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

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

Gavin Fay[4] Alex Hansell[2] Christopher M. Legault[2]

Brandon Muffley[5] John Wiedenmann[6]

29 May, 2025

[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 02740, 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

We completed a large scale simulation study with 288 operating models, each with 100 simulated data sets, and 12 estimated models fit to each simulated data set. The factors defining operating model configuration included the source of process error on the population (recruitment, recruitment and survival, recruitment and natural mortality), the degree of temporal variation and autocorrelation of the environmental covariate, the uncertainty in the observation of the covariate, the uncertainty in indices and age composition observations, fishing history, and the magnitude of the effect of the covariate on natural mortality. The estimating models make alternative assumptions on whether to include the environmental effect, whether the mean/intercept log natural mortality (log(0.2)) is estimated or known, and whether process errors are on just recruitment, recruitment and survival, or recruitment and natural mortality.

We found convergence of all estimation models was generally best when operating models assumed process errors in recruitment and survival, constant fishing rate, greater contrast in the true environmental covariate, and lower uncertainty in corresponding observations. Reliable convergence of all estimating models also occurred with the same process errors in the operating model and a step-change in fishing, but this also required lower uncertainty in index and age composition observations. Estimating models with process errors on recruitment and survival were unlikely to converge when the process errors in the operating model did not match whereas estimating models with process errors in recruitment and natural mortality converged for operating models without this match in certain cases. Probability of convergence generally decreased when the mean/intercept log-natural mortality rate parameter was estimated.

Whether the mean log-natural mortality was estimated or not, the best accuracy of AIC for model selection occurred for models with process errors on recruitment and survival. AIC accuracy was poor for models with process errors on recruitment and natural mortality.

estimating the mean natural mortality rate had small effects on the accuracy of AIC in selecting the appropriate process error. Estimating the mean log-natural mortality resulted in a small decrease in AIC accuracy. AIC was conservative for determining whether the environmental covariate affected natural mortality. AIC was very accurate in determining no effect when there was no effect in the operating model, but AIC often ranked the null model best when there was an effect. Accuracy of AIC for covariate effects improved with increased effect size, increased temporal contrast in the covariate, and lower uncertainty in observations.

We found no evidence of bias in estimation of environmental effects regardless of process error assumptions when there was low uncertainty in the environmental observations and large constrast in the environmental covariate. In most cases the relative error of the estimated environmental effect did not depend on the source of process error assumed in the estimating model. The worst bias was observed when OMs assumed R+S process errors, high uncertainty in covariate observations, low variability in the covariate, and low uncertainty in index and age composition observations. Simultaneously estimating the mean/intercept log natural mortality resulted in larger variation in the relative errors of the estimated environmental effect. Estimation of the intercept was reliable for all EM process error assumptions when the operating models assumed process errors on recruitment and natural mortality, contrast in fishing pressure over time, and lower observation error. Estimating the mean/intercept for log natural mortality generally resulted in highly variable estimates of annual natural mortality and spawning biomass and evidence of bias for some operating and estimation model assumptions about process error source. Again reliability of annual natural mortality estimates was generally improved with lower observation error uncertainty and contrast in fishing pressure.

Reliable detection of covariate effects requires informative data. AIC preferred simpler models than the true model when information content in data and contrast in covariates and abundance were low. The null model for environmental covariate effects (no covariate effect) was selected when contrast in the time series was low and/or uncertainty in observations was high. The selection of the null model by AIC also likely decreases with strength of the effect of the covariate on M. Similarly, when there was process error in recruitment and natural mortality, estimation models with process error only in recruitment were preferred presumably due to low variation in simulated natural mortality process errors. Covariate effect estimation can be robust to process error assumptions with high contrast in covariate and low observation error.

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 formulation treating latent population characteristics as statistical time series with periodic observations that also may have error due to sampling or other sources of measurement error and therefore separating these sources of biological and measurement variability. 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 ways we make inferences and the effects of various factors on reliabilty of inferences from state-space assessment models is growing [Li et al. (2024);Miller et al. (In reviewa);cadiganetal]. The importance of contrast in population size and fishing mortality and quality of data used to fit assessment models including the state-space variety is known (Magnusson and Hilborn, 2007; Miller et al., In reviewa). Furthermore, estimation of natural mortality, and even temporal variablity is possible in many scenarios (Lee et al., 2011; Johnson et al., 2016; Cadigan, 2016; Miller and Hyun, 2018; Miller et al., In reviewa).

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

Temporal and environmental effects on growth and weight at age are also more important for short term projections and the reliability of estimation of those effects (Correa et al., 2023, In review).

However, temporal variation in natural mortality and covariate effects on natural mortality are less studied Cadigan (2016). State-space assessment models currently used can treat the changes in cohort abundance over time as random effects and or despite the importance of natural mortality in inferences for the size of fish populations, their productivity and projections necessary for making catch advice. In fact, natural mortality plays a more significant role in short term projections than recruitment due to immediate effects on older age classes that constitute spawning biomass and catch.

See Miller et al. (In reviewa) for relevance of project 0 results (estimability of natural mortality). Some difficulty in distinguishing variation in natural mortality or effects of explicit covariates when variation in M random effects or Ecov is low relative to observation error.

Deriso et al. (2008) formulated the same natural mortality model as a function of covariates and random effects as we use in WHAM.

Mis-specified temporal population process errors could lead to misidentification of stock status and biased population estimates and poor fisheries management decisions (Trijoulet et al., 2020; Legault and Palmer, 2016; Szuwalski et al., 2018; Cronin-Fine and Punt, 2021; Liljestrand et al., 2024).

In the present study, we conduct a simulation study with operating models (OMs) varying by degree of observation error, source and variability of process error, and fishing history. The simulations from these OMs are fitted with estimation models (EMs) that make alternative assumptions for sources of process error, whether a SRR was estimated, and whether natural mortality is estimated. Given the confounding nature of process errors, developing diagnostic tools to detect model misspecification is of great scientific interest and could aid the next generation of stock assessments (Auger-Méthé et al., 2021). We evaluate whether convergence and Akaike Information Criterion (AIC) can correctly determine the source of process error and the existence of a SRR. We also evaluate when retrospective patterns occur and the degree of bias in the outputs of the assessment model that are important for management.

Estimation of natural mortality is known to be challenging in stock assessment models (Lee et al. (2011), others). However, estimation has been shown to be reliable in some situations (Miller and Hyun, 2018)

Miller et al in review (project 0) found estimation of natural mortality was feasible when data are good and contrast in fishing pressure etc.

Let's focus results here on those situations so that we can reduce the plots in figs.

Variation in natural mortality has an immediate impact on projections unlike recruitment. Understanding this varation in M leads to better understanding of post-recruit productivity and therefore management.

Here we conduct a simulation study with 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 natural mortality. The simulations from these operating models are fitted with estimating models that make alternative assumptions for sources of process error, and whether (mean) M is estimated. We evaluate effects of these factors on convergence of fitted models, whether AIC can correctly determine the correct source of process error and correct assumption about covariate effects on natural mortality, and the degree of bias in relevant parameters and outputs of the assessment model.

Methods

All of our analyses used the Woods Hole Assessment Model (WHAM) to construct both OMs and EMs (Miller and Stock, 2020; Stock and Miller, 2021; Miller et al., In reviewb). 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., In review; Li et al., In reviewa) 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 for to generate all results.

We completed a simulation study with a 288 operating models. The factors defining the configuration of each operating model which are 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), latent environmental covariate process stochasticity (4 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 estimating models (EMs). The factors that distinguish the estimating models which are also described in detail below include source of population process error type (3 levels) whether (median) natural mortality rate was estimated or assumed known (2 levels), and whether the effect of the environmental covariate on natural mortality 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 changes in cohort abundance over time (R+S), or recruitment and natural mortality (R+M). We did not use the log-normal bias-correction feature for process errors or observations described by Stock and Miller (2021) for operating and EMs (Li et al., In reviewb). 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 configuration of the latent environmental covariate in the operating models assume the 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 S2 provides example simulations of the latent process 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 S1, top left), growth (Figure S1, top right), time of spawning (1/4 of the year), and recruitment (Figure S1, bottom right) are identical to Miller et al. (In reviewa). The maturity at age is a logistic function of age 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 $\theta_2 = 3.2$.

Here we consider both covariate effects and process errors for natural mortality and the general model for natural mortality in year y is a log-linear function of a parameter β_M that defines median natural mortality and possibly effects of the environmental covariate X_y described above and autocorrelated normally distributed annual process error $\varepsilon_{M,y}$

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

$\epsilon_{M,y} | \epsilon_{M,y-1} \sim \mathrm{N}\left(\epsilon_{M,y-1}, (1-\rho_M^2)\sigma_M^2\right)$ (Stock and Miller, 2021). We assume the median natural mortality rate $e^{\beta_M} = 0.2$ is constant across ages. For R and R+S OMs and EMs, $\epsilon_{M,y} = 0$. For all R+M OMs, we assume the same standard deviation $\sigma_M = 0.3$ and 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 natural mortality random effects result in a range of different levels of variation in annual natural mortality rates (Figure S3).

We assumed expected recruitment each year 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 natural mortality at age) are constant in the R and R+S OMs without covariate effects on natural mortality. 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 reviewa), our assumed biological inputs and selectivity (defined below) with constant natural mortality result in equilibrium fishing mortality 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 $F_{\text{MSY}} = F_{[\%}40]$ 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 natural mortality, 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 process for M. For operating models with time-varying random effects for fishery selectivity, F_{MSY} is also not constant however we use the same fishing history as other operating models which corresponds to F_{MSY} at the mean selectivity parameters.

We also used the same two fishing scenarios as Miller et al. (In reviewa) for OMs. In the first scenario, the stock experiences overfishing at $2.5F_{\text{MSY}}$ for the first 20 years followed by fishing at F_{MSY} for the last 20 years (denoted $2.5F_{\text{MSY}} \rightarrow F_{\text{MSY}}$). In the second scenario, the stock is fished at F_{MSY} for the entire time period (40 years).

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 reviewa) for R+M OMs and other OMs with fishery selectivity and index catchability process errors. For R+S OMs, cohort temporal transition process errors were uncorrelated with $\sigma_{2+} = 0.3$

Catch and index observations

We define the generation of observations of total catch, aggregate indices, and corresponding age composition identical to Miller et al. (In reviewa). 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 S1, 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 aer 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 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, standard deviation for log-aggregate observations was assumed known whereas that for the logistic-normal age composition observations was estimated.

Estimating models

For each data set simulated from an operating model 12 estimating models were fit. There were three factors defining the configuration of each estimating model: 1) whether the mean natural mortality 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 natural mortality (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 to ensure comparability of AIC. All fixed effects parameters for selectivity, catchability, fully-selected fishing mortality, 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 natural mortality were also estimated. The observation error variance of the environmental observations and aggregate catch and indices were all assumed known at the true values.

Performance measures

EM convergence

The first measure of reliability we investigated was frequency of convergence when fitting each estimating model to the simulated data sets. There are various ways to assess convergence of the fit, but we defined successful convergence as the hessian of the marginal log-likelihood being invertible and providing variance estimates for the fixed effects parameters.

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 (mean) natural mortality rate or it is assumed at the true value $\beta_M = log(0.2)$. 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 $n\left(\boldsymbol{\theta}\right)$ estimated,

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

All model fits that successfully completed the optimization were used in these results. We used all of these fits 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 natural mortality random effects and the estimated variance parameter going to zero would cause poor convergence.

Parameter inference reliability

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 calculated median errors (ME) of

- β_E, the effect of environmental covariates on natural mortality,
- β_M, the mean log-natural mortality rate,

and the median relative errors (MRE) of terminal year

- natural mortality rate ,
- spawning stock biomass, and
- fully-selected fishing mortality rate.

In preliminary analyses we inspected MRE of these parameters over the whole time series, but we found no appreciable differences in patterns across the various factors defining the OMs and EMs.

We also calculated the root mean square error (RMSE) and estimated probability of coverage of constructed 95% confidence intervals for β_E and β_M for EMs that estimated these parameters.

Results for fishing mortality rate are provided in the Supplementary Materials. For the ith simulated data the relative error for a parameter θ provided from the fitted estimation model is

$$\mathrm{RE}_i(\theta) = \frac{\widehat{\theta}_i - \theta_i}{\theta_i}$$

and constructed 95% confidence intervals using the binomial distribution approach as in Stock and Miller (2021) and Miller et al. (In reviewa). Because of the inverse relationship of the the confidence interval and the null hypothesis significance test, when the confidence interval contains zero we do not have evidence of bias from the simulation study.

For natural mortality rate, SSB, and fully-selected fishing mortality rate, we summarized results for each of the annual values, but we present results only for estimates from the terminal year, because there were no appreciable differences between the results among all years.

Results

Because Miller et al. (In reviewa) found inferences are most reliable for scenarios with low observation error for indices and age composition data and with temporal contrast in fishing pressure, we restrict our attention to scenarios with these characteristics in the main paper (temporal contrast in fishing and low observation error), but include corresponding results for the other scenarios in the Supplementary Materials.

EM Convergence

When the median natural mortality rate was assumed known, corresponding to usual practice in application of assessment models, good convergence for all EMs was observed for R+S operating models with F_{MSY} fishing history, more variation in the latent environmental covariate ($\sigma_E = 0.5$), and lower error in the associated observations $\sigma_e = 0.1$ (Figure 1).

Good convergence for all EMs was also observed for R+S OMs with the step change in fishing history, but with low error in indices and age comp observations.

EMs that assumed random effects just on recruitment (R), nearly always converged across all operating models that assumed more variation in the latent environmental covariate ($\sigma_E = 0.5$) and lower error in the associated observations $\sigma_E = 0.1$. EMs that assumed recruitment and survival random effects (R+S), had poor converge probability when the OMs had alternative process errors (R or R+M). The R+S EMs showed best convergence for the R+S OMs where there was lower error in the environmental observations $\sigma_E = 0.1$ or higher error in environmental observations of latent true environmental covariates with greater temporal variation. EMs that assumed recruitment and natural mortality random effects (R+M), had poor convergence probability for all OMs that had a step change in fishing history and higher uncertainty in indices and age composition.

When the OMs and EMs both assume process errors only on recruitment (R) or on recruitment and survival (R+S), convergence was worst with less variation in the true environmental covariate and larger uncertainty in associated observations. When the OMs and EMs both assume process errors on recruitment and natural mortality (R+M) convergence was problematic for all OMs with the step-change in fishing history. The best convergence was observed with this match between OMs and EMs was when the fishing history was constant and there was low uncertainty in environmental observations.

As might be expected, there was an overall drop in the probability of convergence when the mean natural mortality rate was estimated rather than assumed at the true value (Figure 1). Otherwise, general trends described above with mean natural mortality fixed, apply when estimated.

AIC performance

We only present results for EMs where the median natural mortality rate parameter was estimated. Results differed little between results where this parameter was estimated or assumed known, but all of these results can be found in the Supplementary materials.

When fitting the EMs to R and R+S OMs, AIC was generally accurate in determining the correct source of process errors and, when no effect of the covariate was simulated in the OMs, AIC was also accurate for the correct treatment of the environmental covariate $\beta_E = 0$ (Figs. 2, S7, S8). However, when OMs included an effect of the covariate AIC accuracy for estimating an effect in the EM was only found when the OMs also included lower covariate observation error and higher effect size and higher temporal variability in the covariate. On the other hand, for R+M Oms, AIC generally was not able to accurately determine the correct source of process errors, but was able to accurately determine the correct treatment of the environmental covariate (Figs. 2 and S9).

When estimating models assumed the mean natural mortality rate was known, the best accuracy of AIC for model selection occurred for models with R+S process errors. R+S estimating models ranked best with very low frequency for R or R+M operating models and with very high frequency for R+S operating models (Figure 2). R estimating models were determined best with high frequency for R operating models, but also for R+M operating models. R+M estimating models were rarely determined best for any operating models including those where the process errors matched.

AIC was conservative for determining whether the environmental covariate affected natural mortality. AIC was highly accurate in determining no effect when there was no effect in the operating model, but AIC ranked the null model best with high frequency even when there was an effect in the operating model in many cases. However the accuracy of AIC improved in certain operating models. Increased effect size, increased temporal contrast in the covariate, and lower uncertainty in all observations types lead to increased accuracy of determining covariate effects.

Relative to the assumption that the mean natural mortality rate was known, estimating the mean natural mortality rate had small effects on the accuracy of AIC in selecting the appropriate process error and whether the covariate affect natural mortality (Figure 2). Where there were differences there were small decreases in accuracy for determining the appropriate process error and determining a covariate effect when there was one.

Bias

Environmental effect

When the EMs assumed the median natural mortality rate was known, we observed generally accurate estimation of environmental effects across all EM and OM process error assumptions and all true covariate effect sizes, when ther was low uncertainty in environmental observations and larger temporal contrast in the simulated true environmental covariate (Figure 3). We observed a negative trend in bias of the environmental effect with increased effect size when temporal variation in the covariate was lower and/or uncertainty in the covariate observations was higher. When the OMs had R+S process errors with low temporal variation in the true environmental covariate and lower uncertainty in the indices age age composition, estimated covariate effects were highly variable. In most cases the relative error of β_E did not depend on the source of process error assumed in the EM. When there was an effect of the EM process error assumption it was when OMs had R+S process errors. The worst bias was observed when OMs assumed R+S process errors, high uncertainty in covariate observations, low variability in the covariate, and low uncertainty in index and age composition observations.

When the median natural mortality rate was estimated in the EM, results were similar except estimated effects were even more variable for data simulated with R+S process errors (Figure 3). There was also more separation of reliability of the estimation of the effect among EMs with different process error assumptions. The separation was most apparent when OMs simulated R+S process errors and larger variability in the environmental covariate wehere EMs with process errors other than R+S showing more bias than when the mean natural mortality rate was assumed known.

Median natural mortality rate

We found high accuracy for estimation of the median natural mortality rate parameter (β_M) for all EM process errors assumptions when OMs had step changes in fishing mortality, lower uncertainty in index and age composition observations, and either R or R+M process errors (Figure 5). The most variation in estimates occurred when fishing mortality was constant and there was higher uncertainty in index and age composition observations. For OMs with R+S process errors, the most reliable estimation of β_M was obtained when the EM also assumed R+S process errors across all other factors. For OMs with R+M process errors, the matching assumption for the EM only showed the best reliability when there was low uncertainty in index and age composition observations and fishing mortality was constant.

Annual natural morality rate

We present results for error in annual natural mortality rate conditional on three alternative EM configurations for the natural mortality parameters: 1) mean natural mortality rate parameter is fixed at the true value ($\beta_M = log(0.2)$) and no covariate effect is assumed ($\beta_E = 0$), 2) $\beta_M = log(0.2)$ and β_E is estimated, and 3) both β_M and β_E are estimated.

When OMs and EMs assume $\beta_M = log(0.2)$ and $\beta_E = 0$, there is no annual variation in natural mortality for OMs with R or R+S process errors simulated or assumed in the EM and, therefore, estimation bias is not possible (Figure 7). We also observe little or no evidence of bias (confidence intervals include 0) for R or R+S EMs for any OMs even when the OMs included an effect of the covariate on natural mortality ($\beta_E > 0$). Including process error on

M (R+M) produces more variability in errors for R and R+S EMs than including increased level of effect of the covariate on M. For R and R+S OMs, errors in annual M for R+S EMs were less variable than those for R EMs.

R+M EMs fit to R OMs with constant fishing mortality, or a step change in fishing mortality and lower uncertainty in index and age composition observations exhibited no bias in annual estimation of M indicating that the variance of the estimated random effects for M in these EMs collapsed to 0. R+M EMs were the only EMs to exhibit differences in the sign of the median errors across the time series. When OMs had a step change in fishing mortality and higher uncertainty in index and age composition observations, we observed positive median errors at the beginning of the time series and lower median errors in later years. However, whether there was evidence of bias, depended on the uncertainty in covariate observations and the degree of variability in the latent covariate. We observed negative median relative errors in annual M estimates for R+M EMs fit to R+S OMs with lower uncertainty in index and age composition observations across all levels of simulated effects of the covariate, but evidence of bias was strongest at the beginning of the time series.

When β_E was estimated, R EMs performed worse for R+S OMs with evidence of negative bias for OMs with constant fishing mortality, lower uncertainty in index and age composition observations, the largest covariate effect size, lower variability in the latent covariate, and larger uncertainty in the covariate observations (Figure 7). There was little difference in the results for R+S EMs whether β_E was estimated or not. The results for R+M EMs were generally similar to those when $\beta_E = 0$ was assumed, except that the differences between the estimates at the beginning of the time series and later years for certain OM configurations did not occur.

Allowing the EMs to estimate β_M resulted in very large variability in estimates for all EMs for R and R+M OMs with constant fishing mortality and higher uncertainty in index and age composition observations (Figure 7). The same variability occurred for R+S OMs, but strong bias was estimated for R and R+M EMs. Annual M estimation was most reliable when OMs had step changes in fishing mortality and lower uncertainty in index and age composition observations. For R and R+M OMs with those characteristics, all EMs generally provided accurate estimation of annual M, but only R+S EMs provided accurate estimation for R+S OMs. Little or no bias in annual M estimation was observed for R+S EMs across all OM process error assumptions as long as there was a step change in fishing mortality and lower uncertainty in indices and age composition.

Spawning stock biomass

Like natural mortality, we present results for error in spawning biomass conditional on three alternative EM configurations for the natural mortality parameters: 1) mean natural mortality rate parameter is fixed at the true value ($\beta_M = log(0.2)$) and no covariate effect is assmed ($\beta_E = 0$), 2) $\beta_M = log(0.2)$ and β_E is estimated, and 3) both β_M and β_E are estimated.

When EMs assume $\beta_M = log(0.2)$ and $\beta_E = 0$, there is little evidence of bias for R and R+M OMs except when there is a step change in fishing mortality and higher uncertainty in index and age composition observations, particularly at the end of the time series (Figure 8). We observed a similar trend in median relative error for R+S OMs, but there was more evidence of positive bias at the beginning of the time series whereas the confidence intervals for the negative median errors at the end of the time series often included 0. However, there was more indication of positive bias of the incorrect process error assumption of the EMs for the R+S OMs. For R+S EMs fit to R+S OMs, there was indication of small positive bias at the beginning of the time series when there was constant fishing mortality and higher uncertainty in index and age composition observations.

When EMs assume $\beta_M = log(0.2)$, but estimate β_E, the median errors for SSB for R and and R+M OMs are similar to those when $\beta_E = 0$ (Figure 8). For R+S OMs, the median errors for EMs that also assume R+S process errors are also similar to those with no covariate effect assumed. However, for R and R+M EMs fit to R+S OMs with constant fishing mortality rate and lower uncertainty in indices and age composition observations, we observed evidences of negative bias when higher covariate observation uncertainty and lower variation in the latent covariate was simulated and positive bias for other configurations of the covariate and corresponding observation uncertainty. When R+S OMs had constant fishing mortality rate and higher uncertainty in indices and age composition observations or step changes in fishing mortality and lower uncertainty in indices and age composition, R and R+M EMs often provided positively biased SSB estimates when there was lower uncertainty in covariate observations.

When EMs estimated both β_M and β_E, we observed large variation in the errors in SSB under the same OM configurations where we observed large variation in errors for annual natural mortality rates (Figure 8). Similarly, we observed little or no bias in SSB estimation for R+S EMs across all OM process error assumptions as long as there was a step change in fishing mortality and lower uncertainty in indices and age composition.

Discussion

We found estimation of environmental effects on M was possible and reliable in certain scenarios even when the process error was mis-specified (R+S EM and R+M OM) also R EMs were not reliable for Ecov effects when PE was mis-specified. Given that milleretal in review found AIC was accurate for R+S OMs and applications of wham to real assessment data typically finds R+S preferrable to R only, it suggests inferences on covariate effects on M without just R process errors (e.g., deriso et al 2008) should be reevaluated.

The estimating models assumed variances of aggregate catch and index observations was known. This approximation may be appropriate for indices where we have a reliable estimate of uncertainty based on the survey design (), but there may be better approaches for the aggregate catch such as an informed prior on the standard errors with realistic bounds.

We found EMs with R+M process errors were rarely determined as the appropriate model when OMs simulated R+M process errors. This was unexpected but is likely due to the size of the variance assumed for those process errors ($\sigma_M = 0.3$) relative to the variances assumed for index and age composition observations. This is related to the difficulty in separately estimating observation and process error variances when the ratio of process to observation uncertainty is low.

Note that the results for bias of covariate effect all assume ecov beta is estimated. The lack of bias in certain situations might suggest including the effect even if AIC says it isn't better at least when contrast in Ecov is high.

See project 0 paper for relevance of Li et al. (2024) and Liljestrand et al. (2024). See Miller et al. (In reviewa) fro relevance of project 0 results (estimability of natural mortality)

Deriso et al. (2008) was the first to model natural mortality as a function of explicit covariate and residual random annual variation. Any relevant points?

We conditioned our investigations of reliability of parameter estimation on simulations with fits that converged satisfactorily. However, these simulations could exclude many realizations of covariate, population, and observation time series that are not sufficiently informative for all the parameters that we were estimating. Why might this matter?

Poor quality of biological, fishery, and observation may adversely affect ability to distinguish between true process and observational error (Punt et al., 2014; Stewart and Monnahan, 2017; Cronin-Fine and Punt, 2021; Fisch et al., 2023; Li et al., In reviewa).

Uncertainty in estimated population attributes such as spawnign biomass can be increased considerably when natural mortality is estimated. This is also a consideration for estimating covariate effects on natural mortality even if the median natural mortality is assumed known.

The bias in natural mortality estimation resulted in biased estimation of stock size and likely harvest rates, but it will also propagate into biological reference points and possibly stock status.

The lack of bias observed for annual M, using R+S EMs when no effect was assumed must be due to the annual M being equal to the assumed value (0.2) on average across simulations because the simulated environmental covariate has mean 0. However, if one were to condition on the covariate time series we would expect biased estimation of annual M when there was a covariate effect.

