Comparison of equilibrium and nonequilibrium estimators for the generalized production model

Erik H. Williams and Michael H. Prager

Abstract: Parameter estimation for the logistic (Schaefer) production model commonly uses an observation-error estimator, here termed the NM estimator, combining nonlinear-function minimization and forward projection of estimated population state. Although the NM estimator can be used to fit the generalized (Pella–Tomlinson) production model, an equilibrium approximation method (EA) is often used, despite calls in the literature to abandon equilibrium estimators. We examined relative merits of NM and EA estimators for the generalized model by fitting 48 000 simulated data sets describing five basic stock trajectories and widely varying population characteristics. Simulated populations followed the generalized production model exactly but were observed with random error. Both estimators were used on each data set to estimate several quantities of management interest. The estimates from NM were usually more accurate and precise than EA estimates, and overall, NM outperformed EA. This is the most comprehensive study of the question to date—the first to examine the generalized production model—and it demonstrates clearly why equilibrium estimators should be abandoned. Although valuable when introduced, they are no longer computationally necessary, they are no less demanding of data, and their performance is poor.

Résumé : L'estimation des paramètres du modèle logistique de production (Schaefer) fait ordinairement intervenir un estimateur de l'erreur d'observation, l'estimateur NM, qui combine une minimisation d'une fonction non linéaire et une projection vers le futur d'une estimation de l'état de la population. Bien que l'estimateur NM puisse être utilisé pour ajuster la forme généralisée du modèle de production (Pella–Tomlinson), on emploie souvent une méthode d'approximation d'équilibre (EA), même si de nombreux auteurs déconseillent cette utilisation. Nous avons comparé le mérite des estimateurs NM et EA utilisés dans le modèle généralisé en ajustant 48 000 ensembles de données simulées qui décrivent cinq évolutions de stocks fondamentales et des caractéristiques démographiques très variées. Les populations simulées suivent le modèle avec précision, mais elles affichent des erreurs aléatoires. Les deux estimateurs ont été employés sur chaque ensemble de données pour calculer plusieurs valeurs d'intérêt pour la gestion. Les estimations faites à partir de NM sont généralement plus précises et, tout compte fait, l'estimateur NM fonctionne beaucoup mieux que EA. Notre étude de ce problème est la plus vaste à ce jour et elle démontre clairement pourquoi les estimateurs EA doivent être abandonnés. Bien qu'utiles lors de leur introduction, ces estimateurs ne sont plus nécessaires, étant donné les ressources informatiques actuellement disponibles. De plus, ils exigent autant de données et leur performance est faible.

[Traduit par la Rédaction]

Introduction

Simple surplus production models (those without age structure) continue to prove useful in stock assessment, both when age data are lacking or unreliable and to provide direct estimates of maximum sustainable yield (MSY), an important measure of potential stock productivity (e.g., Zhang et al. 1991). As with all models, care in implementation and choice of statistical methods can be as important as model definition itself (Xiao 1997, 2000). For example, Fox (1971) described the importance of the form of objective function in fitting the generalized production model of Pella and Tomlinson (1969); Schnute (1989) demonstrated the importance of choice of conditioning (the assumption on which observed quantity, yield or relative abundance, should be assumed known without error) in fitting a logistic surplus-production model (Schaefer 1954, 1957), which he used to exemplify a relatively simple fishery model. Ludwig et al. (1988) and Polacheck et al. (1993) compared observation-error and process-error estimators, among other issues of statistical model structure.

Current practice in parameter estimation for the simplest production model, the logistic (Graham–Schaefer) model (Graham 1935; Schaefer 1954, 1957), usually takes the form of an observation-error estimator comprising nonlinear-function minimization and forward projection of estimated population state. That method of parameter estimation, here denoted NM (for nonlinear minimization), was introduced to fishery science by Pella (1967) and Pella and Tomlinson (1969) and was modified by Prager (1994) to incorporate a logarithmic objective function (reflecting assumed lognormal errors in the observed data) and conditioning on yield, in keeping with findings of Fox (1971) and Schnute (1989). The method

Received 7 February 2002. Accepted 4 September 2002. Published on the NRC Research Press Web site at http://cjfas.nrc.ca on 26 September 2002.
J16752

E.H. Williams[1] and M.H. Prager. National Oceanic and Atmospheric Administration, Center for Coastal Fisheries and Habitat Research, 101 Pivers Island Road, Beaufort, NC 28516, U.S.A.

[1]Corresponding author (e-mail: erik.williams@noaa.gov).

 DOI: 10.1139/F02-123

is also described by Hilborn and Walters (1992), Quinn and Deriso (1999), and Haddon (2001). Similar methods were introduced for catch-at-age models by Methot (1989) and, because of their flexibility, are increasingly used in fishery science for both simple and complex models.

