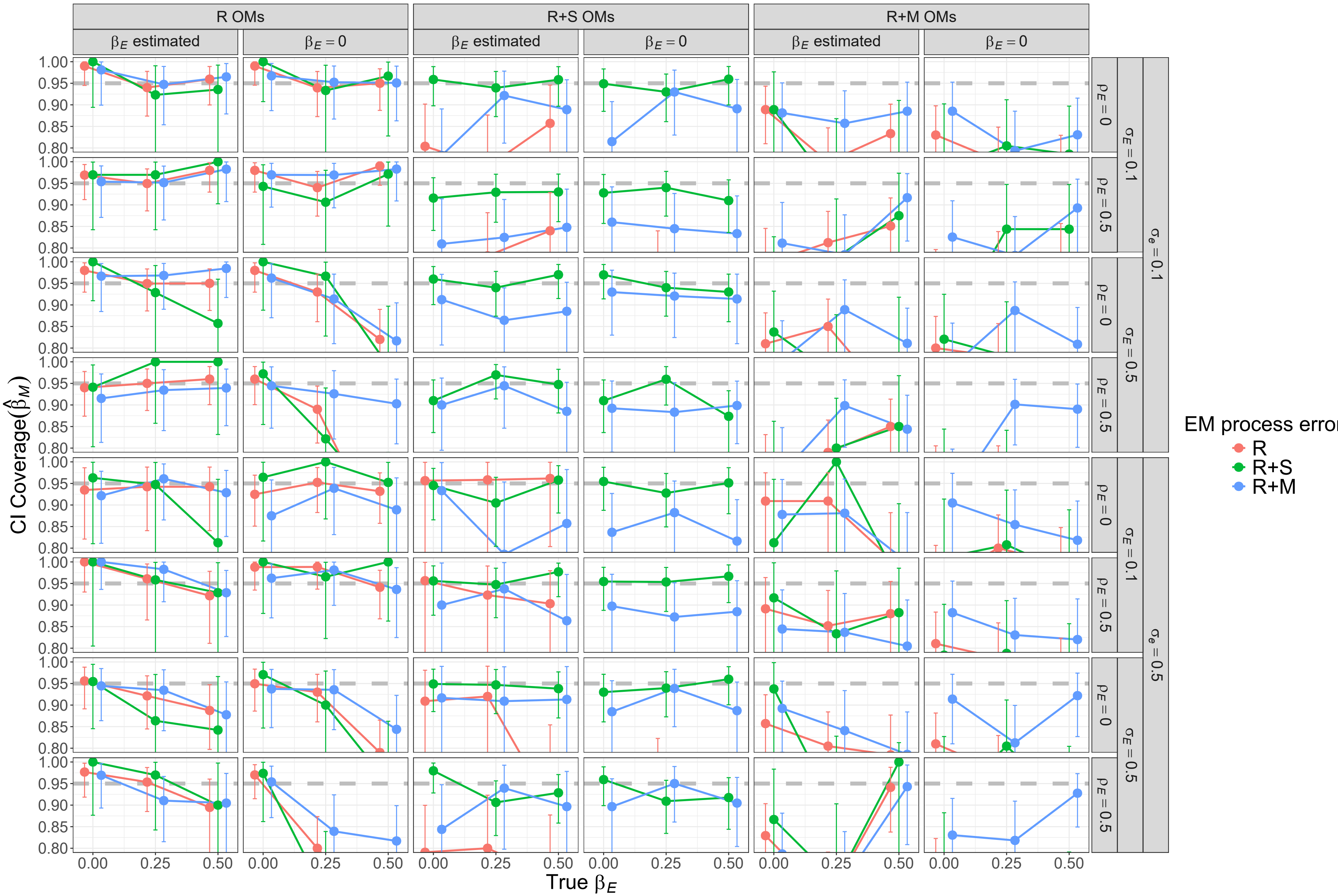

Fig. S75. Probability of 95% confidence interval for β_M containing the true value for EMs with alternative process error assumptions and treatment of covariate effect ($\beta_E = 0$ or estimated). All OMs had low observation error and contrast in fishing mortality. Vertical lines represent 95% confidence intervals.