Conclusions

As we would expect, reliable detection of covariate effects requires informative data. AIC preferred simpler models than the true model when information content in data and contrast in covariates and abundance were low. Null model for environmental covariate effect (no covariate effect) was selected when contrast in the time series was low and/or uncertainty in observations was high. Null selection likely decreases with strength of the effect on M. We only examined two non-zero effect sizes: 0.25, and 0.5, but our results suggest larger effect sizes, with the same observation error and contrast in time series, would allow better AIC performance for determining covariate effects. When there was process error in recruitment and M (R+M), models with process error only in recruitment were preferred. This could be because the variation in the M random effects was small relative to observation variability. The marginal variance for log M random effects was 0.3. Reliable estimation of environmental effect on M regardless of the process error assumed by the EM as long as the contrast in the covariate is sufficient and the uncertainty in the observations is low.

Acknowledgements

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

References

Auger-Méthé, M., Newman, K., Cole, D., Empacher, F., Gryba, R., King, A.A., Leos-Barajas, V., Mills Flemming, J., Nielsen, A., Petris, G. et al. 2021. A guide to state–space modeling of ecological time series. Ecological Monographs **91**(4): e01470. doi: 10.1002/ecm.1470.

Britten, G., Brooks, E.N., and Miller, T.J. In review. Identification and performance of environmentally-driven stock-recruitment relationships in state space assessment models. For Canadian Journal of Fisheries and Aquatic Sciences, State-space assessment models special collection.

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.

Correa, G., Monnahan, C.C., Miller, T.J., and Sullivan, J.Y. In review. Performance of age-only state-space assessment models under diverse somatic growth scenarios. For Canadian Journal of Fisheries and Aquatic Sciences, State-space assessment models special collection.

Correa, G.M., Monnahan, C.C., Sullivan, J.Y., Thorson, J.T., and Punt, A.E. 2023. Modelling time-varying growth in state-space stock assessments. ICES Journal of Marine Science **80**(7): 2036–2049. doi:10.1093/icesjms/fsad133.

Cronin-Fine, L. and Punt, A.E. 2021. Modeling time-varying selectivity in size-structured assessment models. Fisheries Research **239**: 105927.

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.

Fisch, N., Shertzer, K., Camp, E., Maunder, M., and Ahrens, R. 2023. Process and sampling variance within fisheries stock assessment models: estimability, likelihood choice, and the consequences of incorrect specification. ICES Journal of Marine Science **80**(8): 2125–2149. doi:10.1093/icesjms/fsad138.

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. In reviewa. 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 .

Li, C., Deroba, J.J., Miller, T.J., Legault, C.M., and Perretti, C. In reviewb. Guidance on bias-correction of log-normal random effects and observations in state-space assessment models. For Canadian Journal of Fisheries and Aquatic Sciences, State-space assessment models special collection.

Liljestrand, E.M., Bence, J.R., and Deroba, J.J. 2024. The effect of process variability and data quality on performance of a state-space stock assessment model. Fisheries Research **275**: 107023. doi:10.1016/j.fishres.2024.107023.

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.

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., Britten, G., Brooks, E.N., Fay, G., Hansell, A., Legault, C.M., Li, C., Muffley, B., Stock, B.C., and Wiedenmann, J. In reviewa. 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.

Miller, T.J., Curti, K.L., and Hansell, A.C. In reviewb. Space for wham: a multi-region, multi-stock generalization of the woods hole assessment model with an application to black sea bass. 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.

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.

Punt, A.E., Hurtado-Ferro, F., and Whitten, A.R. 2014. Model selection for selectivity in fisheries stock assessments. Fisheries Research **158**: 124–134. doi:10.1016/j.fishres.2013.06.003.

Stewart, I.J. and Monnahan, C.C. 2017. Implications of process error in selectivity for approaches to weighting compositional data in fisheries stock assessments. Fisheries Research **192**: 126–134. doi:10.1016/j.fishres.2016.06.018.

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.

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.

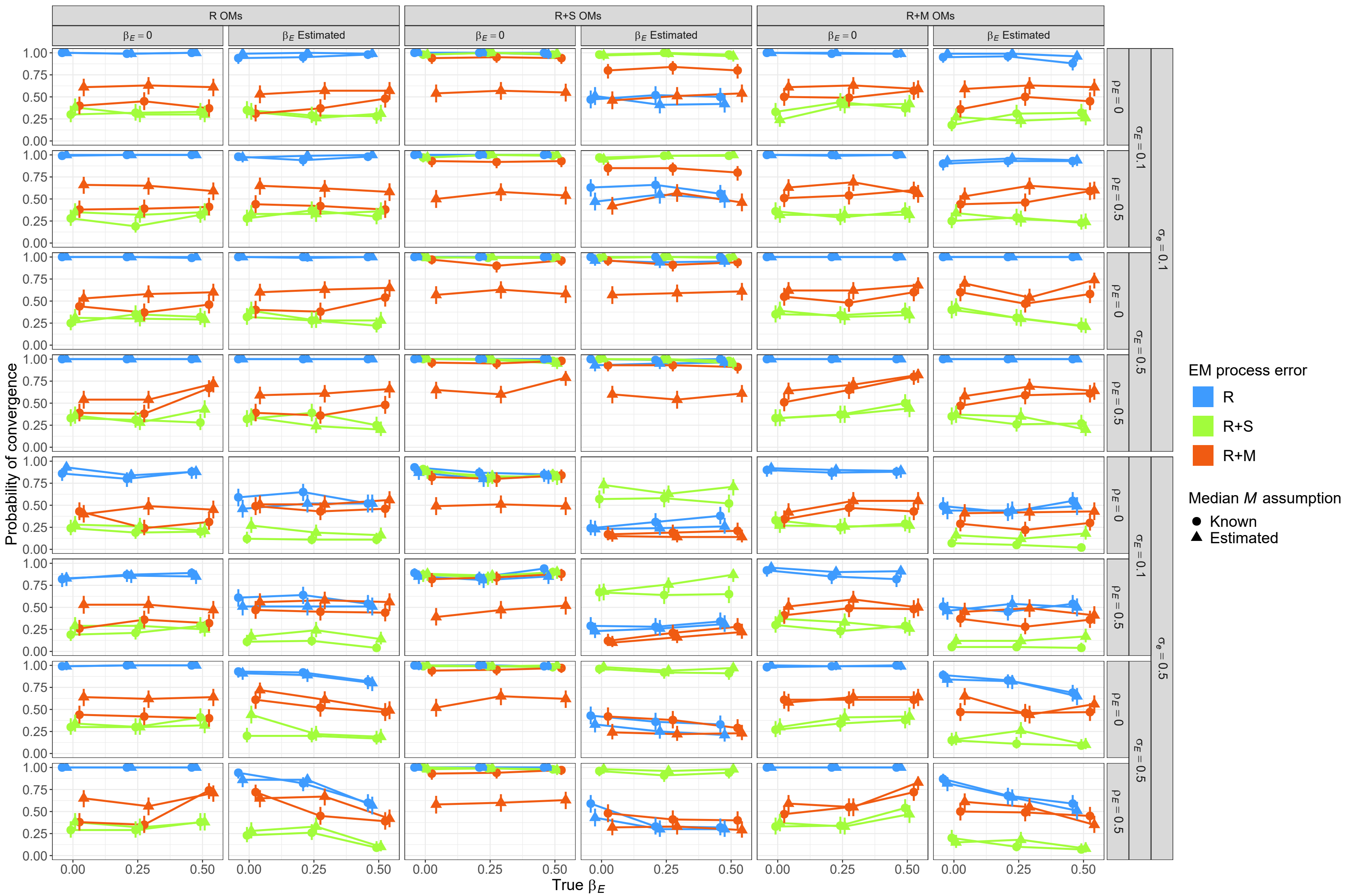

Fig. 1. Estimated probability of fits providing hessian-based standard errors for EMs with alternative process error assumptions, treatment of median natural mortality (e_M^{β} known or estimated), and treatment of covariate effect ($\beta_E = 0$ or estimated). The OMs have R (left) and R+S (middle), 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.

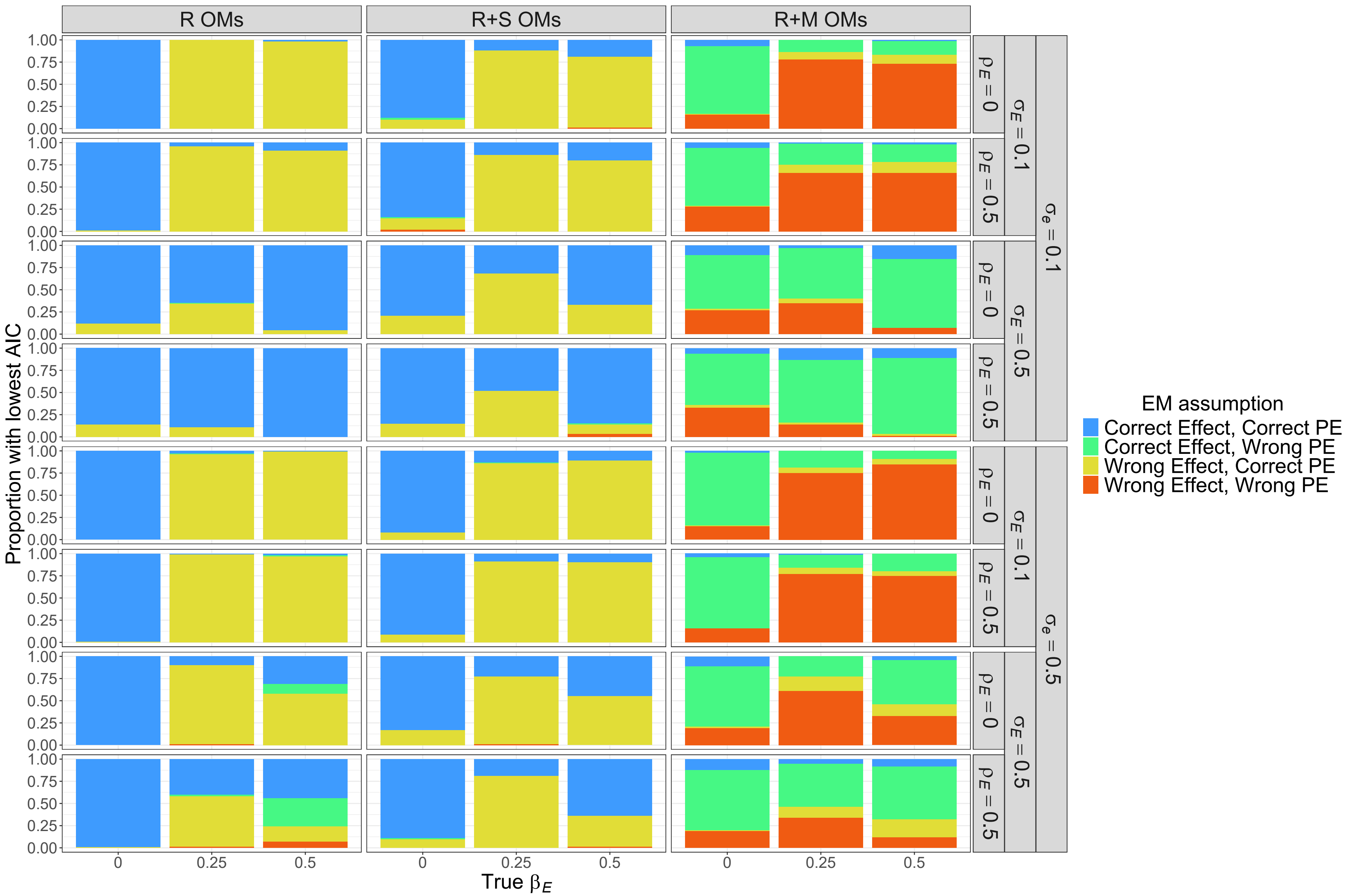

Fig. 2. For each OM, the proportion of simulated data sets where the EM type (treatment of environmental covariate effect and assumed process error type) had the lowest AIC. All OMs had low observation error for fish population observations and temporal contrast in fishing pressure. All EMs estimated median natural mortality rate.

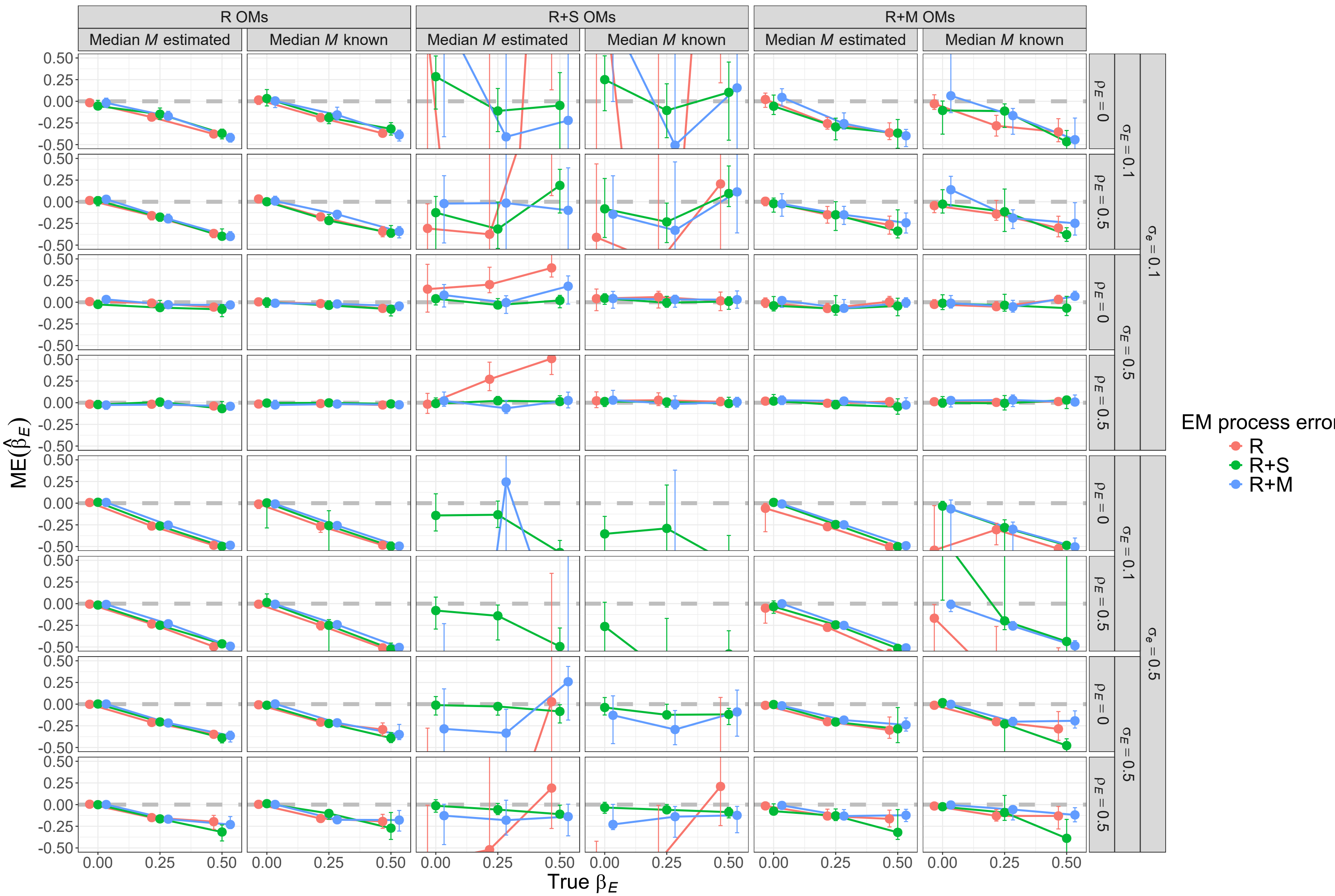

Fig. 3. Median error (ME) of estimates of environmental effect on natural mortality β_E from fitting EMs with alternative process error assumptions and treatment of median natural mortality (e^{β}_{M} known or estimated). All OMs had low observation error and contrast in fishing mortality. Vertical lines represent 95% confidence intervals.

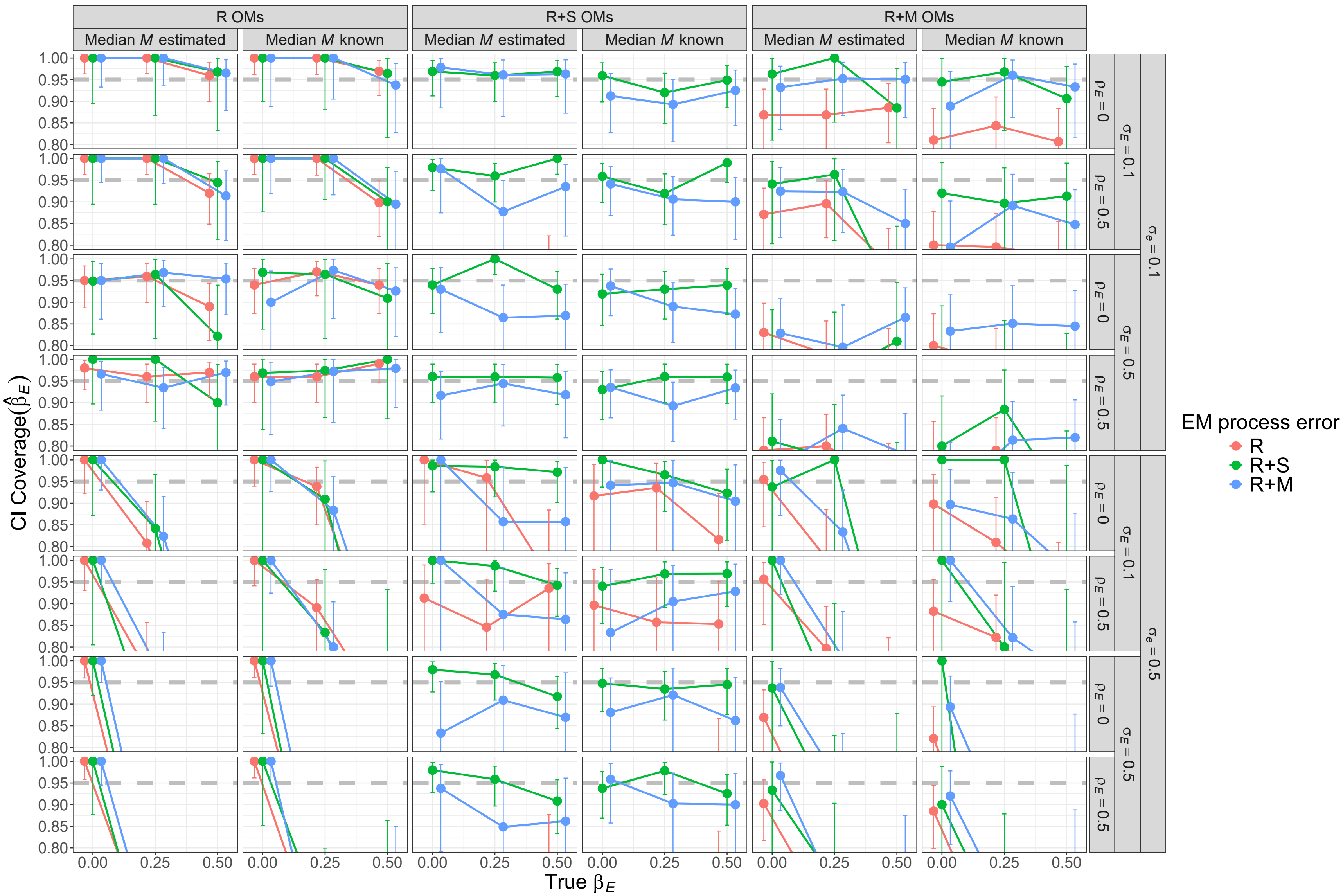

Fig. 4. Probability of 95% confidence interval for β_E containing the true value for EMs with alternative process error assumptions and treatment of median natural mortality (e^{β}_{M} known or estimated). All OMs had low observation error and contrast in fishing mortality. Vertical lines represent 95% confidence intervals.

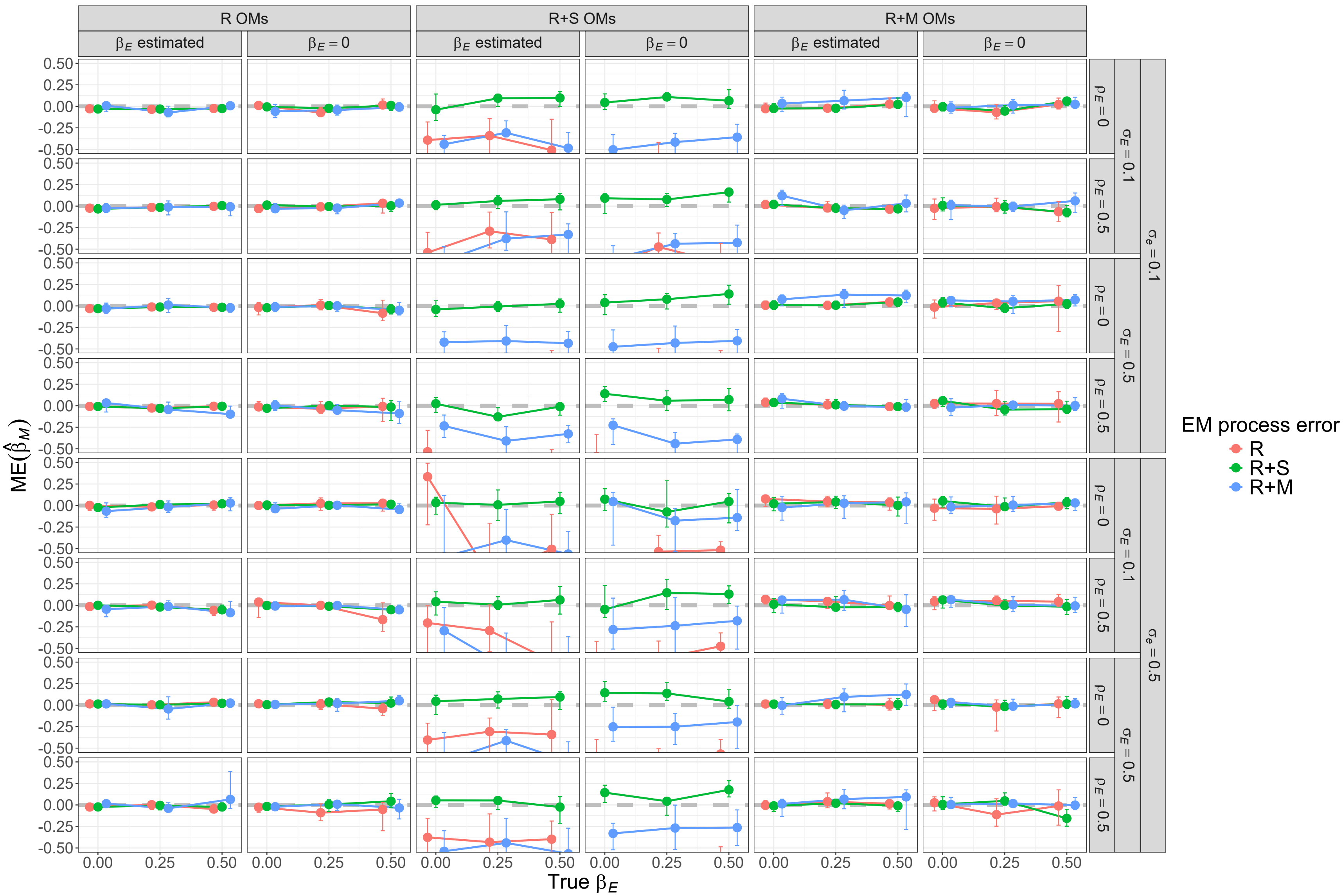

Fig. 5. Median error (ME) of estimates of β_M from fitting EMs with alternative process error assumptions and treatment of covariate effect ($\beta_E = 0$ or estimated). All OMs had low observation error and contrast in fishing mortality. Vertical lines represent 95% confidence intervals.

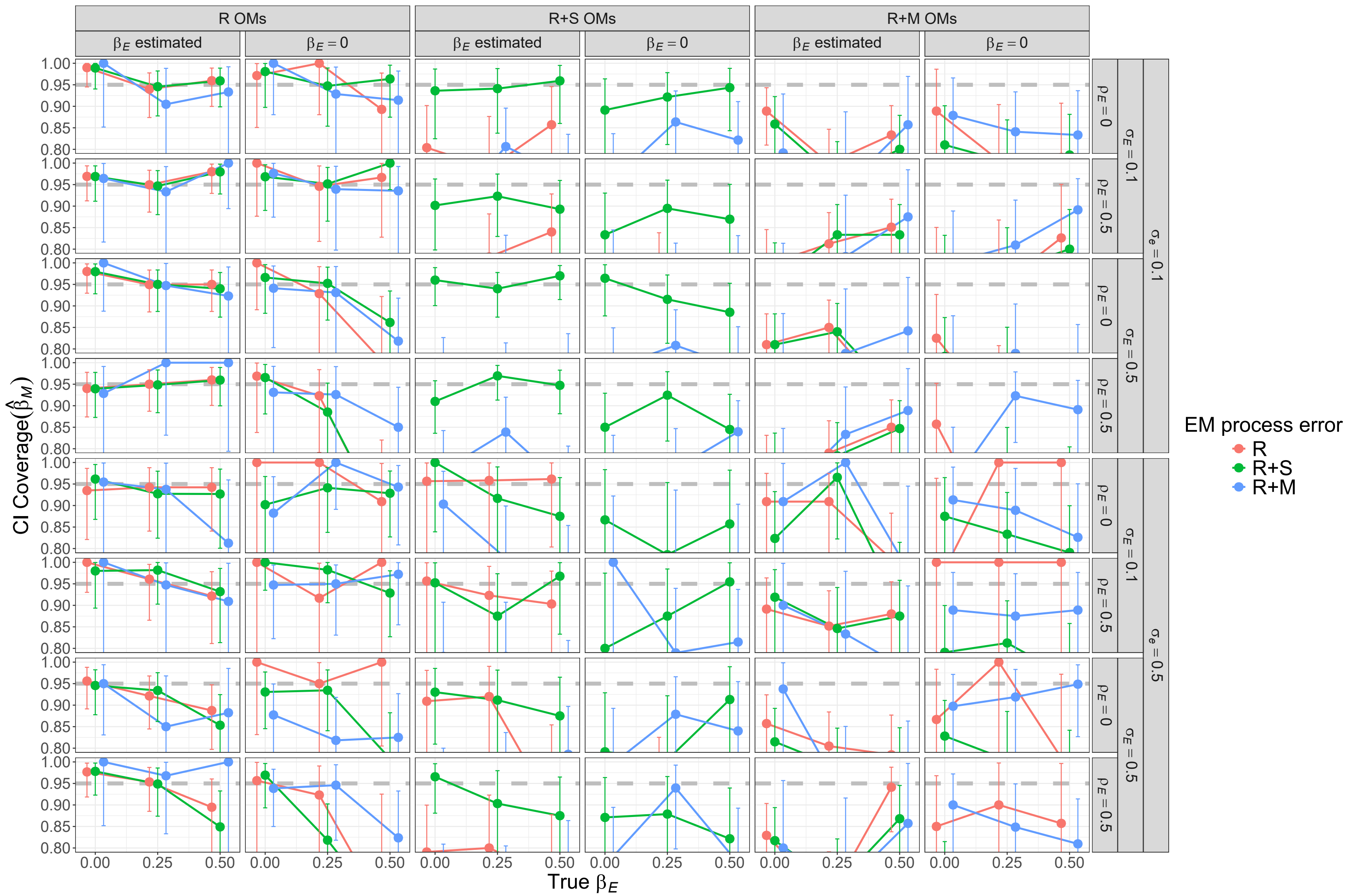

Fig. 6. Probability of 95% confidence interval for β_M containing the true value for EMs with alternative process error assumptions and treatment of covariate effect ($\beta_E = 0$ or estimated). All OMs had low observation error and contrast in fishing mortality. Vertical lines represent 95% confidence intervals.

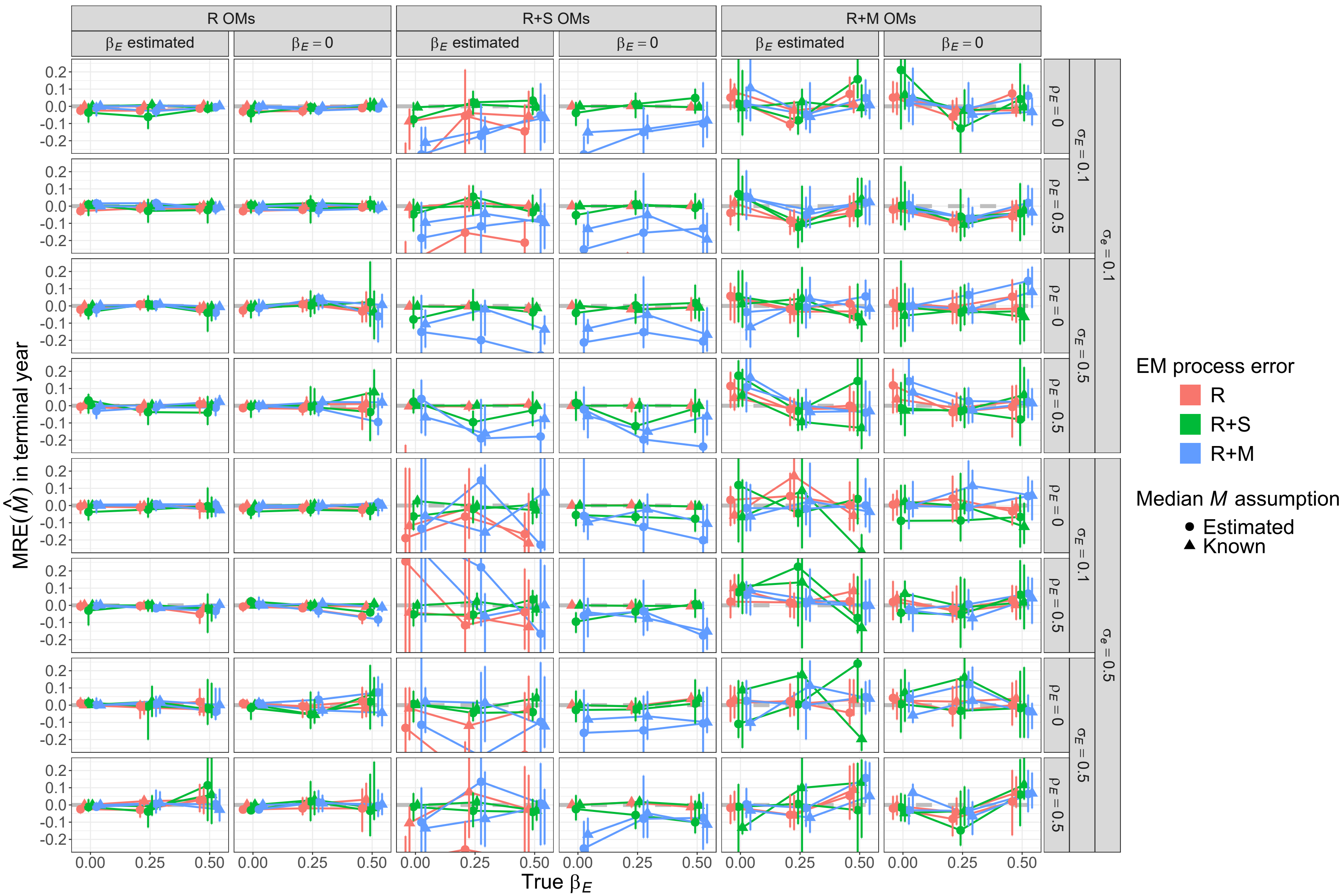

Fig. 7. Median relative error (MRE) of estimates of natural mortality rate in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known). All OMs had low observation error and contrast in fishing mortality.

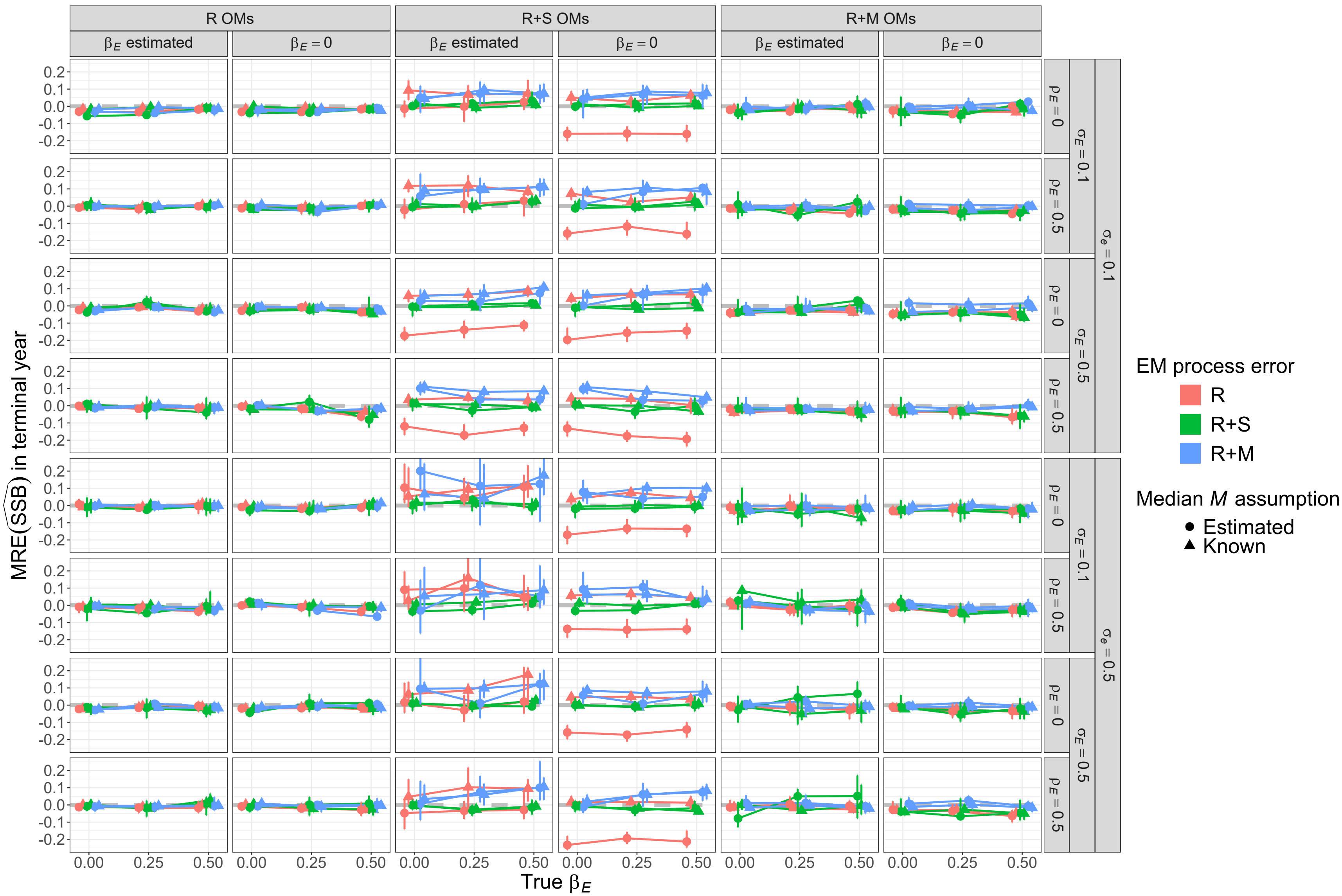

Fig. 8. Median relative error (MRE) of estimates of spawning stock biomass (SSB) in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known). All OMs had low observation error and contrast in fishing mortality.

Supplemental Materials

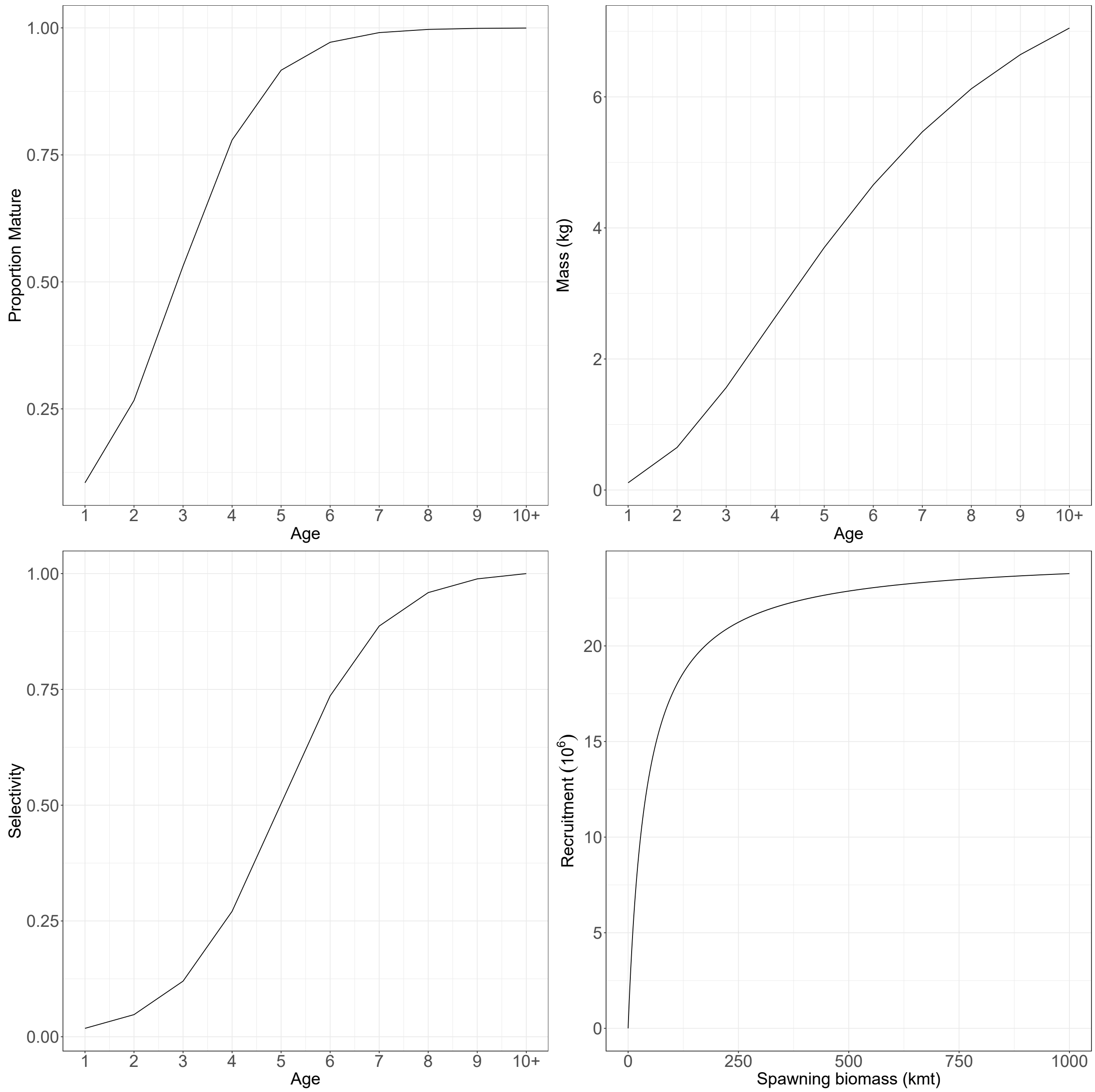

Fig. S1. The proportion mature at age, weight at age, fleet and index selectivity at age, and Beverton-Holt stock-recruit relationship assumed for the population in all operating models. For operating models with random effects on fleet selectivity, this represents the selectivity at the mean of the random effects.

Fig. S2. 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. S3. Example simulations of annual natural mortality rates that may be a function of a temporally varying environmental covariate and autoregressive random effects.

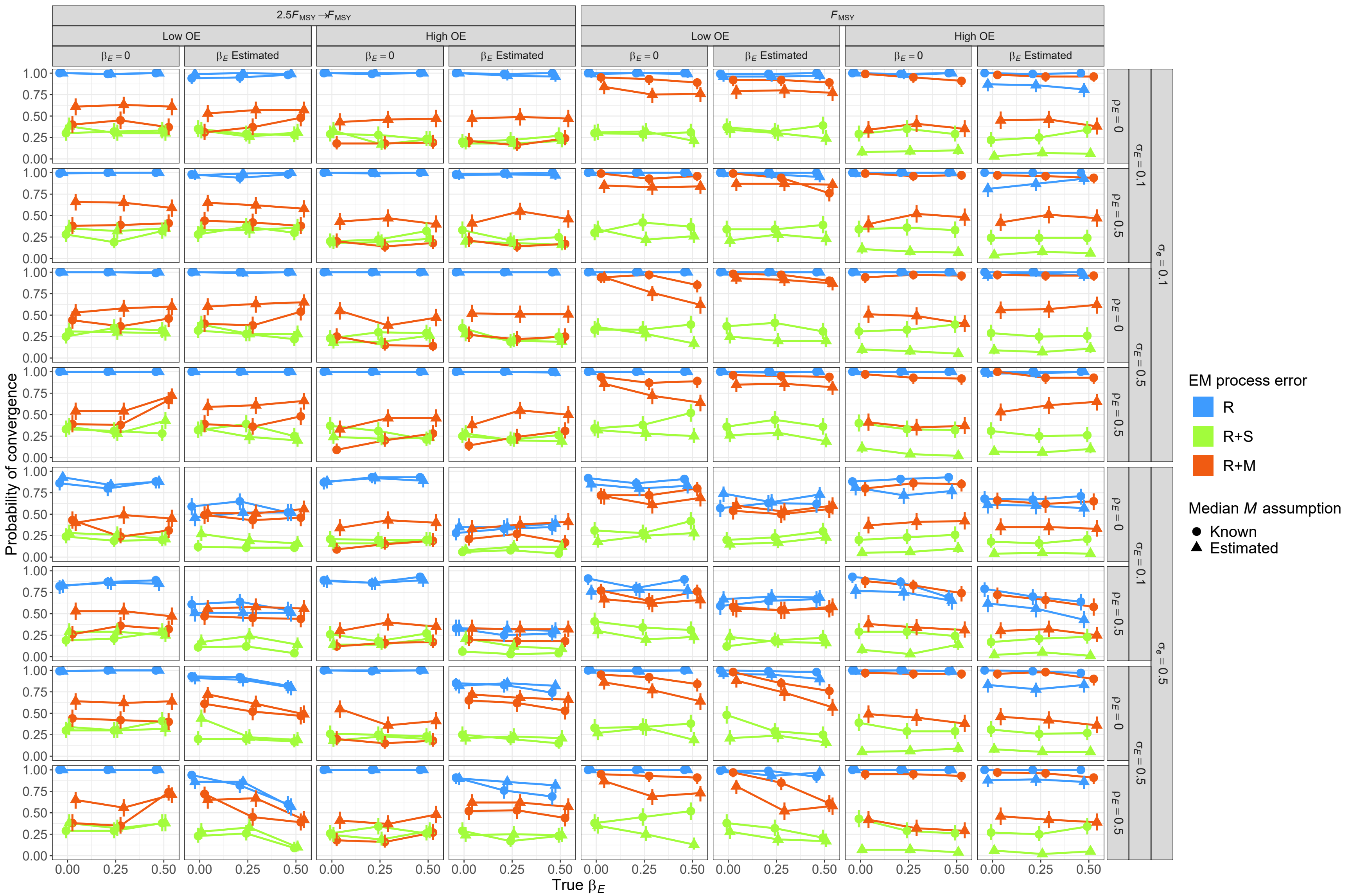

Fig. S4. Estimated probability of fits providing hessian-based standard errors for EMs assuming alternative process error, that estimate or assume known median natural mortality, and that estimate or assume no covariate effect on median natural mortality when fitted to R OMs and three levels of true covariate effect on median natural mortality (x axis). Vertical lines represent 95% confidence intervals.

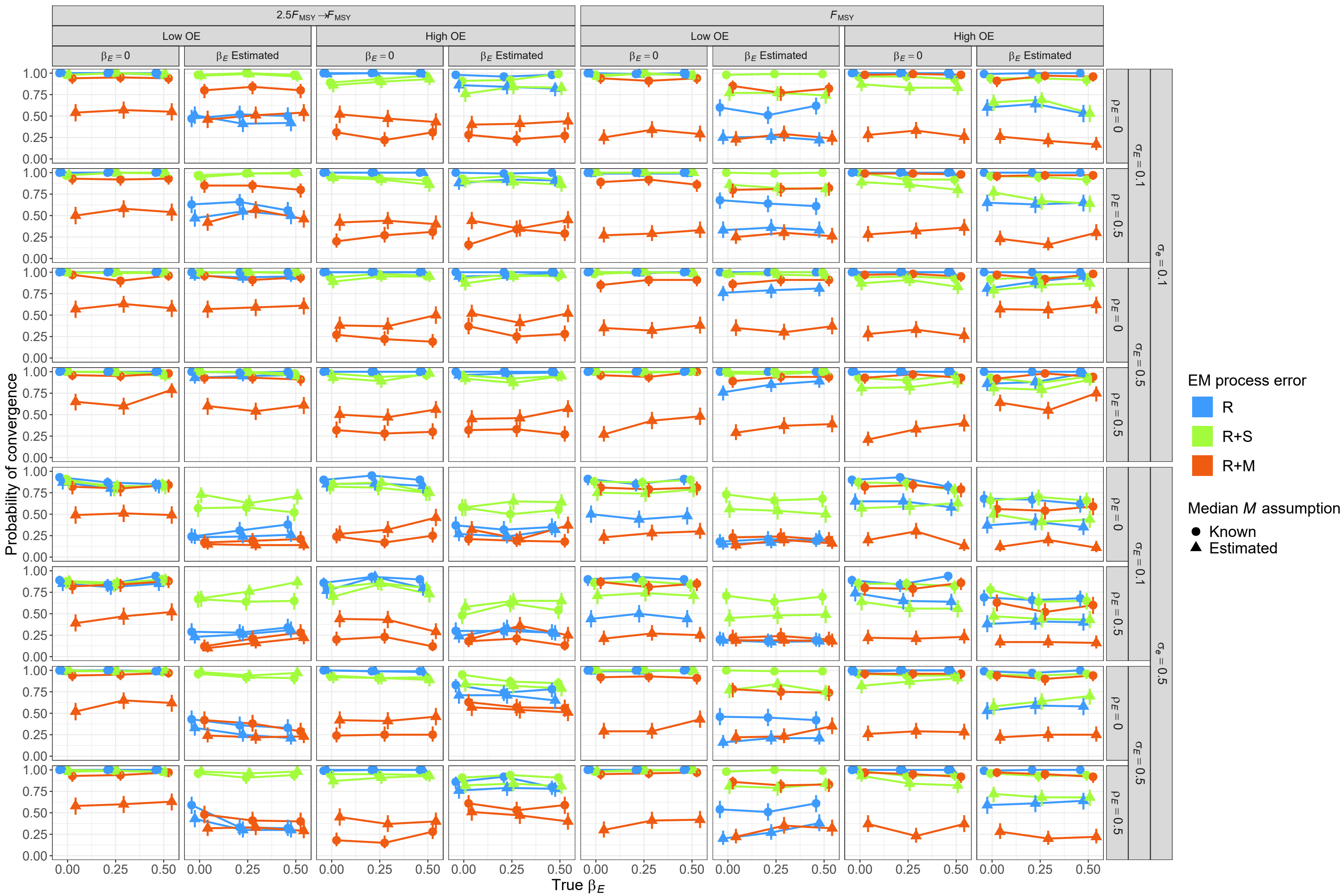

Fig. S5. Estimated probability of fits providing hessian-based standard errors for EMs assuming alternative process error, that estimate or assume known median natural mortality, and that estimate or assume no covariate effect on median natural mortality when fitted to R+S OMs and three levels of true covariate effect on median natural mortality (x axis). Vertical lines represent 95% confidence intervals.

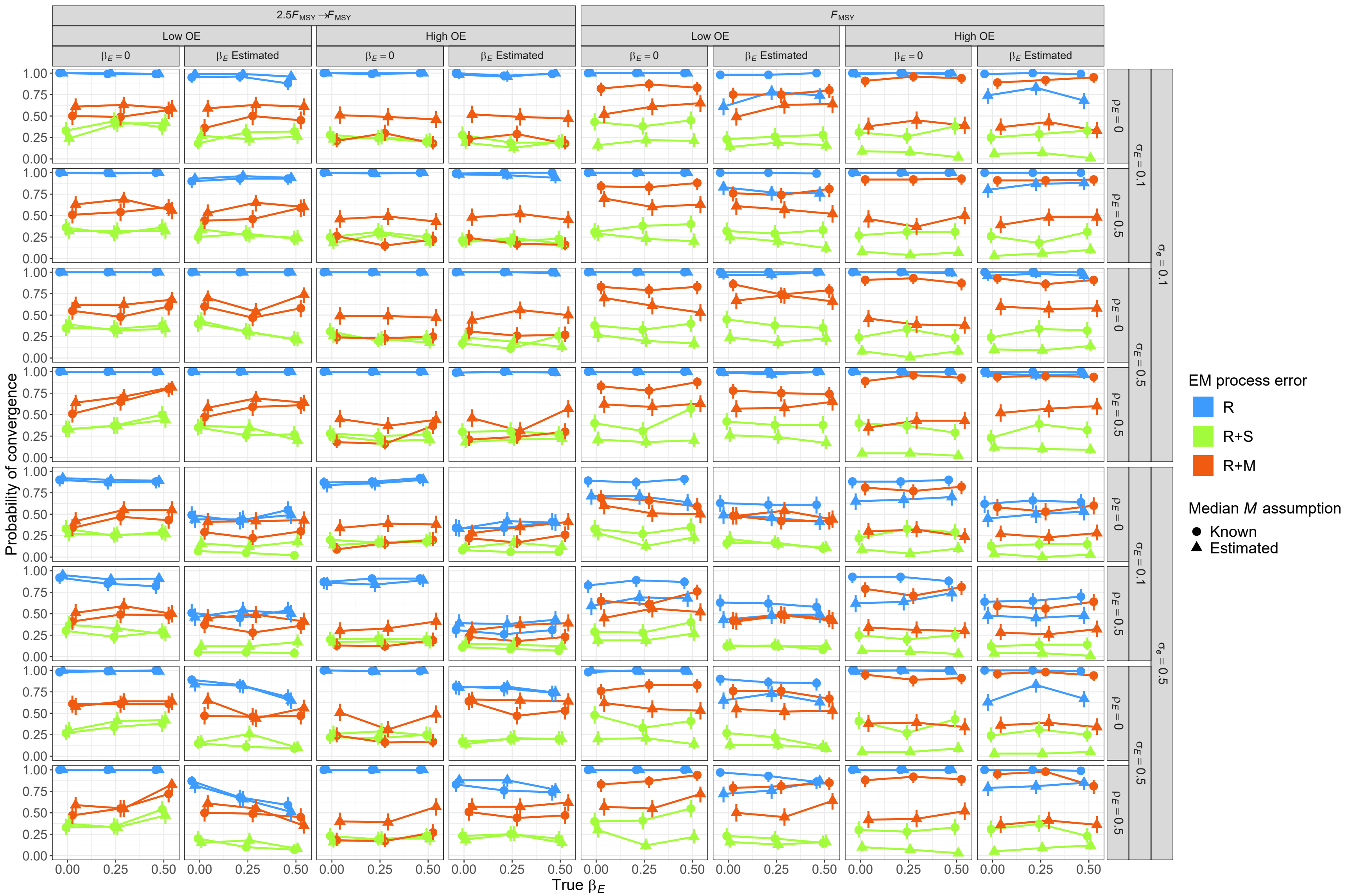

Fig. S6. Estimated probability of fits providing hessian-based standard errors for EMs assuming alternative process error, that estimate or assume known median natural mortality, and that estimate or assume no covariate effect on median natural mortality when fitted to R+M OMs and three levels of true covariate effect on median natural mortality (x axis). Vertical lines represent 95% confidence intervals.