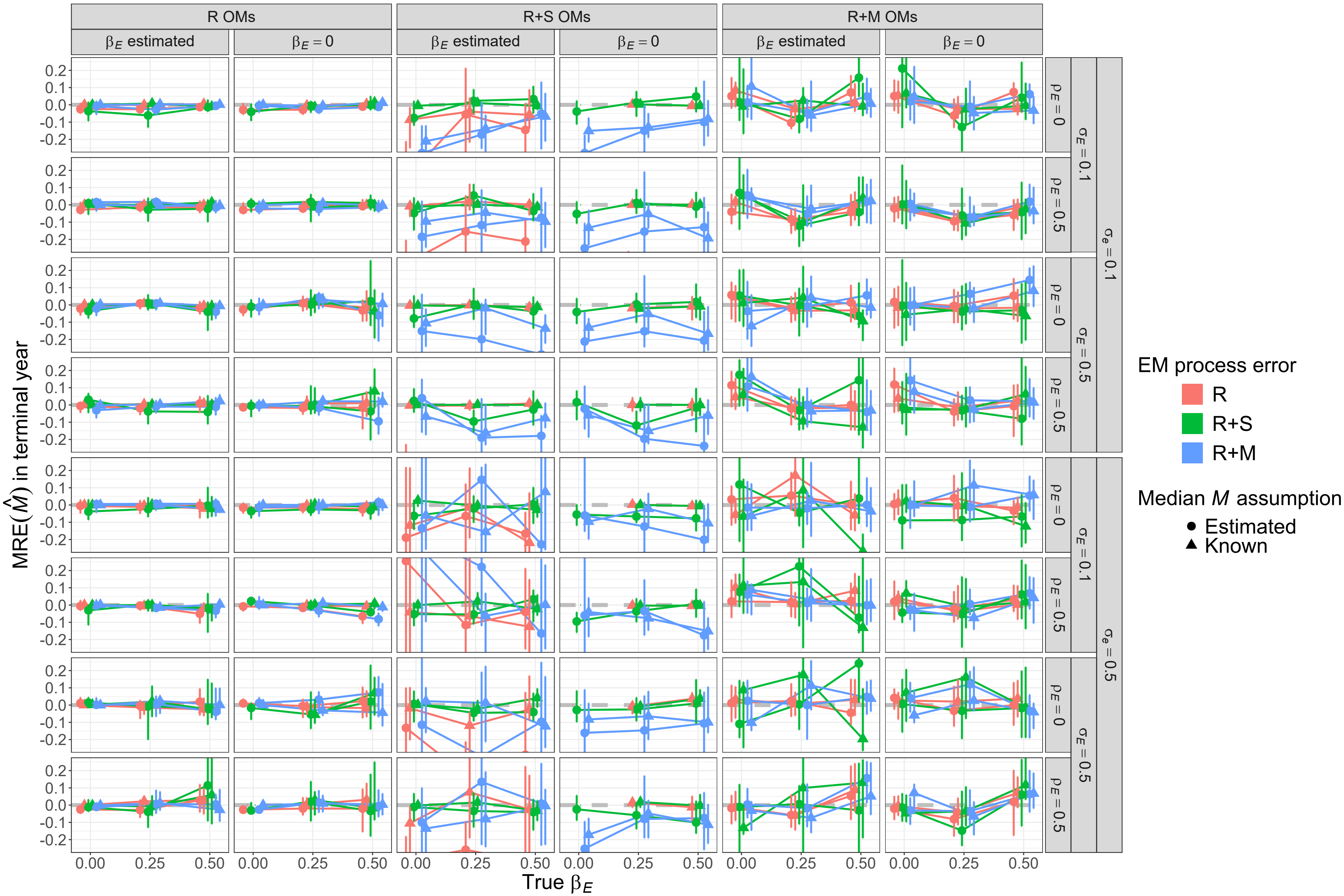

Fig. S76. Median relative error (MRE) of estimates of natural mortality rate in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known). All OMs had low observation error and contrast in fishing mortality.