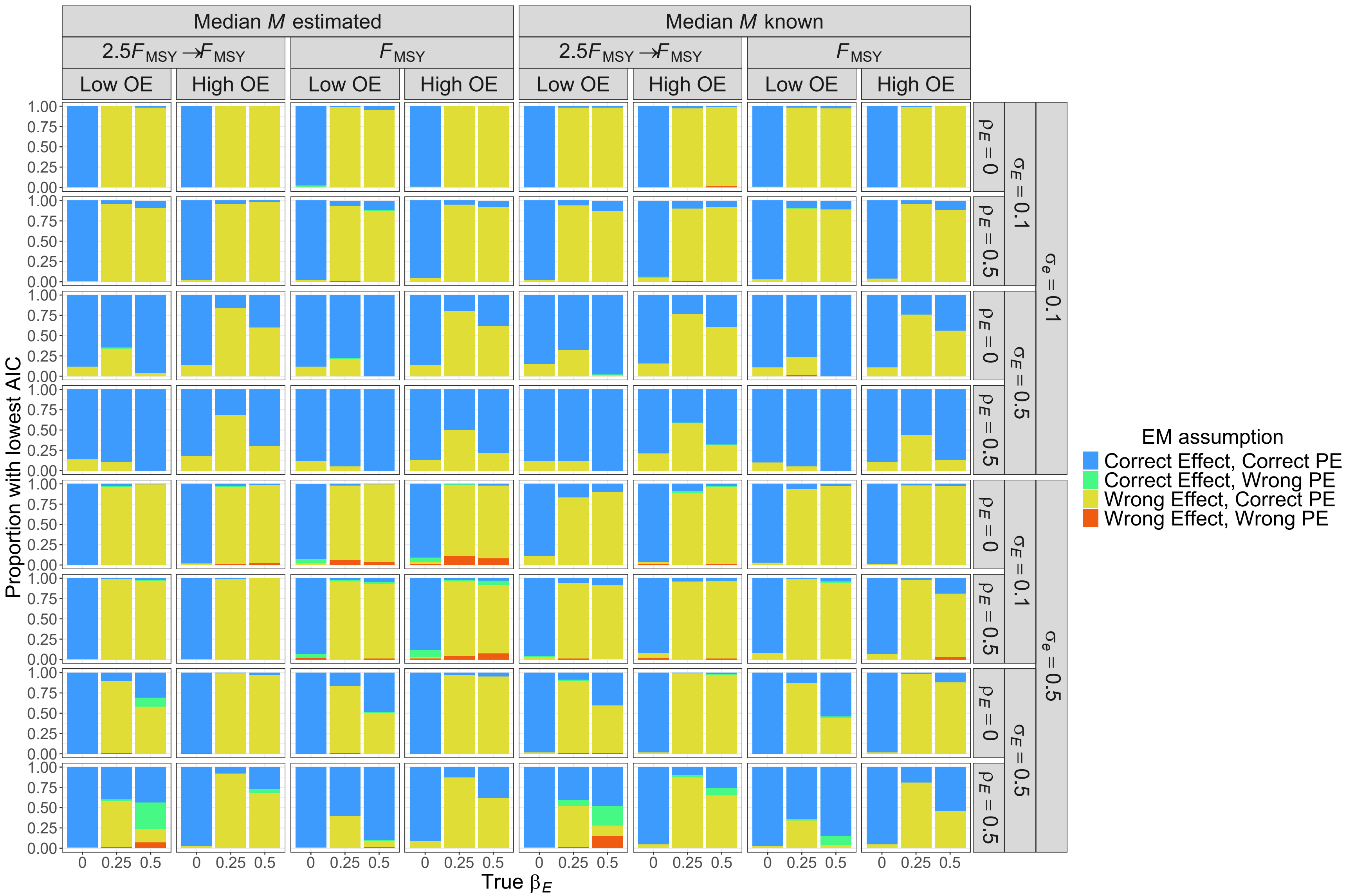

Fig. S7. Proportion of simulated data sets for R OMs where the EM type (treatment of environmental covariate and assumed process error type) had the lowest AIC.

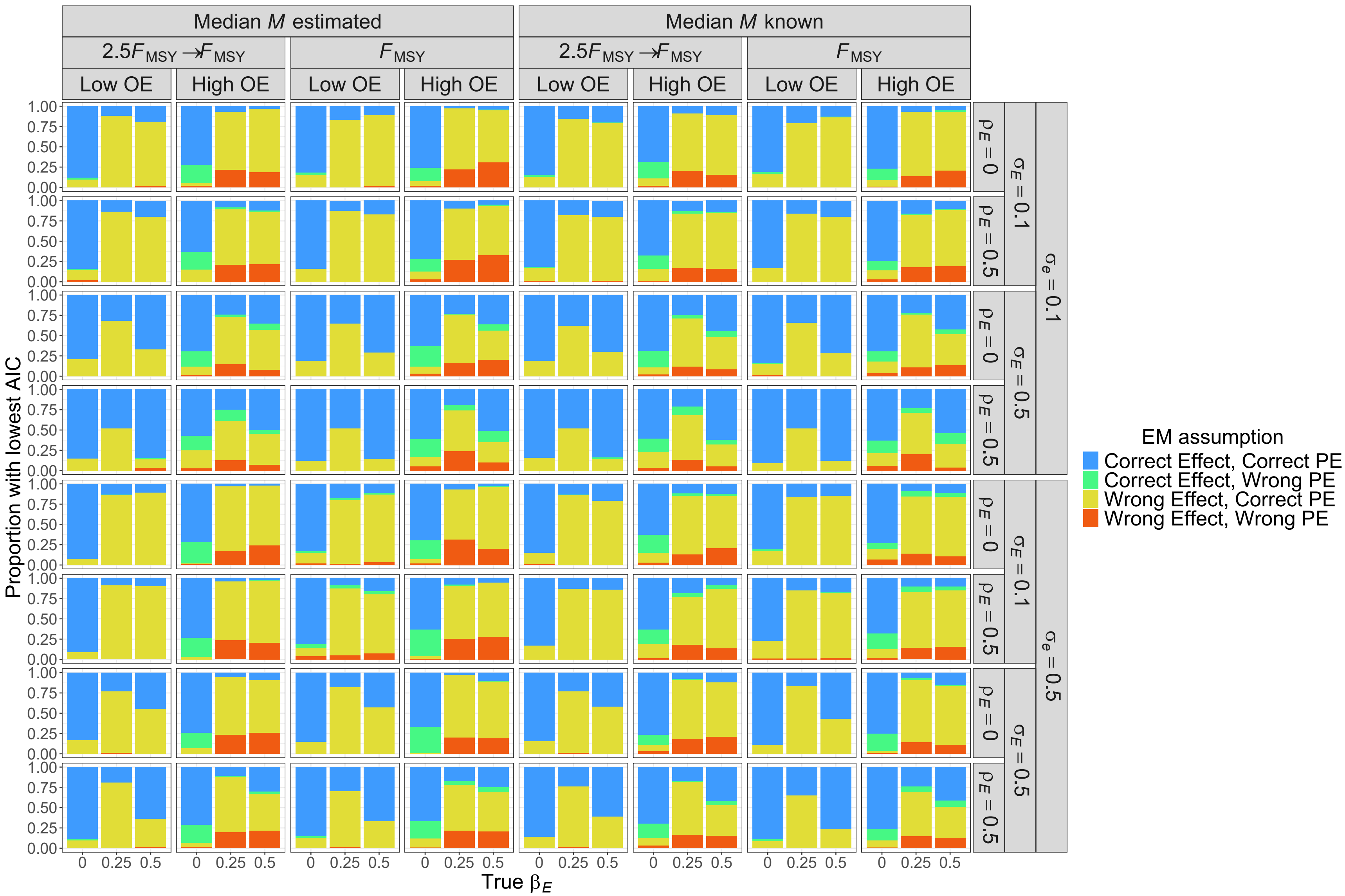

Fig. S8. Proportion of simulated data sets for R+S OMs where the EM type (treatment of environmental covariate and assumed process error type) had the lowest AIC.

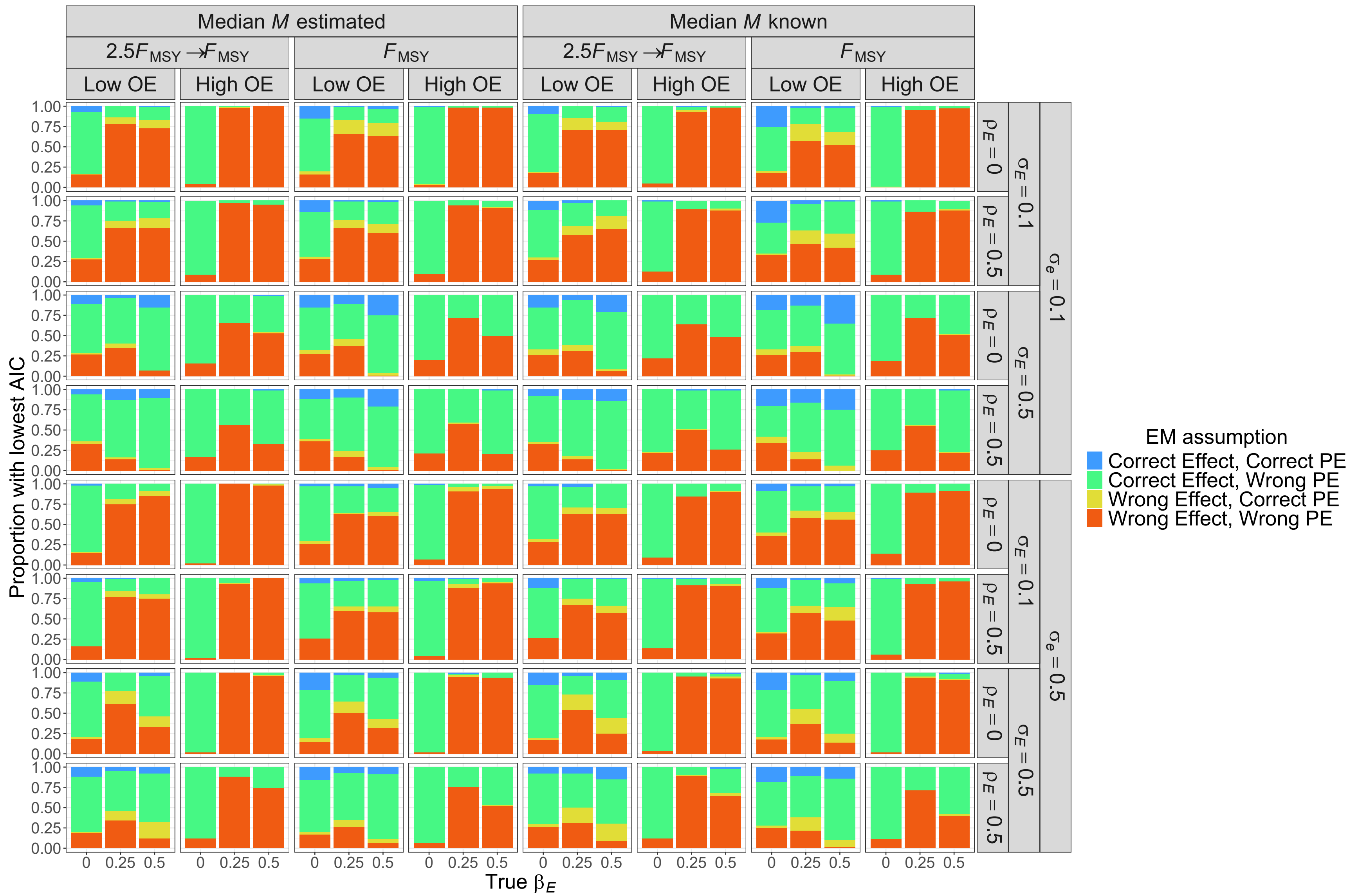

Fig. S9. Proportion of simulated data sets for R+M OMs where the EM type (treatment of environmental covariate and assumed process error type) had the lowest AIC.

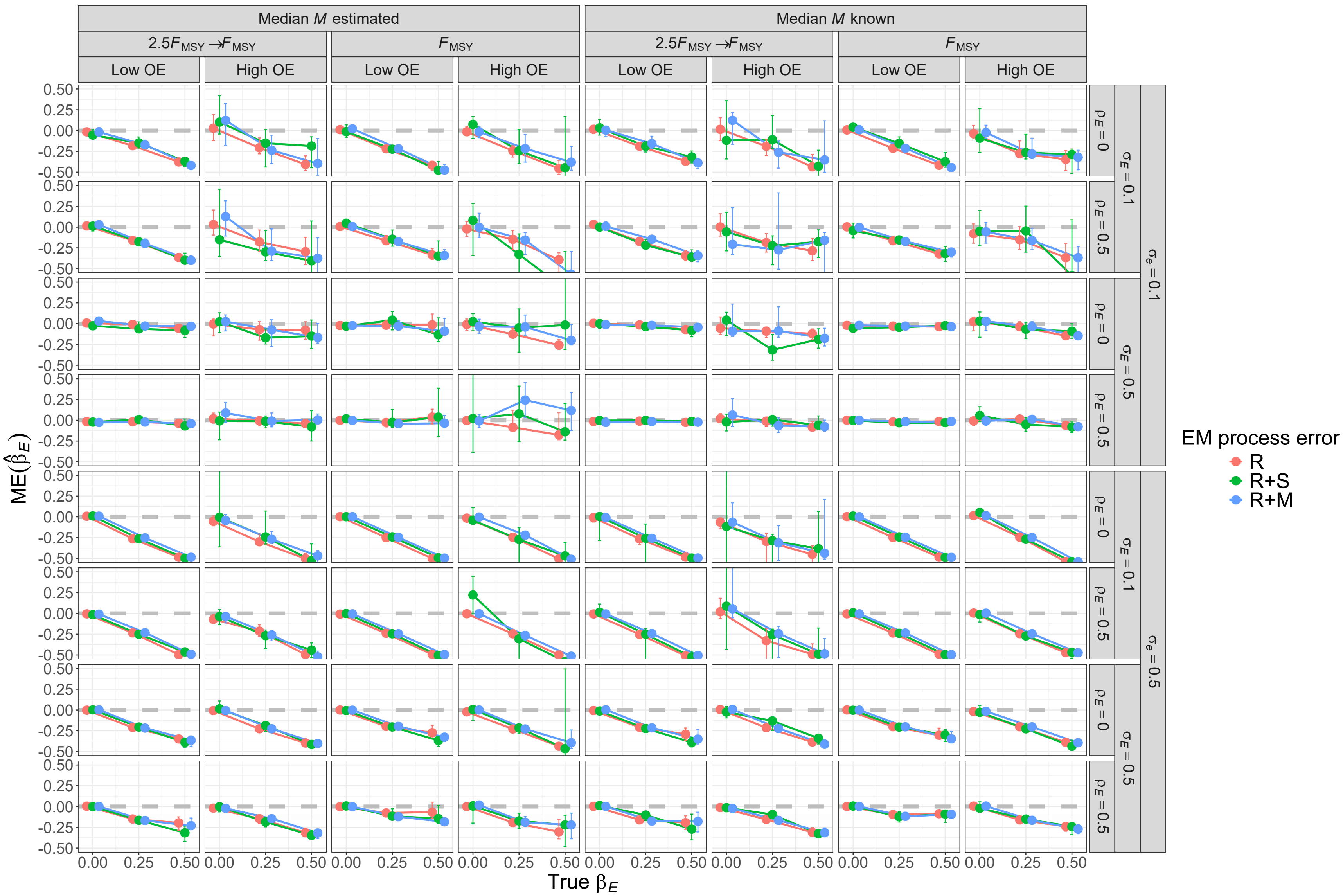

Fig. S10. For R OMs, median error (ME) of estimates of environmental effect on natural mortality β_E from fitting EMs with alternative process error assumptions and treatment of median natural mortality (e^{β}_{M} known or estimated). Vertical lines represent 95% confidence intervals.

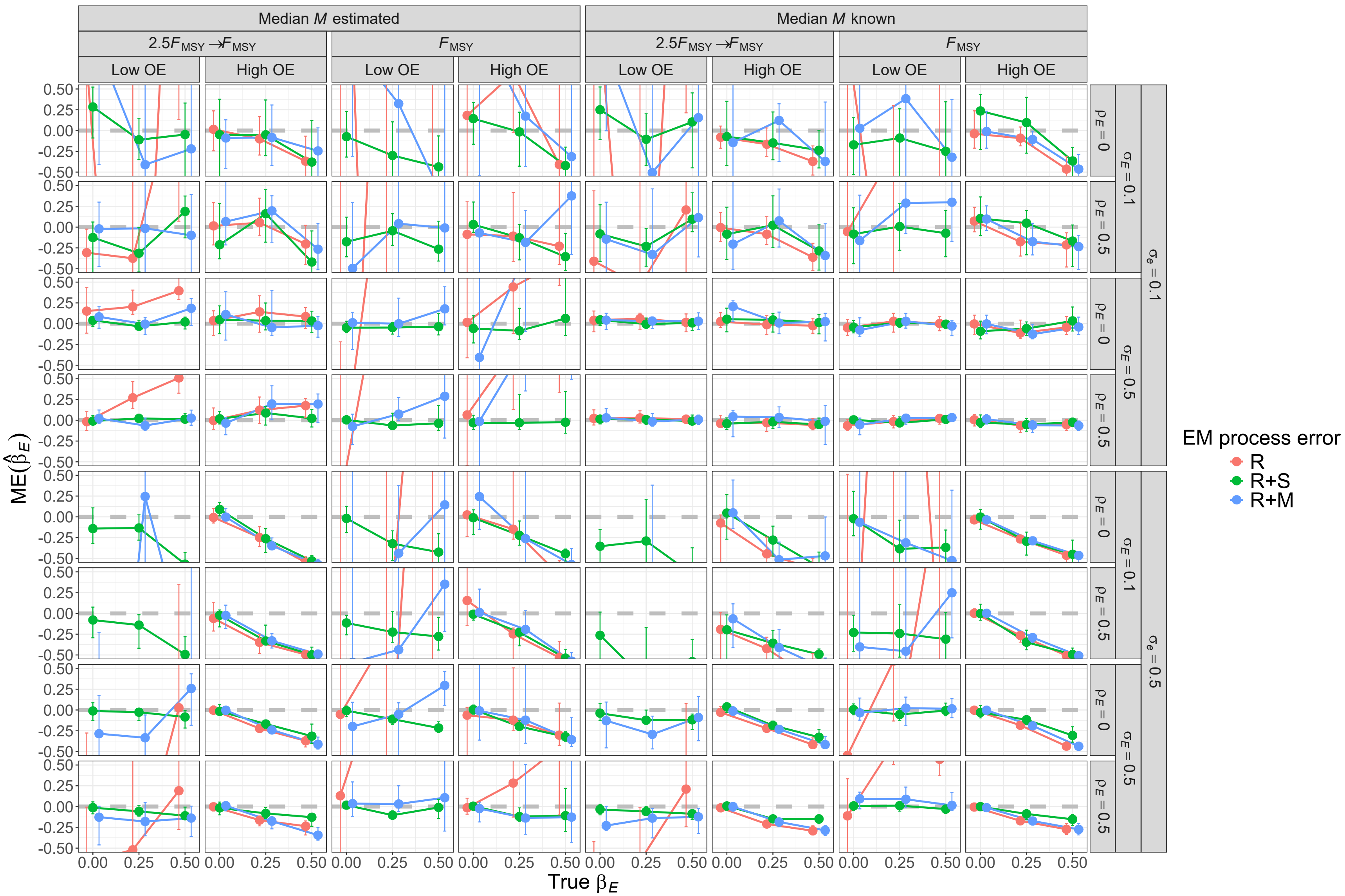

Fig. S11. For R+S OMs, median error (ME) of estimates of environmental effect on natural mortality β_E from fitting EMs with alternative process error assumptions and treatment of median natural mortality (e^{β}_{M} known or estimated). Vertical lines represent 95% confidence intervals.

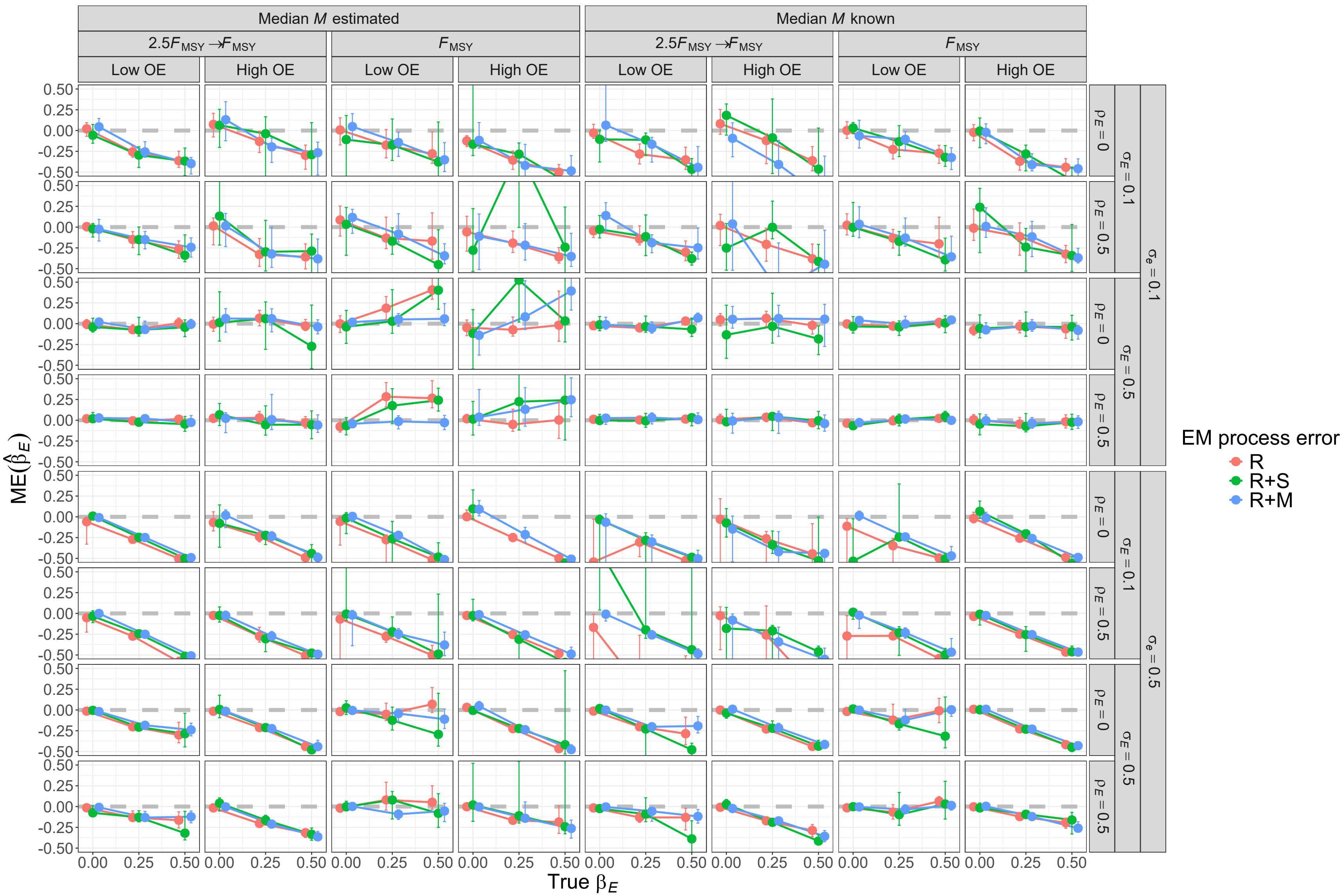

Fig. S12. For R+M OMs, median error (ME) of estimates of environmental effect on natural mortality β_E from fitting EMs with alternative process error assumptions and treatment of median natural mortality (e^{β}_{M} known or estimated). Vertical lines represent 95% confidence intervals.

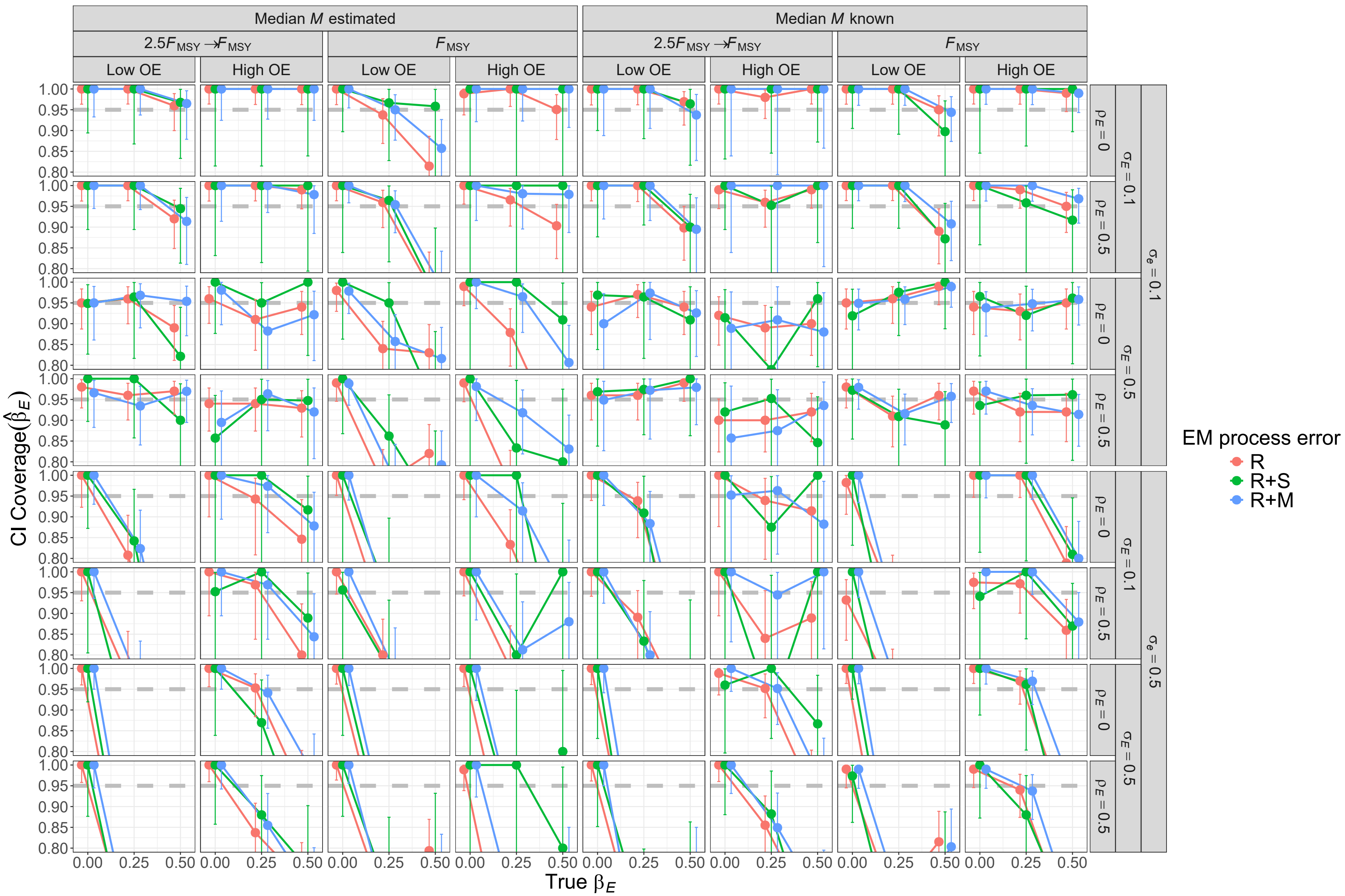

Fig. S13. For R OMs, probability of 95% confidence interval for β_E containing the true value for EMs with alternative process error assumptions and treatment of median natural mortality (e_M^{β} known or estimated). Vertical lines represent 95% confidence intervals.

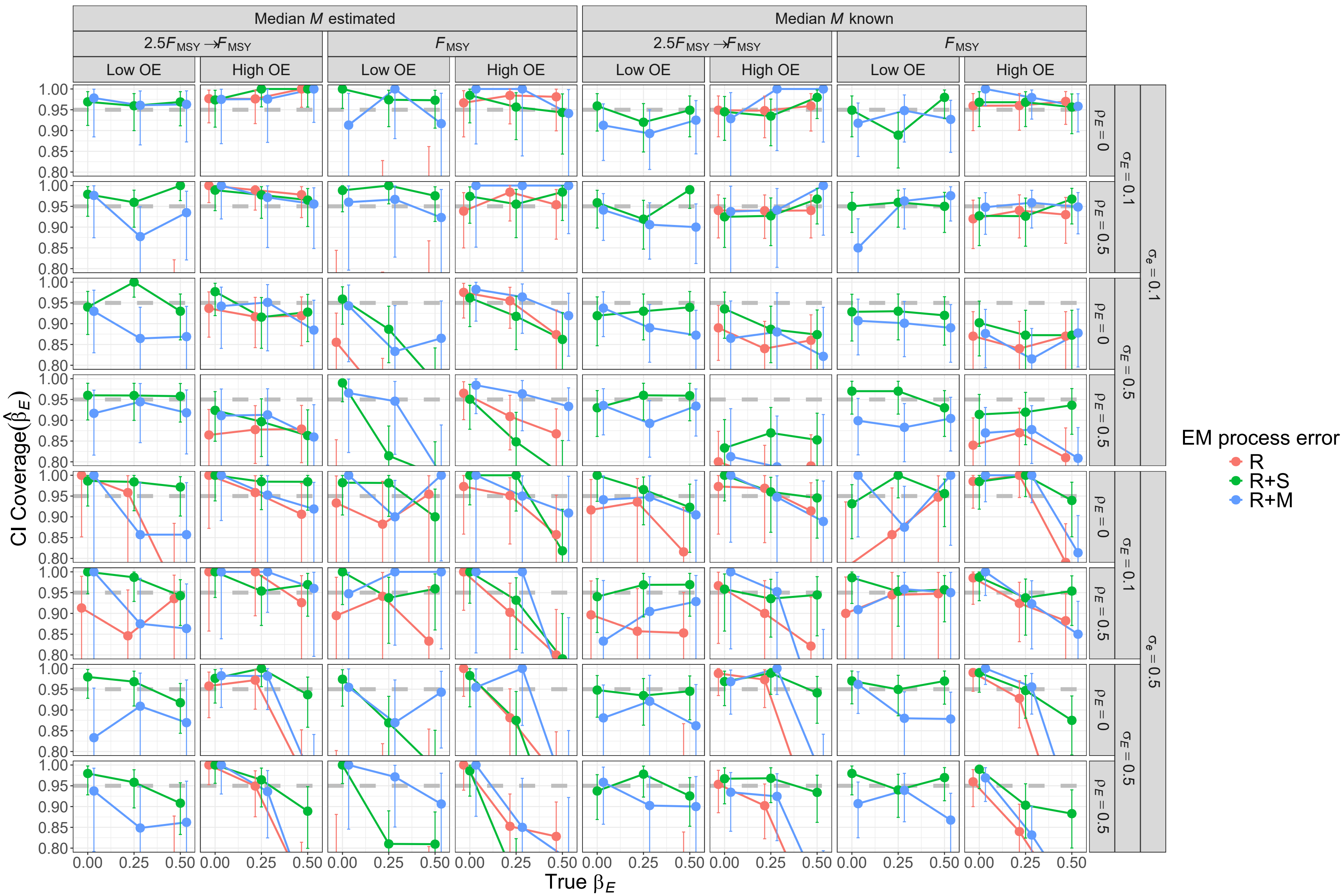

Fig. S14. For R+S OMs, probability of 95% confidence interval for β_E containing the true value for EMs with alternative process error assumptions and treatment of median natural mortality (e^{β}_{M} known or estimated). Vertical lines represent 95% confidence intervals.

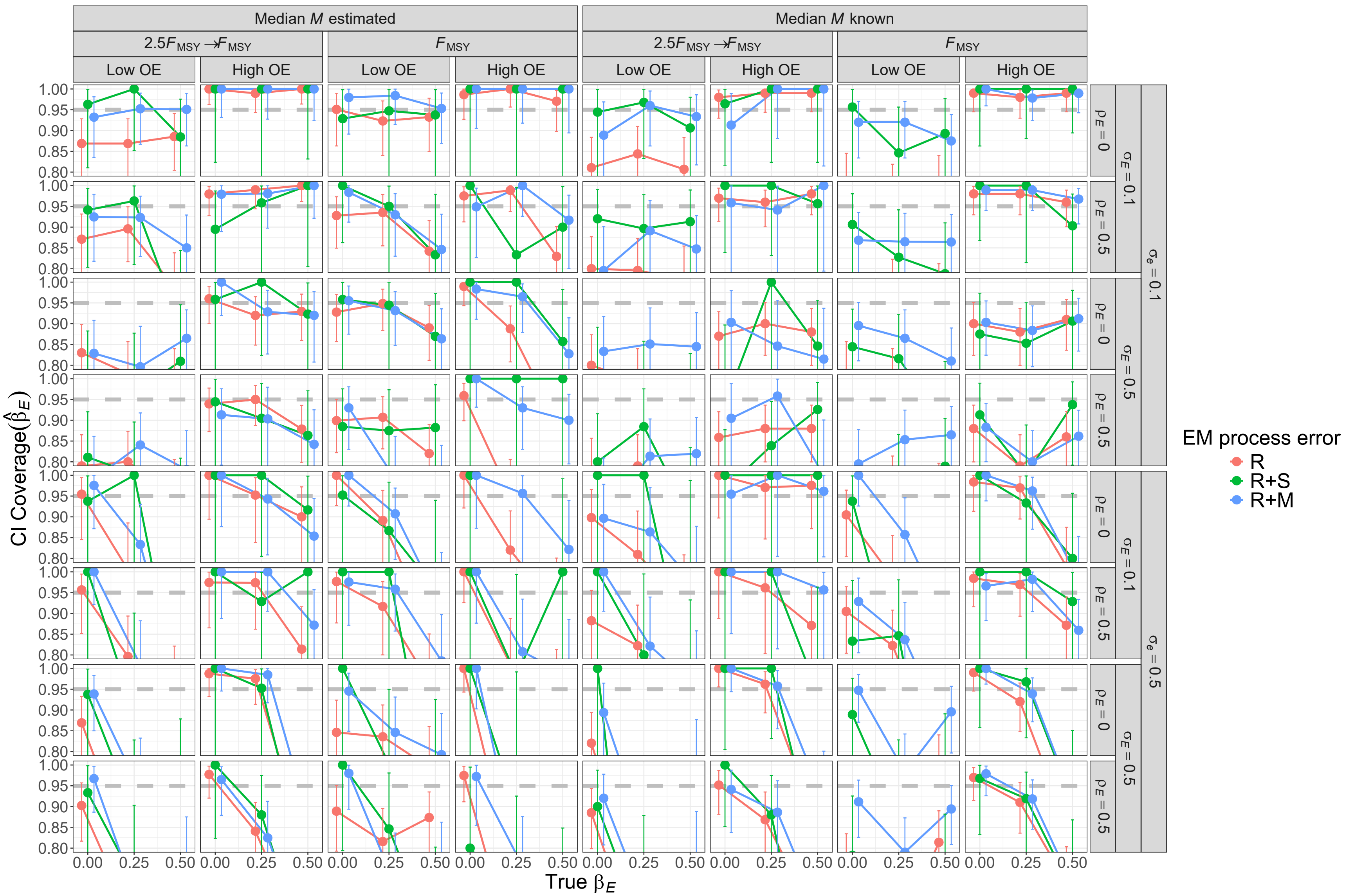

Fig. S15. For R+M OMs, probability of 95% confidence interval for β_E containing the true value for EMs with alternative process error assumptions and treatment of median natural mortality (e^{β}_{M} known or estimated). Vertical lines represent 95% confidence intervals.

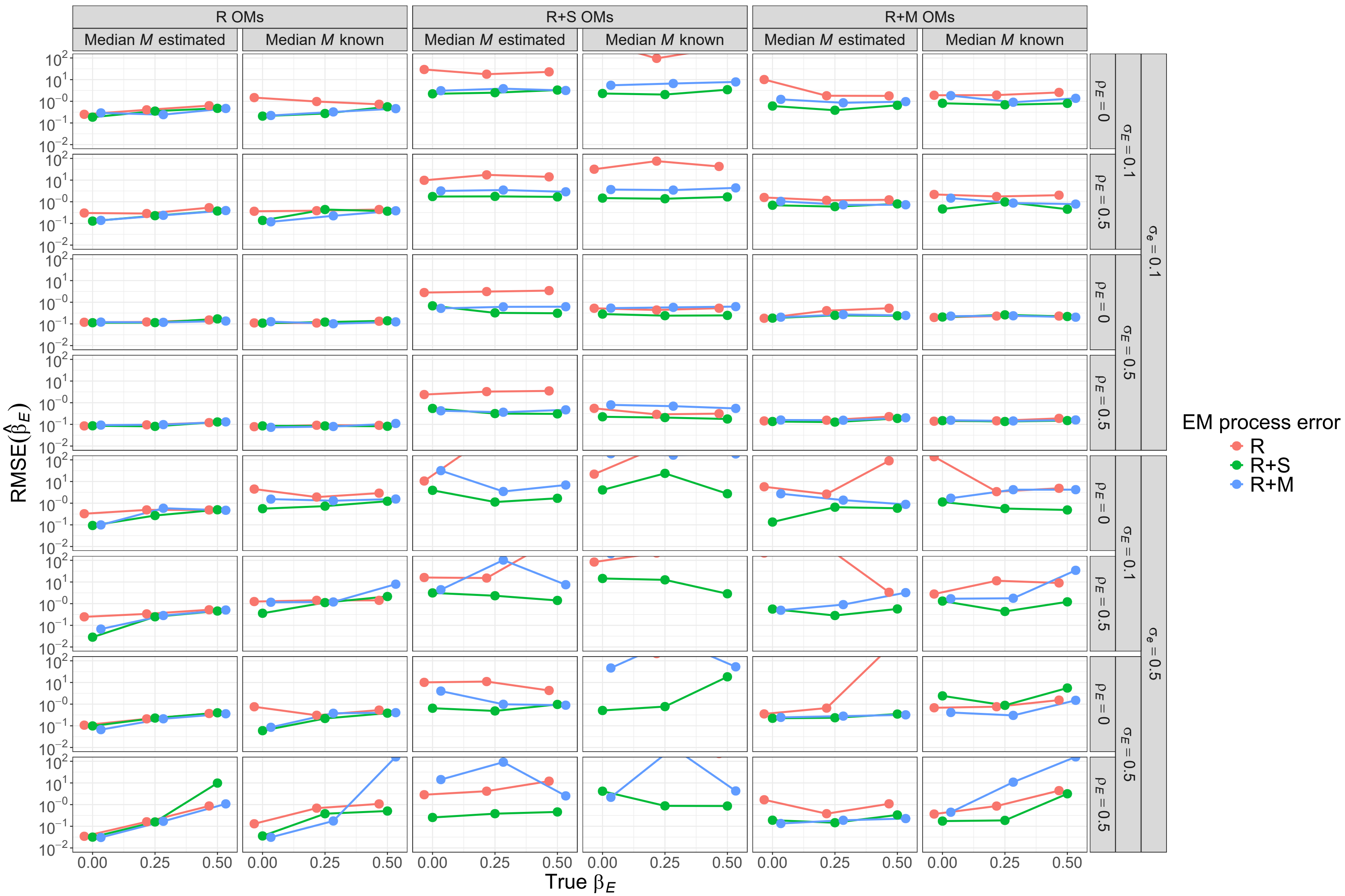

Fig. S16. Root mean square error (RMSE) of estimates of covariate effect on natural mortality β_E from fitting EMs with alternative process error assumptions and treatment of median natural mortality (e^{β}_{M} known or estimated). All OMs had low observation error and contrast in fishing mortality.

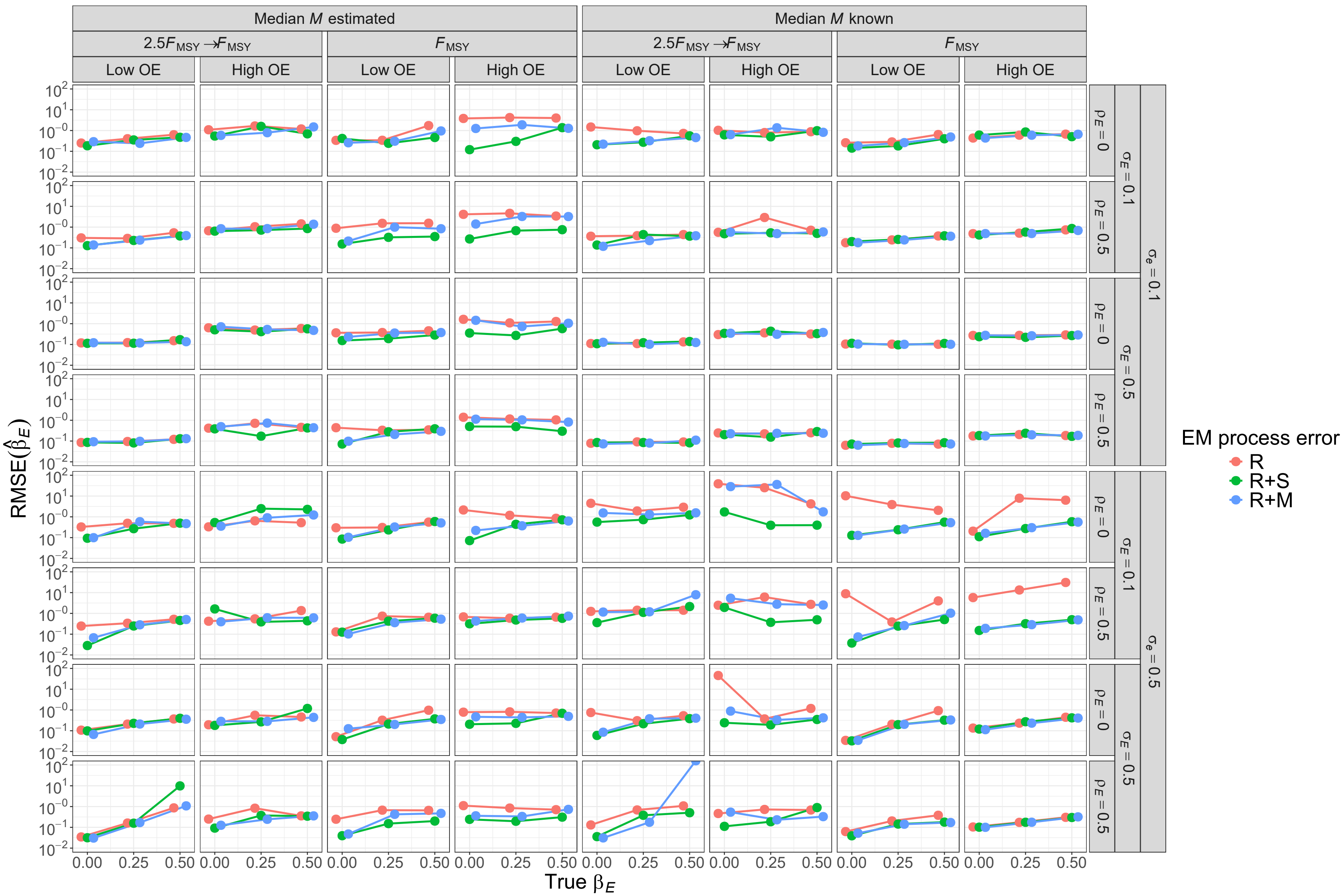

Fig. S17. For R OMs, root mean square error (RMSE) of estimates of covariate effect on natural mortality β_E from fitting EMs with alternative process error assumptions and treatment of median natural mortality (e^{β}_{M} known or estimated).

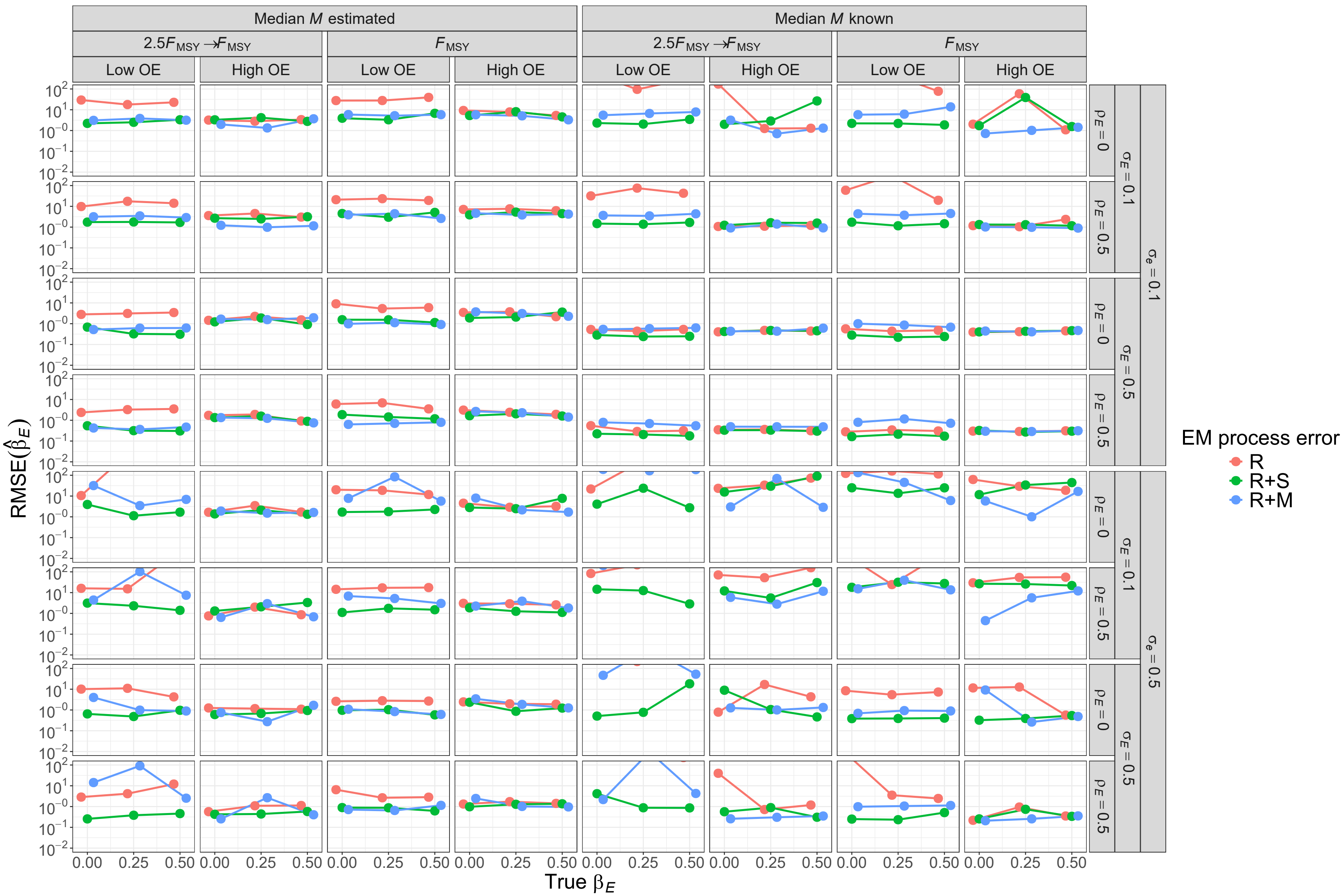

Fig. S18. For R+S OMs, root mean square error (RMSE) of estimates of covariate effect on natural mortality β_E from fitting EMs with alternative process error assumptions and treatment of median natural mortality (e^{β}_M known or estimated).

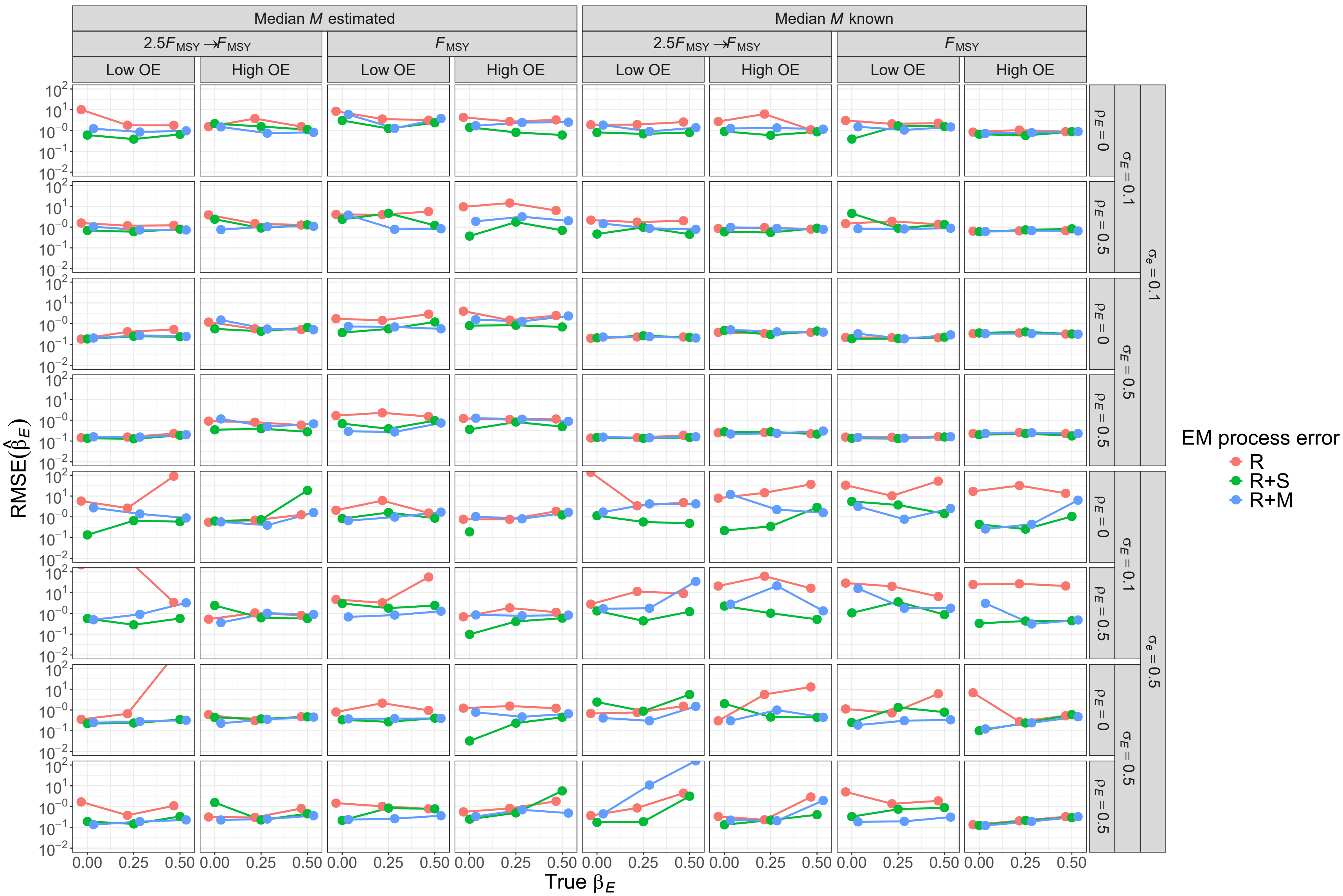

Fig. S19. For R+M OMs, root mean square error (RMSE) of estimates of covariate effect on natural mortality β_E from fitting EMs with alternative process error assumptions and treatment of median natural mortality (e_M^{β} known or estimated).