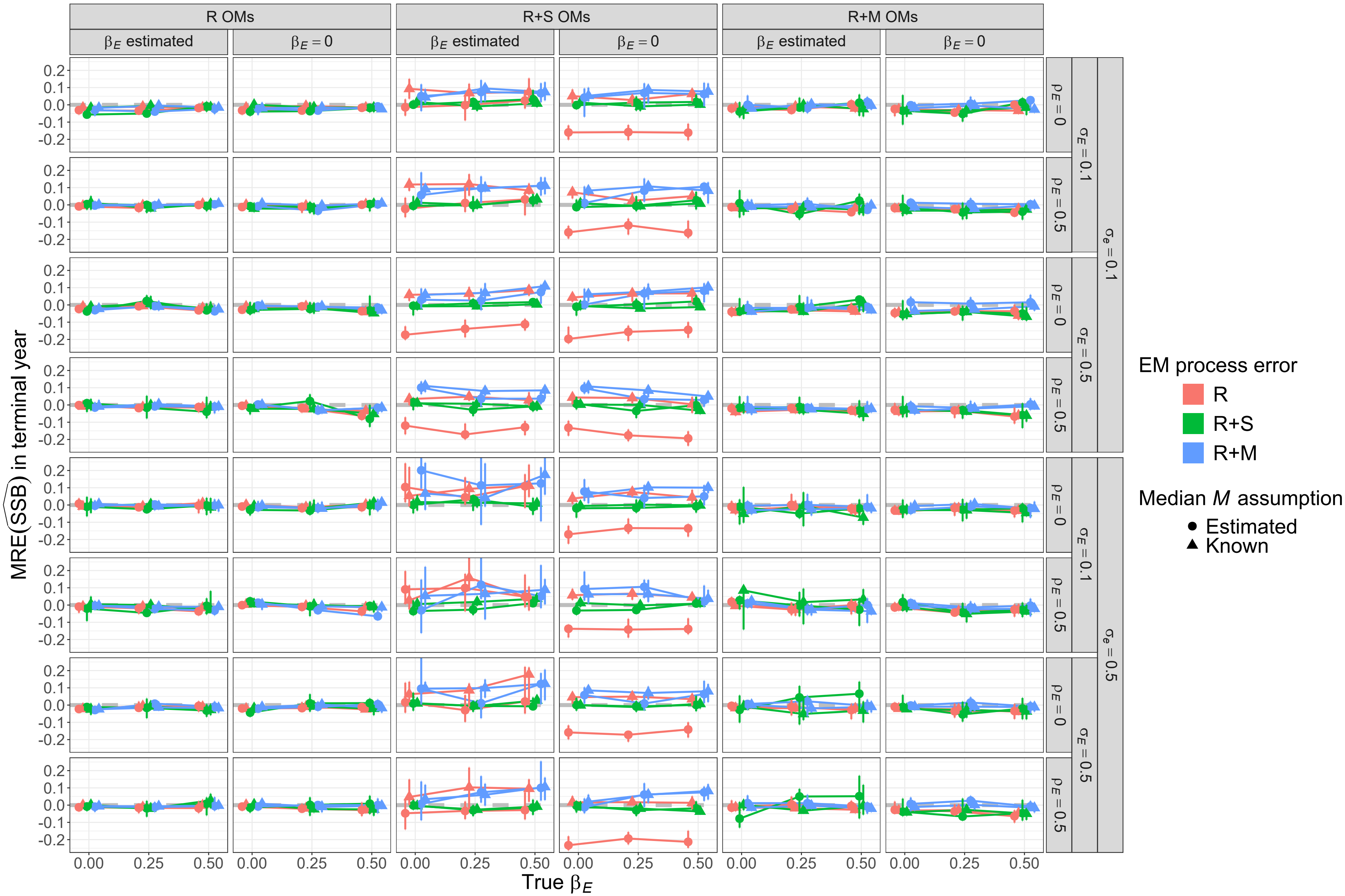

Fig. S77. Median relative error (MRE) of estimates of spawning stock biomass (SSB) in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known). All OMs had low observation error and contrast in fishing mortality.