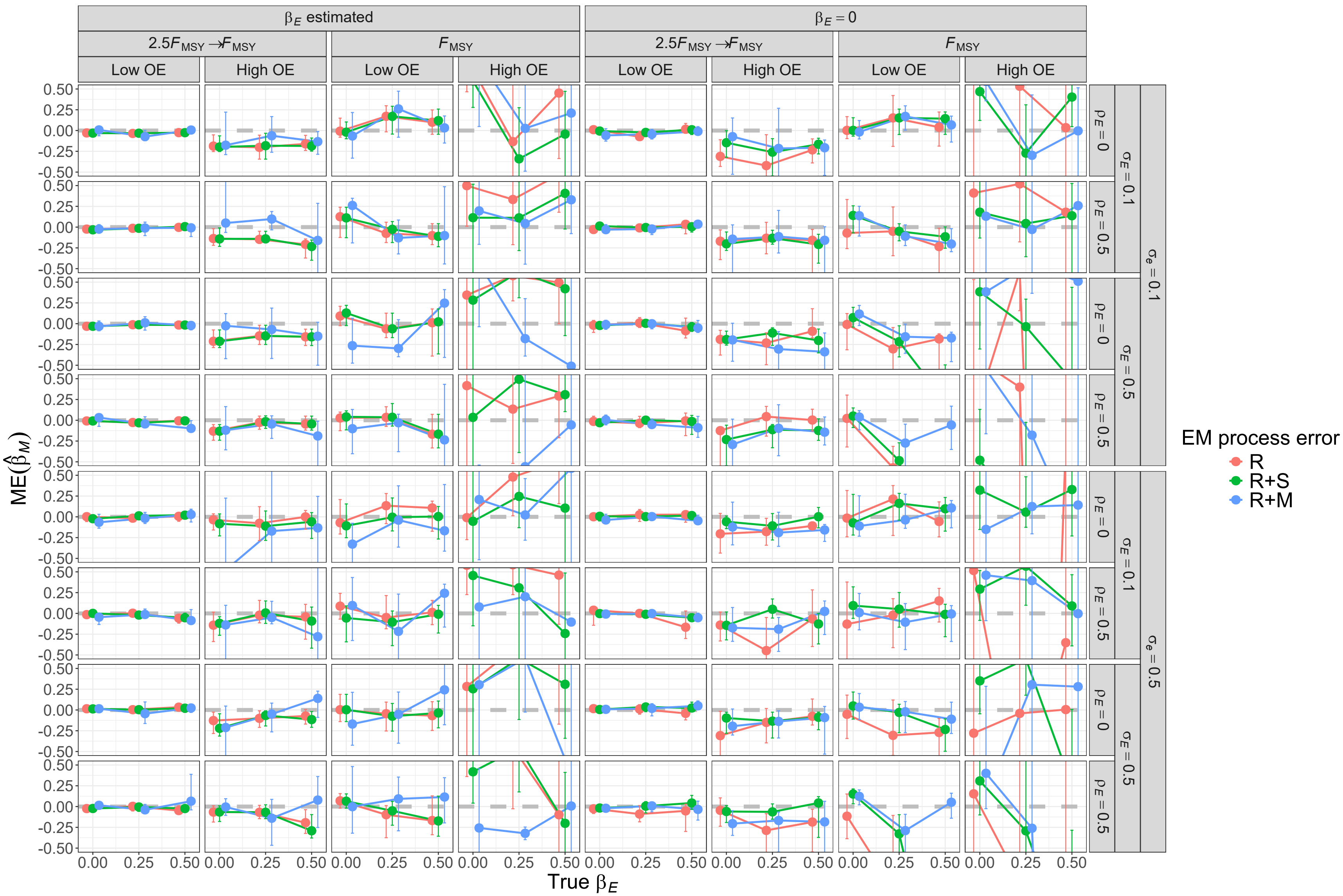

Fig. S20. For R OMs, median error (ME) of estimates of β_M from fitting EMs with alternative process error assumptions and treatment of covariate effect ($\beta_E = 0$ or estimated). Vertical lines represent 95% confidence intervals.

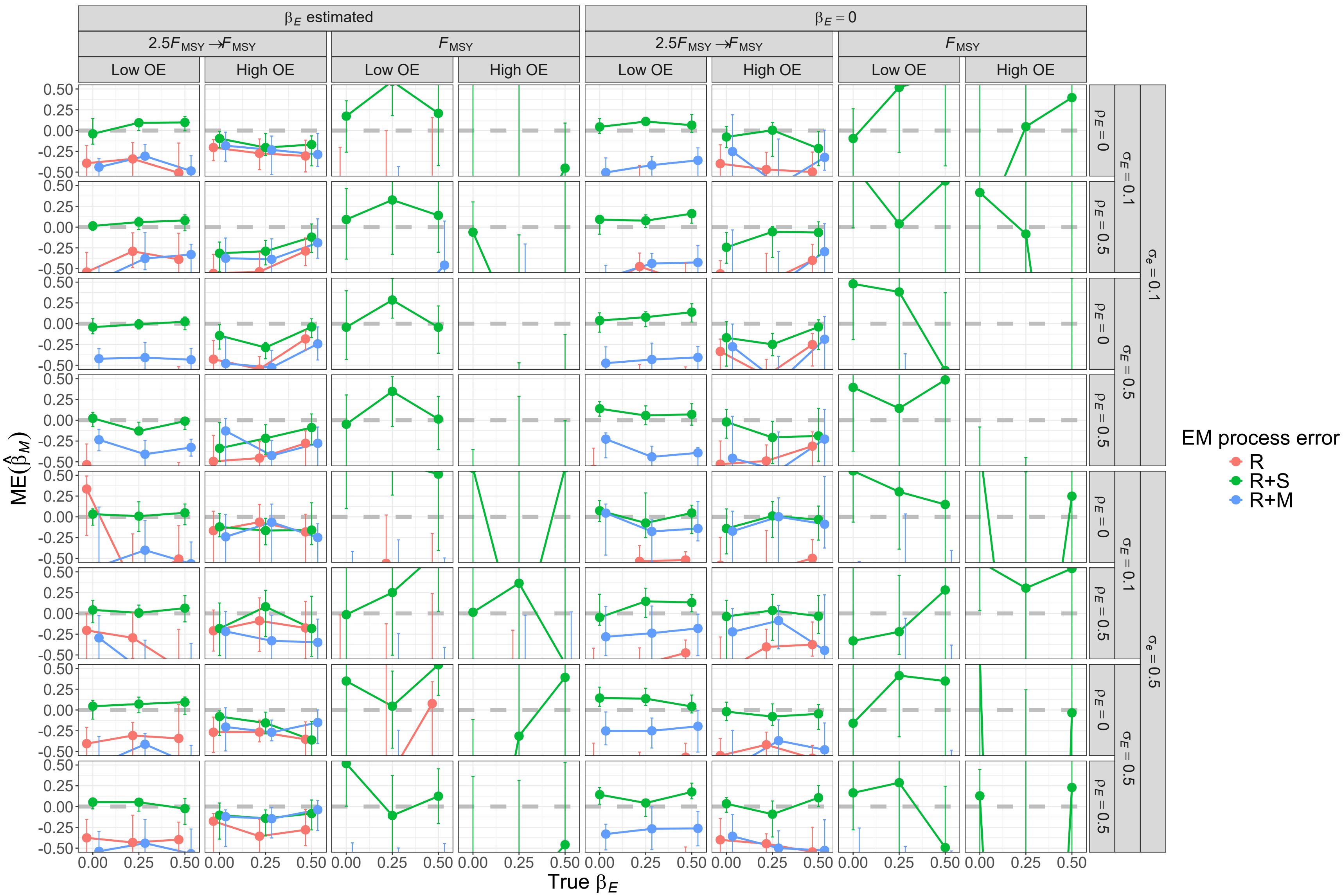

Fig. S21. For R+S OMs, median error (ME) of estimates of β_M from fitting EMs with alternative process error assumptions and treatment of covariate effect ($\beta_E = 0$ or estimated). Vertical lines represent 95% confidence intervals.

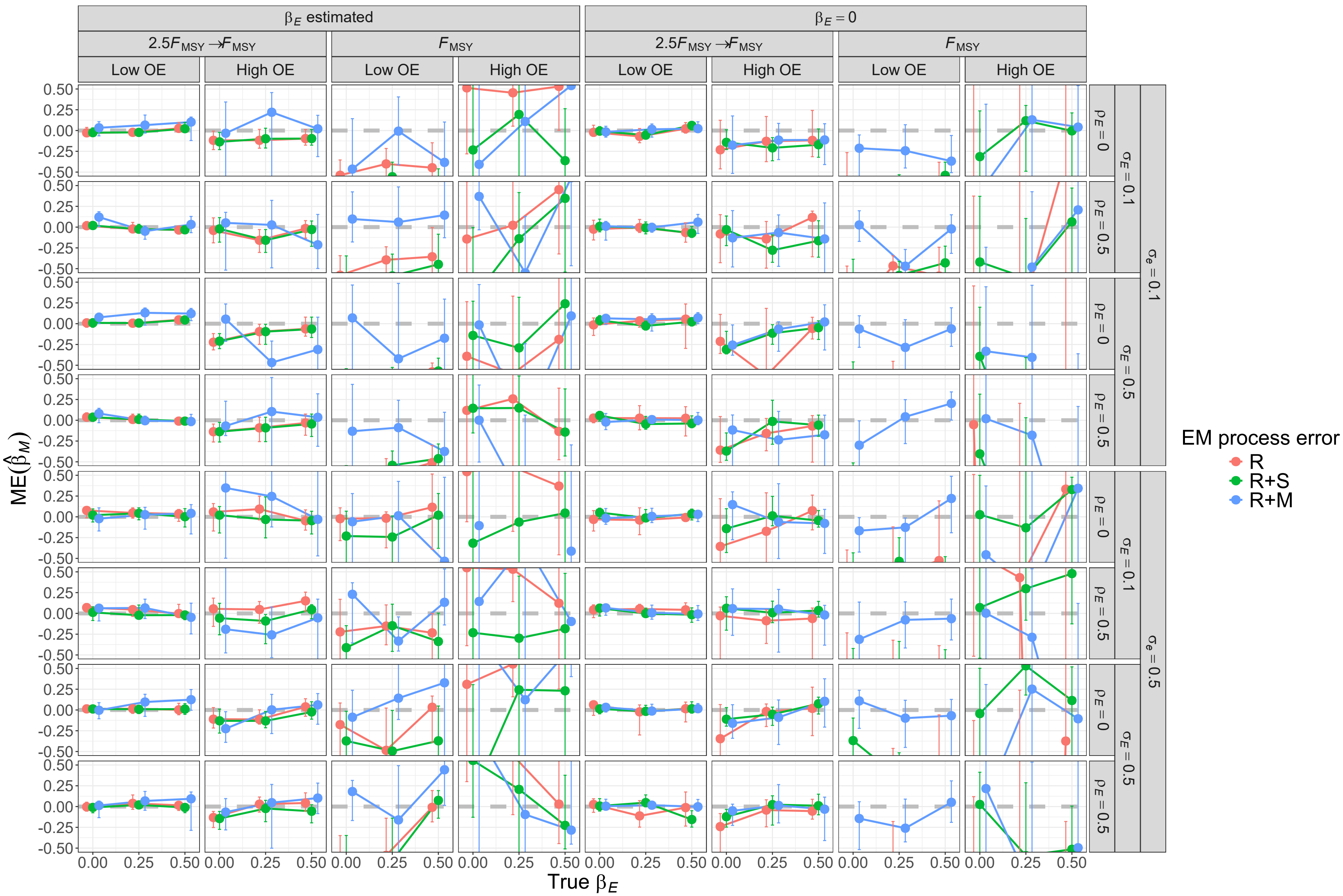

Fig. S22. For R+M OMs, median error (ME) of estimates of β_M from fitting EMs with alternative process error assumptions and treatment of covariate effect ($\beta_E = 0$ or estimated). Vertical lines represent 95% confidence intervals.

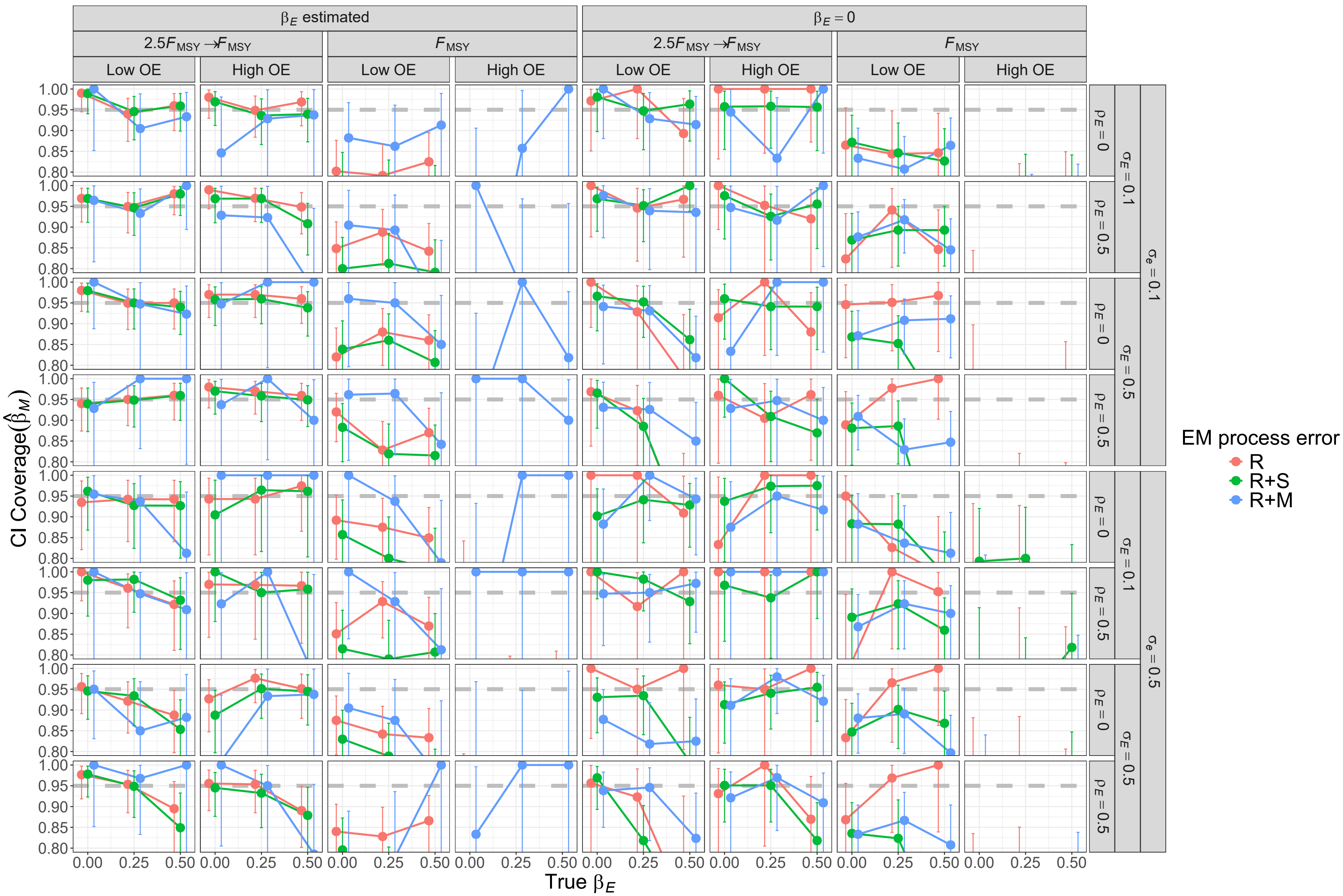

Fig. S23. For R OMs, probability of 95% confidence interval for β_M containing the true value for EMs with alternative process error assumptions and treatment of covariate effect ($\beta_E = 0$ or estimated). Vertical lines represent 95% confidence intervals.

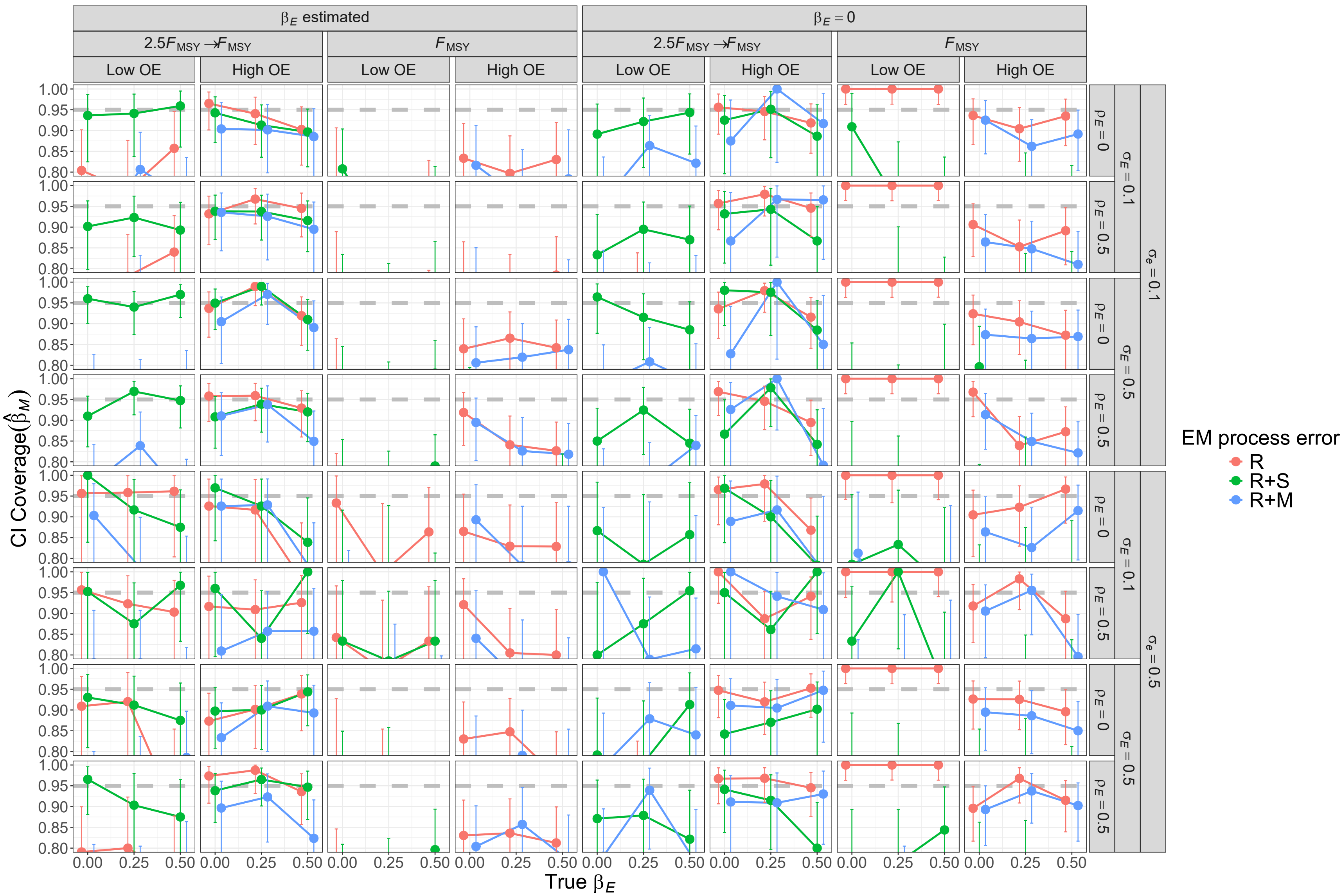

Fig. S24. For R+S OMs, probability of 95% confidence interval for β_M containing the true value for EMs with alternative process error assumptions and treatment of covariate effect ($\beta_E = 0$ or estimated). Vertical lines represent 95% confidence intervals.

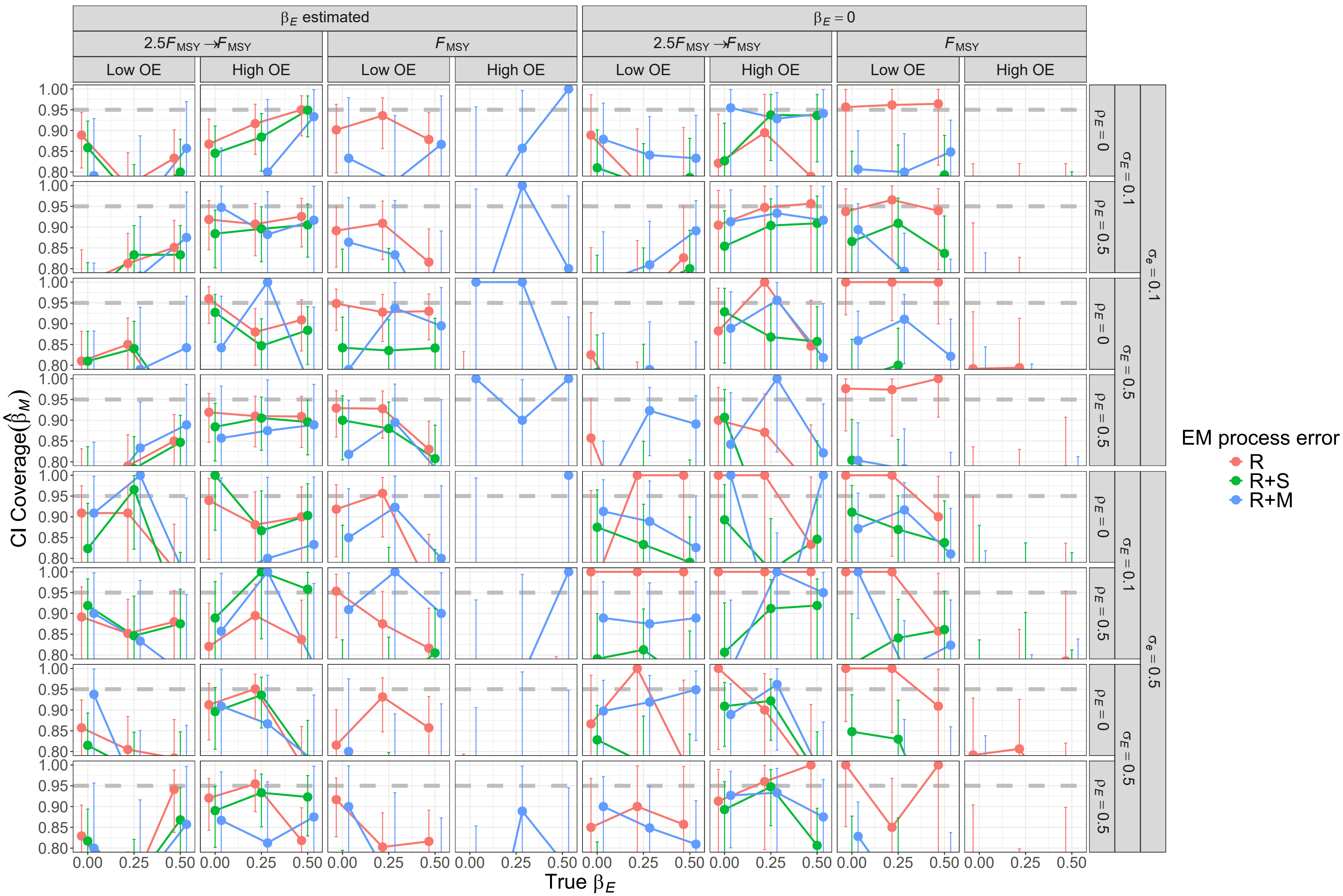

Fig. S25. For R+M OMs, probability of 95% confidence interval for β_M containing the true value for EMs with alternative process error assumptions and treatment of covariate effect ($\beta_E = 0$ or estimated). Vertical lines represent 95% confidence intervals.

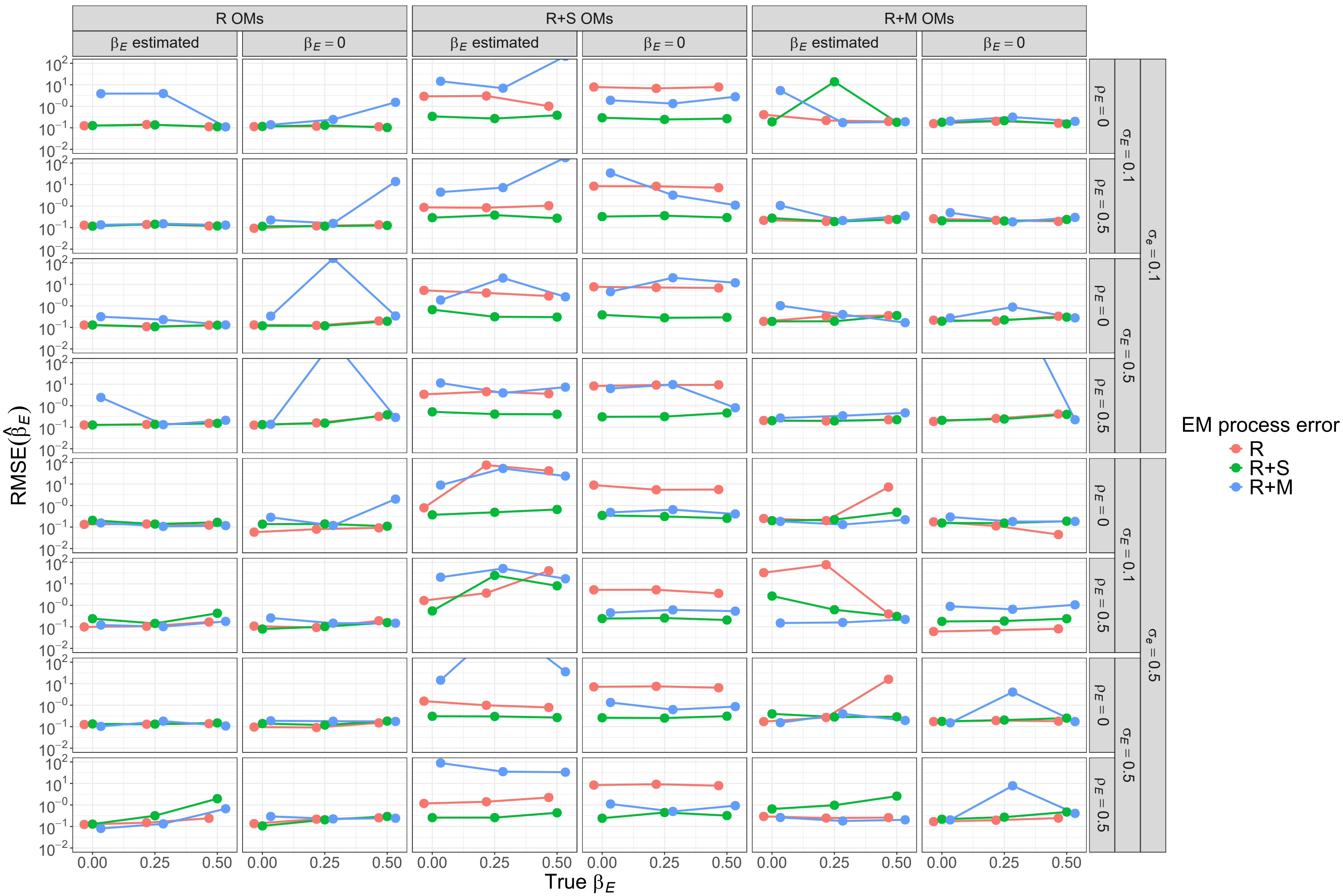

Fig. S26. Root mean square error (RMSE) of estimates of β_M from fitting EMs with alternative process error assumptions and treatment of covariate effect ($\beta_E = 0$ or estimated). All OMs had low observation error and contrast in fishing mortality.

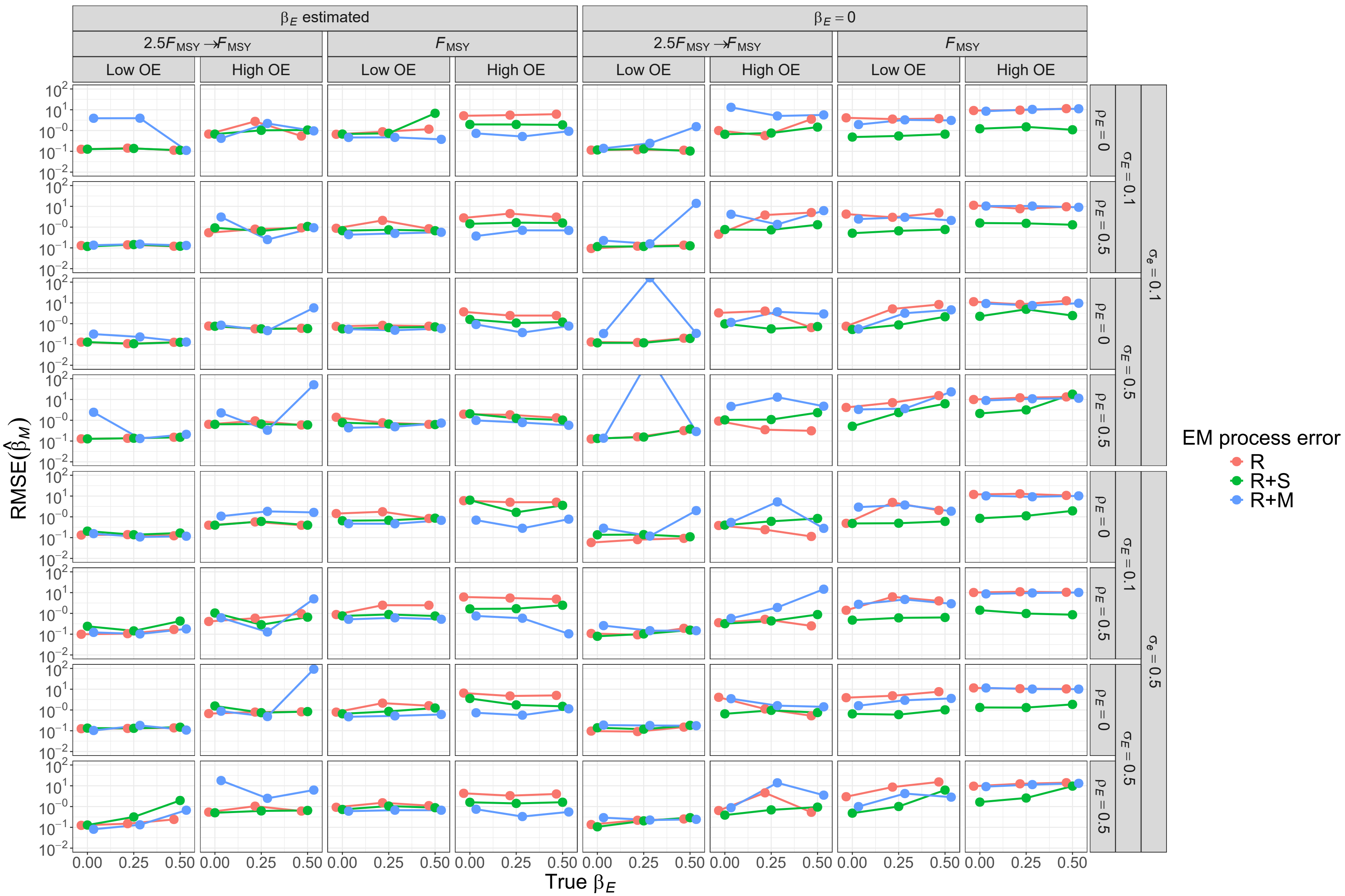

Fig. S27. For R OMs, root mean square error (RMSE) of estimates of β_M from fitting EMs with alternative process error assumptions and treatment of covariate effect ($\beta_E = 0$ or estimated).

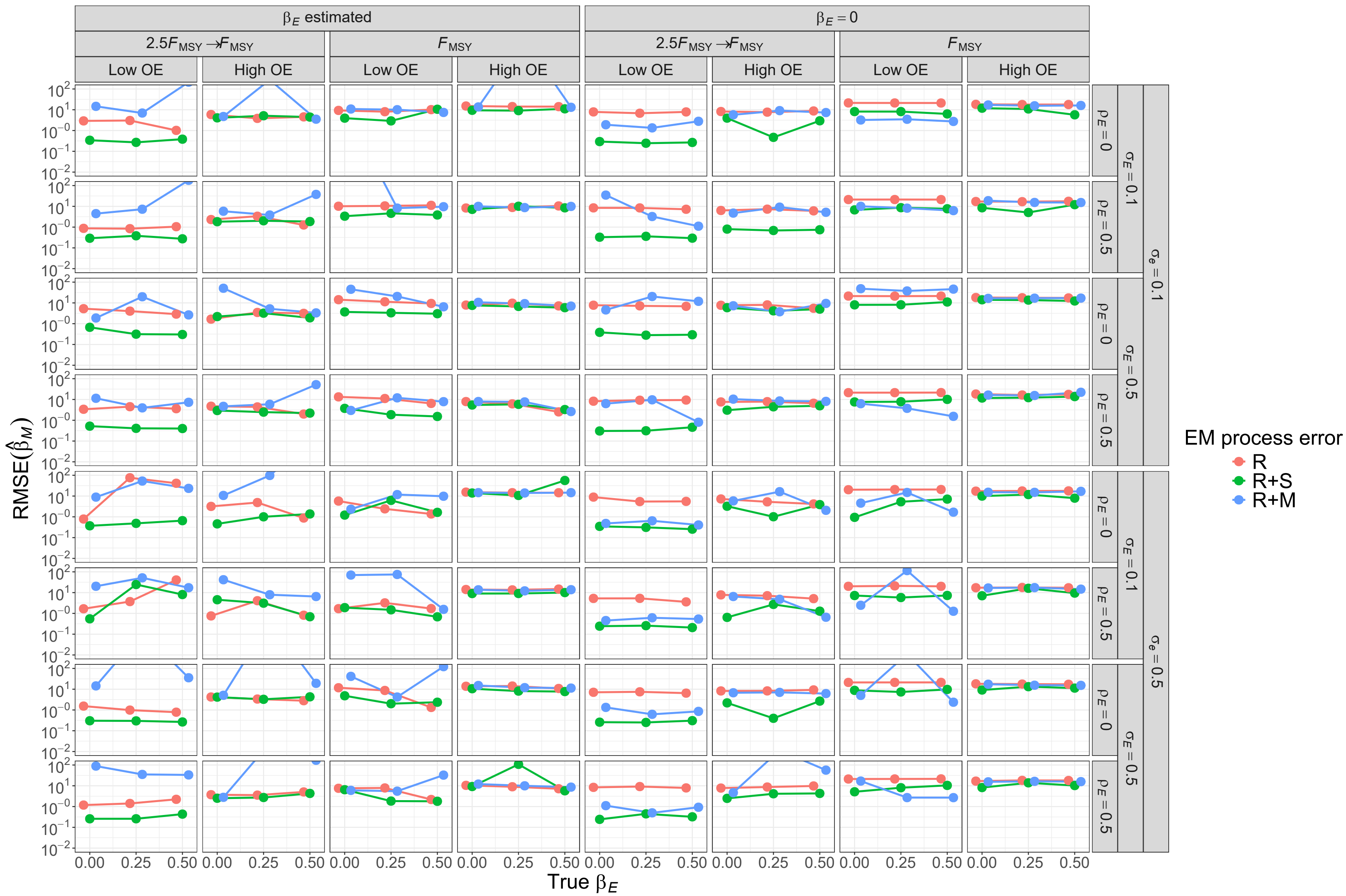

Fig. S28. For R+S OMs, root mean square error (RMSE) of estimates of β_M from fitting EMs with alternative process error assumptions and treatment of covariate effect ($\beta_E = 0$ or estimated).

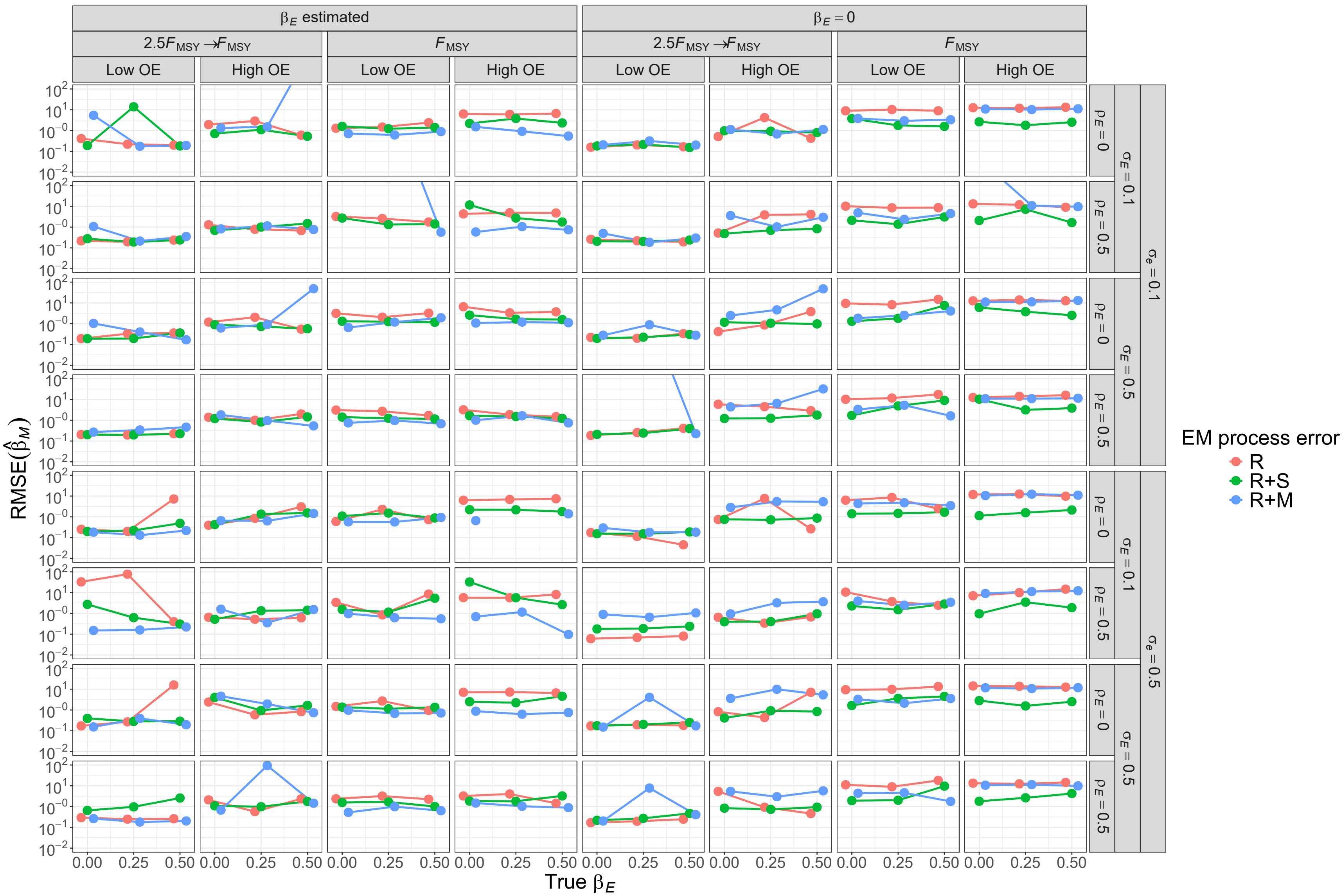

Fig. S29. For R+M OMs, root mean square error (RMSE) of estimates of β_M from fitting EMs with alternative process error assumptions and treatment of covariate effect ($\beta_E = 0$ or estimated).

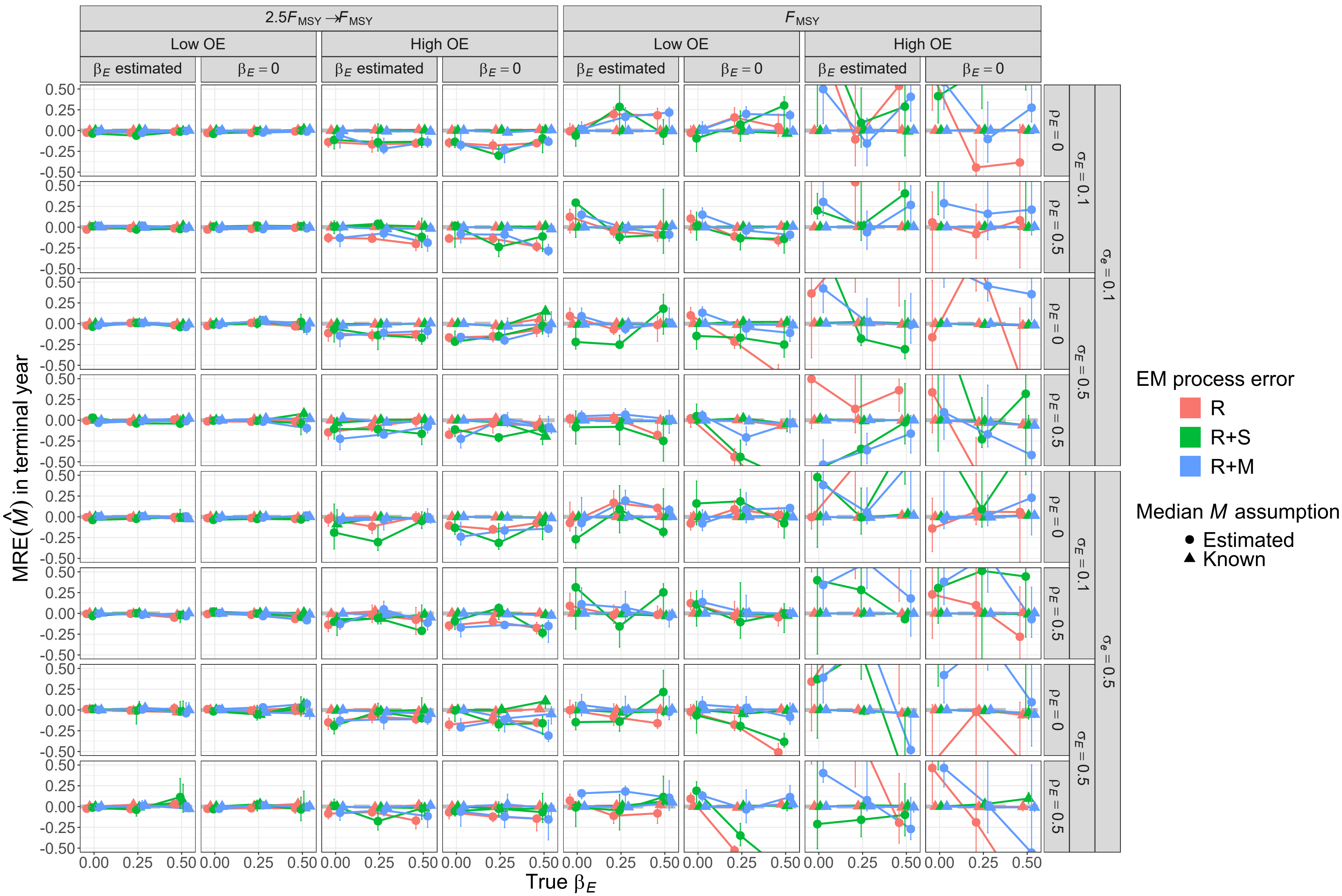

Fig. S30. For R OMs, median relative error (MRE) of estimates of natural mortality rate in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known).

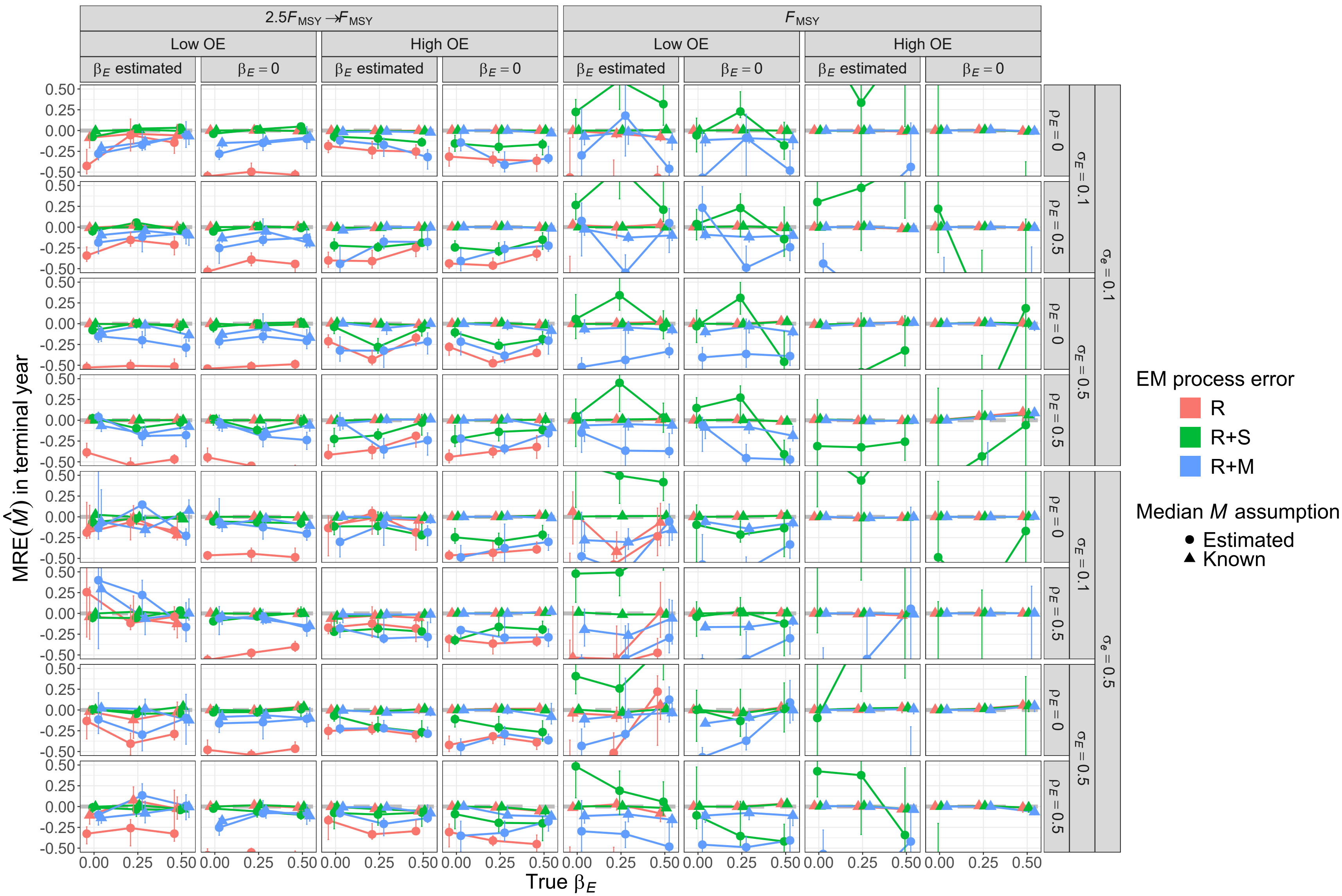

Fig. S31. For R+S OMs, median relative error (MRE) of estimates of natural mortality rate in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known).

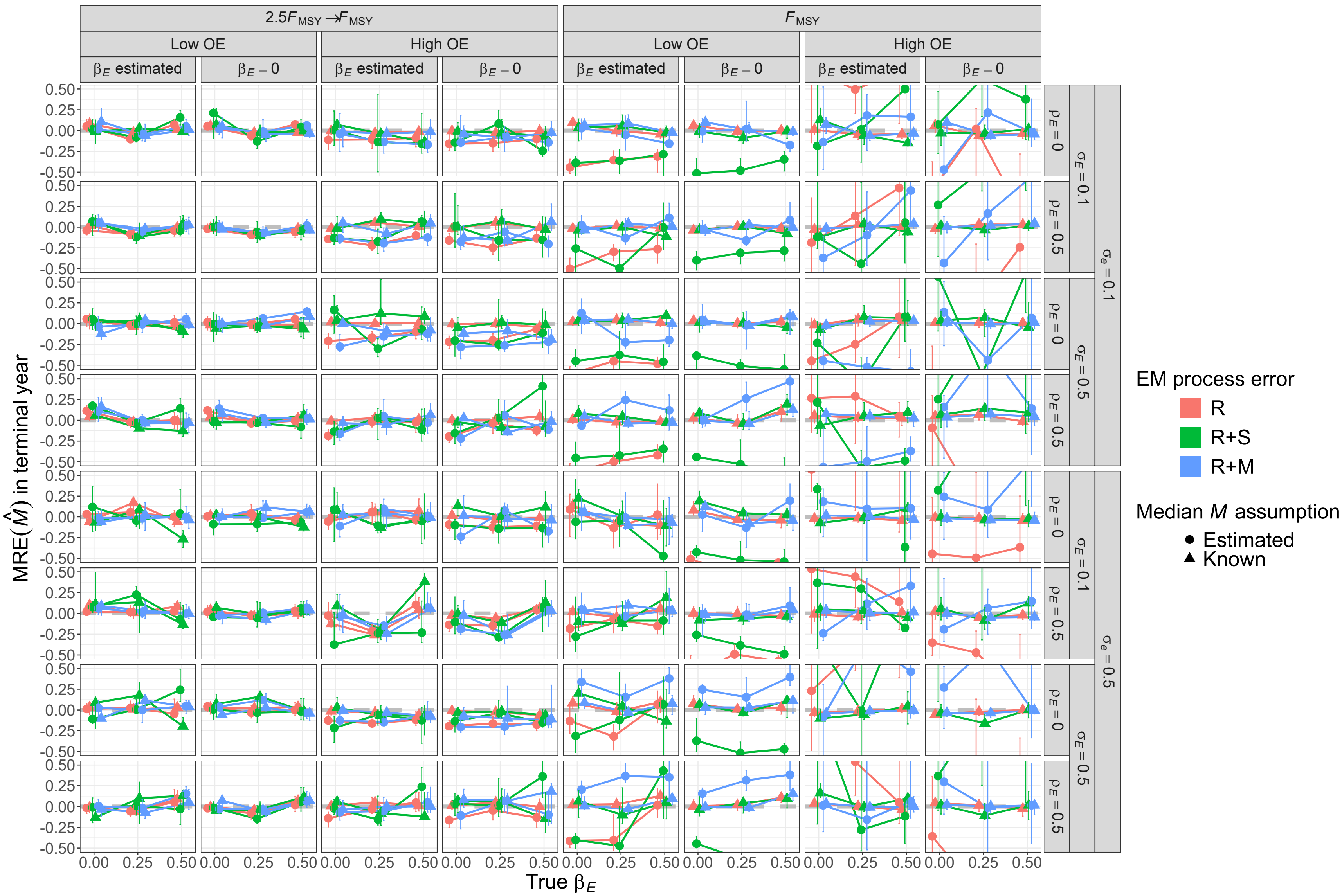

Fig. S32. For R+M OMs, median relative error (MRE) of estimates of natural mortality rate in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known).

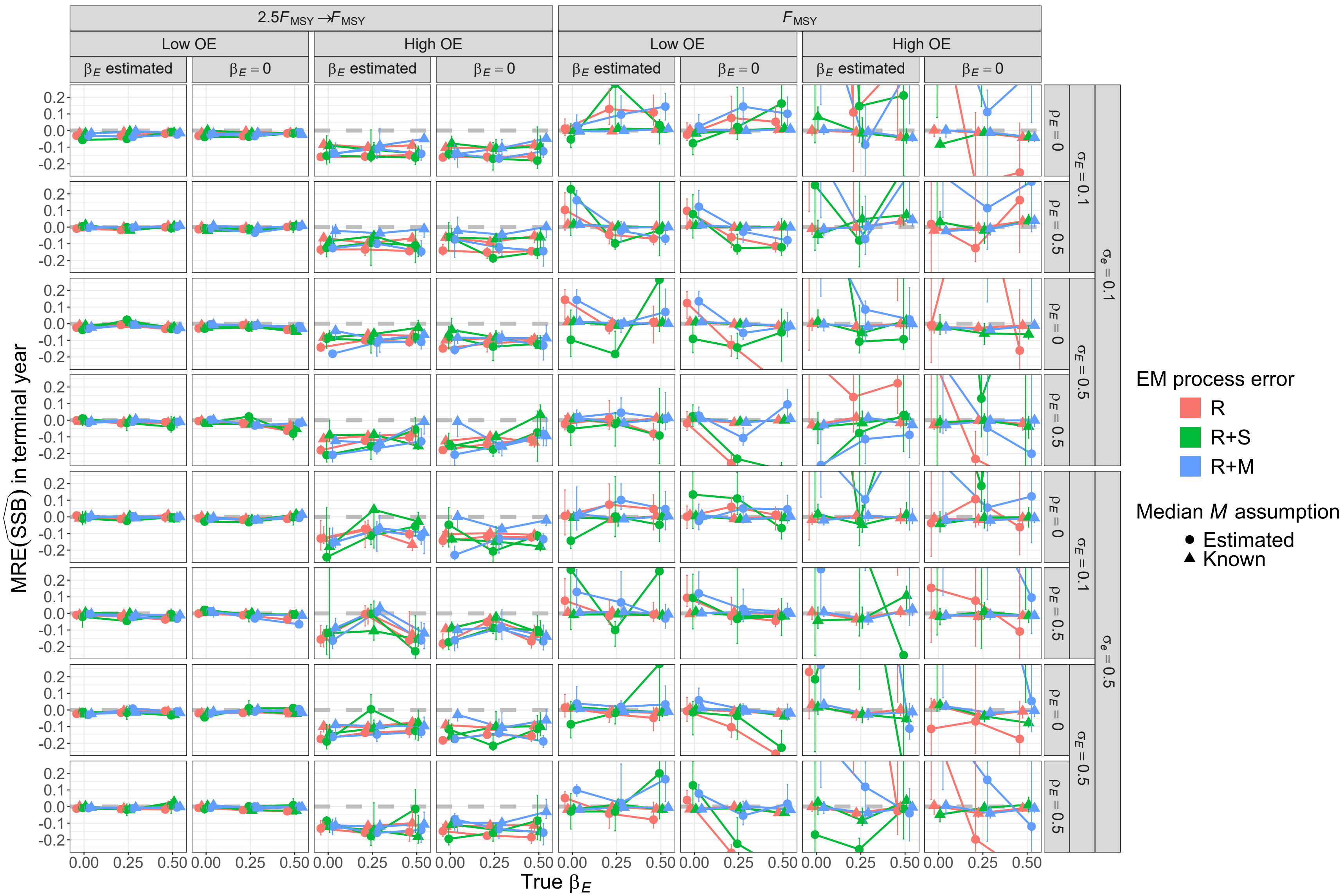

Fig. S33. For R OMs, median relative error (MRE) of estimates of spawning stock biomass (SSB) in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known).

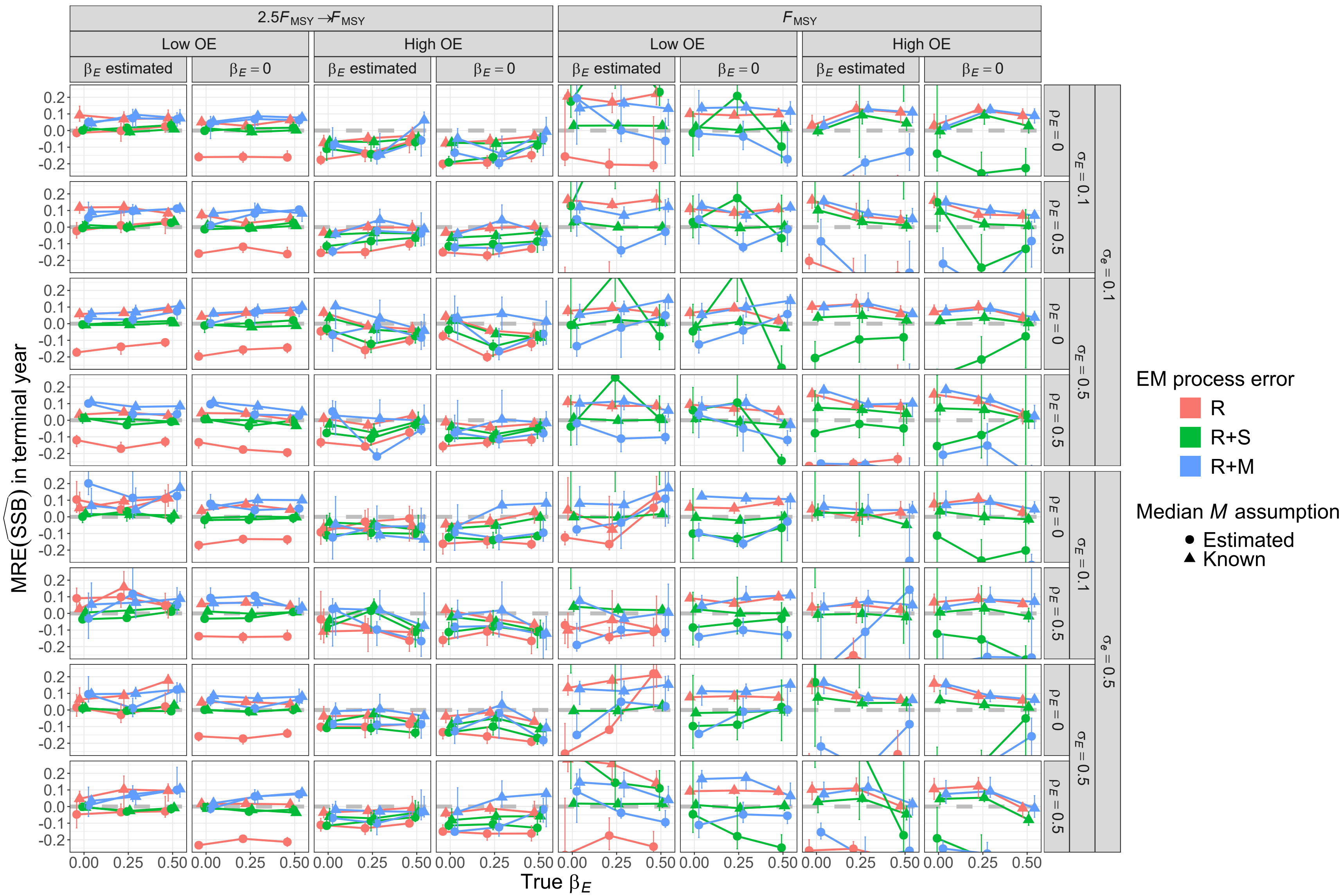

Fig. S34. For R+S OMs, median relative error (MRE) of estimates of spawning stock biomass (SSB) in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known).

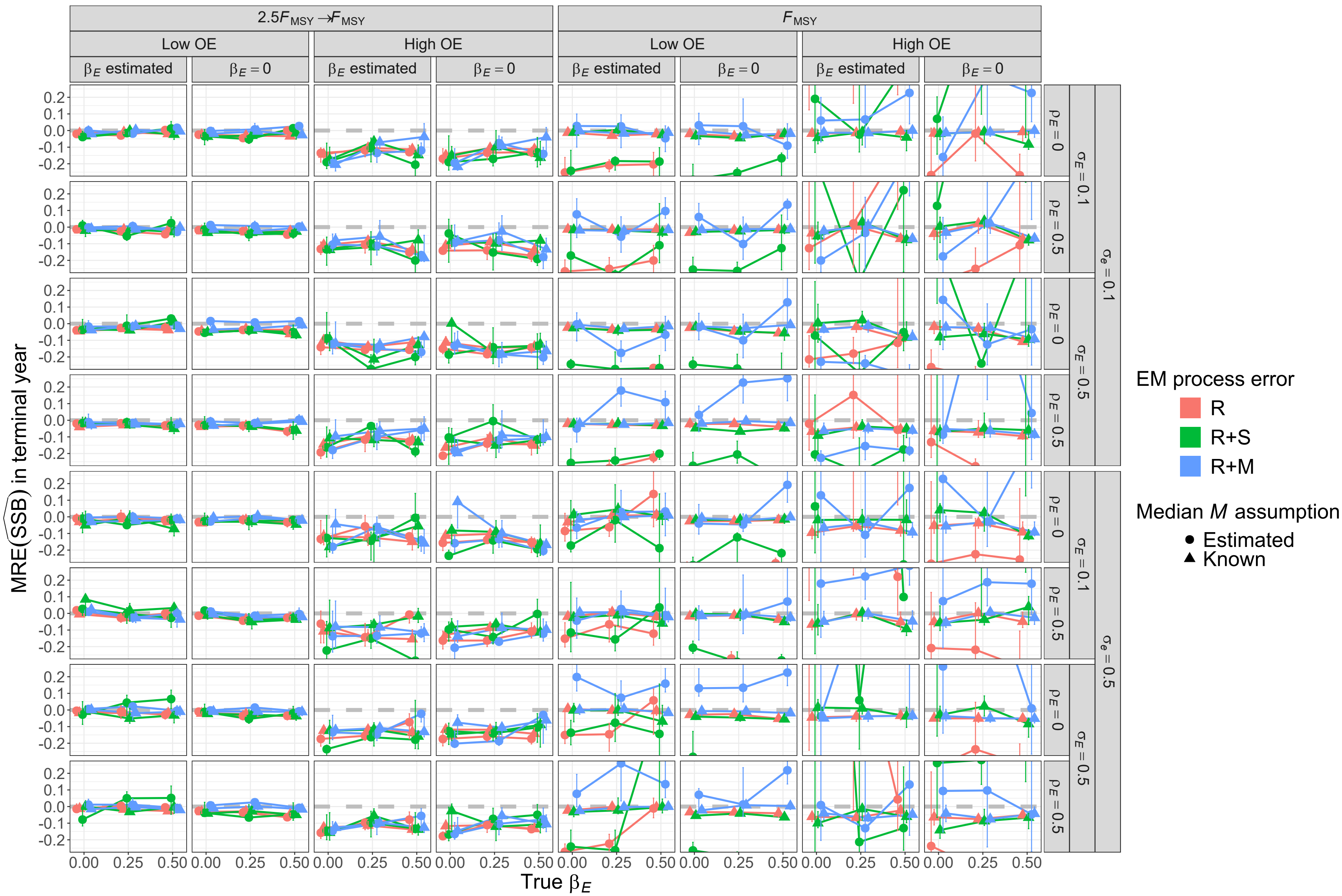

Fig. S35. For R+M OMs, median relative error (MRE) of estimates of spawning stock biomass (SSB) in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known).

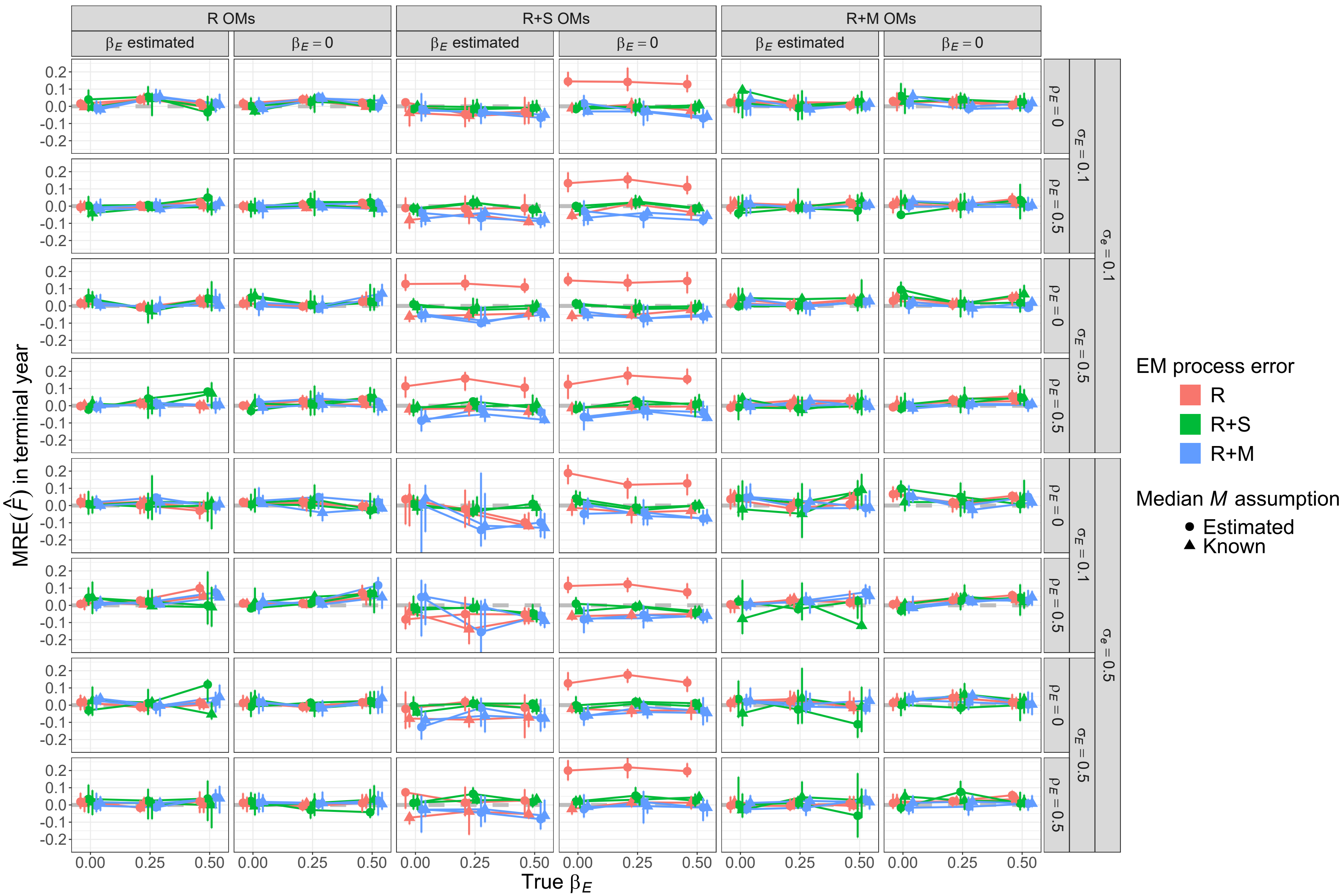

Fig. S36. Median relative error (MRE) of estimates of fully-selected fishing mortality (F) in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known). All OMs had low observation error and contrast in fishing mortality.

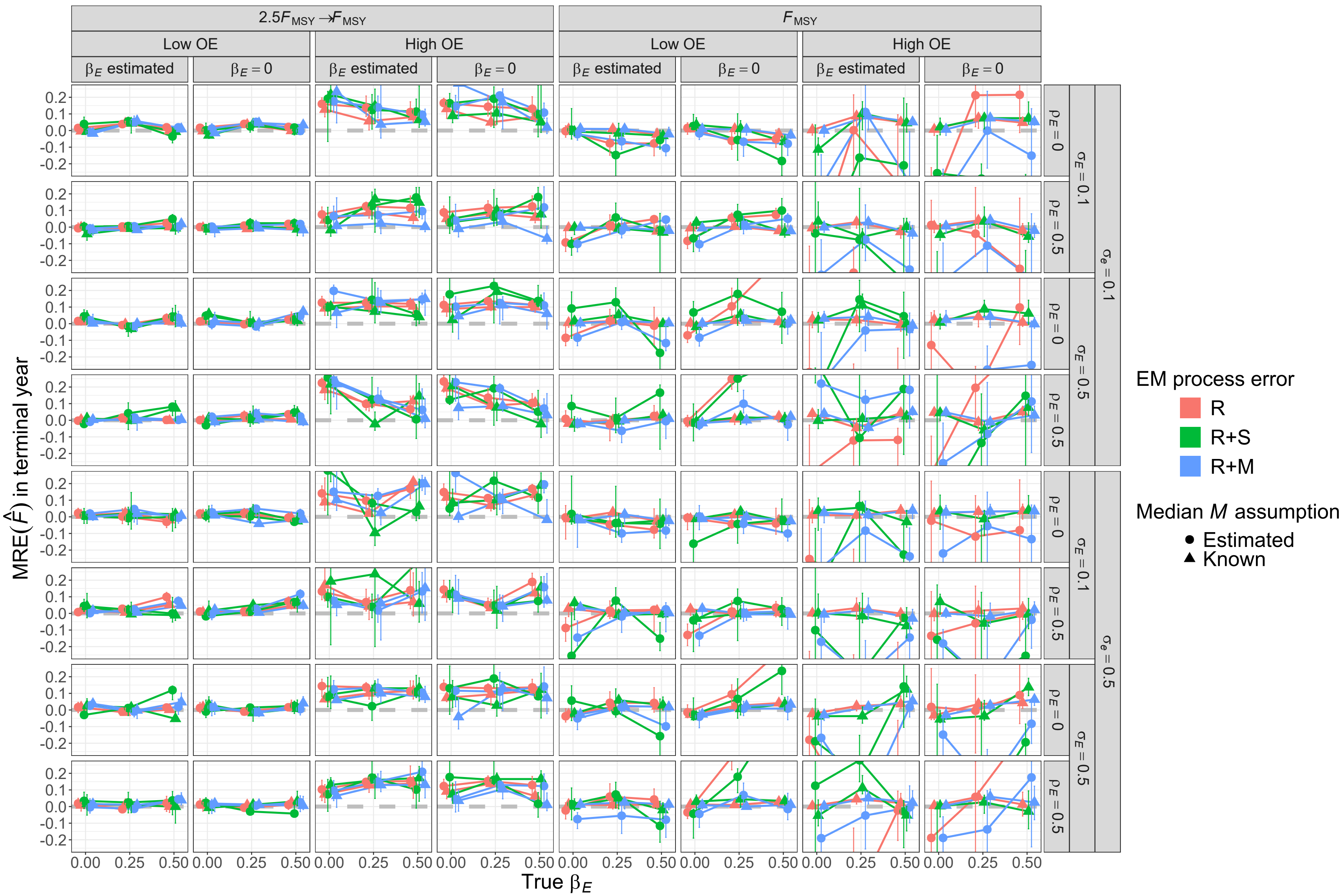

Fig. S37. For R OMs, median relative error (MRE) of estimates of fully-selected fishing mortality (F) in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known). All OMs had low observation error and contrast in fishing mortality.

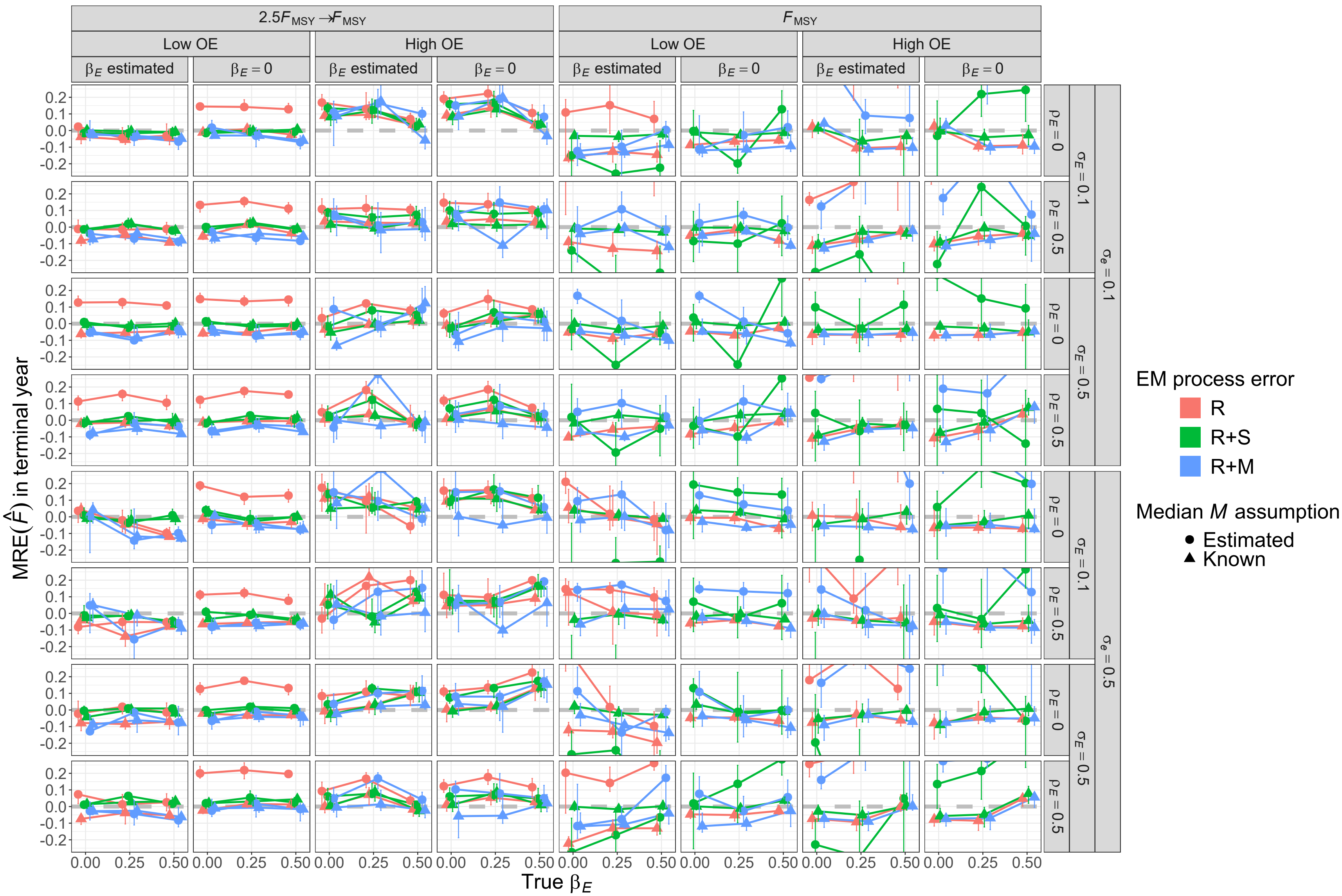

Fig. S38. For R+S OMs, median relative error (MRE) of estimates of fully-selected fishing mortality (F) in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known). All OMs had low observation error and contrast in fishing mortality.

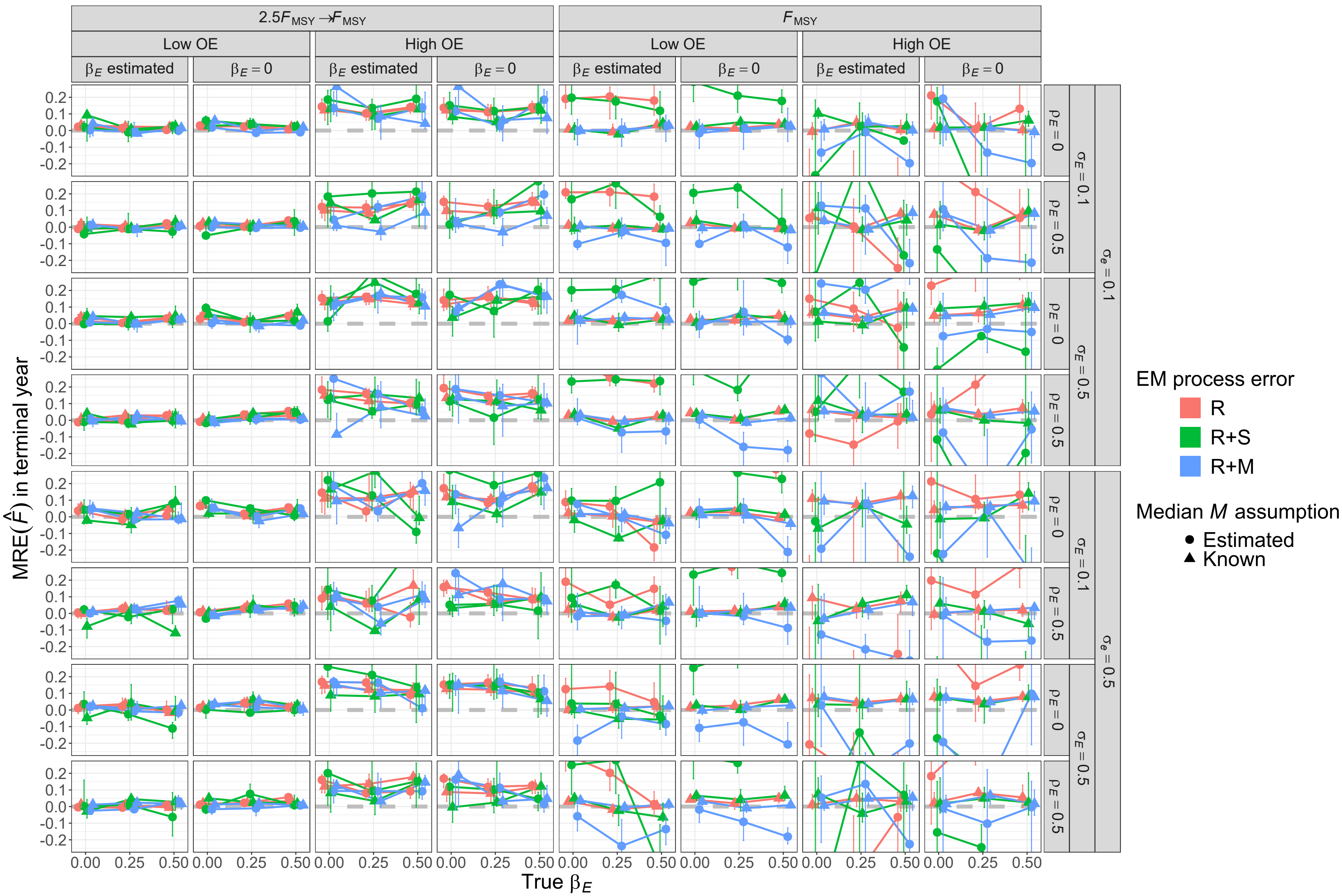

Fig. S39. For R+M OMs, median relative error (MRE) of estimates of fully-selected fishing mortality (F) in the terminal year for EMs with alternative process error assumptions, treatment of covariate effect ($\beta_E = 0$ or estimated), and treatment of median natural mortality parameter (β_M estimated or known). All OMs had low observation error and contrast in fishing mortality.