The NM estimator avoids the so-called "equilibrium assumption", in which observed yields are assumed to represent equilibrium yields, an assumption usually made about data on fishing effort averaged over several years; the averaging is said to approximate equilibrium conditions (Gulland 1961; Fox 1975*a*). Equilibrium estimators make short-cut methods of parameter estimation possible—and that was of great importance before the wide availability of small, powerful computers—but they lack a solid theoretical foundation.

Several studies of the logistic production model have found that significant biases can occur when using equilibrium estimators. Sissenwine (1978) pointed out that equilibrium methods can provide an answer, if a possibly misleading one, when data do not support estimation by nonequilibrium methods. That occurs not because equilibrium estimators require less extensive or less precise data (which they do not), but because they are generally linear approximations, and thus least-squares parameter estimates can be computed easily. In contrast, most nonequilibrium estimators rely on nonlinear optimization techniques that may fail to converge when the data conflict with the model.

Mohn (1980) evaluated equilibrium estimators on a simulated stock trajectory of increasing effort, stock collapse, and recovery. Results from simulated data with two levels of population productivity, four levels of noise, and variations of several other parameters were reported, with 50 replications of each. Mohn (1980) found that equilibrium estimators in general overestimated MSY and f_{MSY}, the rate of fishing effort rate at which MSY can be attained. By fitting sections of the population trajectories, Mohn (1980) found that the trend in population level affected the degree of bias in the equilibrium estimates. Less bias in f_{MSY} but greater overestimation of MSY occurred from an increasing trend in population levels than from a decreasing trend (Mohn 1980).

Roff and Fairbairn (1980) noted that bias in the equilibrium method of parameter estimation is not alleviated by increasing the number of years used in the effort averaging and argued that, on statistical grounds alone, the equilibrium method cannot be justified.

The most extensive previous comparison of equilibrium and nonequilibrium methods was by Polacheck et al. (1993). Using population parameters from two fits to data on three real fish stocks, the authors simulated 500 data sets for each set of parameters. Those data were then fit with three estimators for the logistic production model: an observation-error estimator similar to the NM estimator used here, a process-error estimator, and an effort-averaging method similar to the equilibrium-approximation (EA) estimator used here. Of the three, the observation-error estimator gave the least biased and most precise estimates of MSY and f_{MSY}. The effort-averaging estimator produced seemingly reasonable estimates of MSY and f_{MSY} that were nonetheless positively biased. The performance of the equilibrium estimator led Polacheck et al. (1993) to recommend "that effort-averaging (equilibrium) methods be abandoned". As will be seen, our study is more exhaustive than that of Polacheck et al. (1993), and unlike that study, we examine estimators for the generalized production model.

Current practice in parameter estimation for the generalized production model (Pella and Tomlinson 1969) is more varied than for the logistic model. Despite the introduction of nonequilibrium fitting methods over 30 years ago by Pella and Tomlinson (1969), the equilibrium-approximation (here, EA) estimation method due to Fox (1975*a*) after Gulland (1961) is still in current use, e.g., for assessments of some tuna species (ICCAT 2000), perhaps because fitting the generalized model without the approximation can be difficult. That difficulty stems from at least three causes: model structure, data insufficiency, and implementation quality. Fletcher (1978) restructured the model equations to give the parameters understandable meanings and reduce correlation among parameters; nonetheless, the model shape can be difficult to estimate. By model shape, we mean the degree and direction of asymmetry in the production curve (Fig. 1).

That many data sets possess insufficient information to specify model shape unambiguously was noted by Fletcher (1978) and Hilborn and Walters (1992), among others. Prager (2002) confirmed Fletcher's (1978) observation that the problem of data insufficiency can be accentuated by anomalous observations in the data set.

Implementation of the generalized model is not trivial: if a continuous-time formulation is used, numerical approximations are required to integrate the catch equation and to obtain its solution when conditioning on yield. Even when using a discrete-time model, the objective function can be characterized by many local minima. Thus, the computational simplicity of the EA estimator is appealing. Given this situation, a natural topic of inquiry is the comparative performance of the two methods, EA and NM, on varied data sets. That topic is the main subject of this study.

Methods

Overview

We employed computer simulation to compare the NM and EA estimators for the generalized production model, using methods (Fig. 2) described in detail below. The initial step in the study was generating numerous data sets representing simulated fished populations. Each simulated population followed the dynamics of the generalized production model exactly, a choice made not because real fish populations follow any model exactly, but because a logical start in evaluating estimators is checking their performance on data that meet their most important assumption (model structure). Simulated populations were "observed" by recording total stock biomass and removals (in arbitrary units, here designated kg) taken each time period (designated years). Simulated observation error was added to each year's data, the resulting data sets were fit using both estimators, and the resulting estimates of management-related quantities were compared (graphically and numerically) to the true underlying values, known from the simulations.

Fig. 1. Production curves characteristic of simulated populations used in study. Curves correspond to $\phi = 0.35$ (short-dashed line), 0.50 (solid line), and 0.65 (long-dashed line).

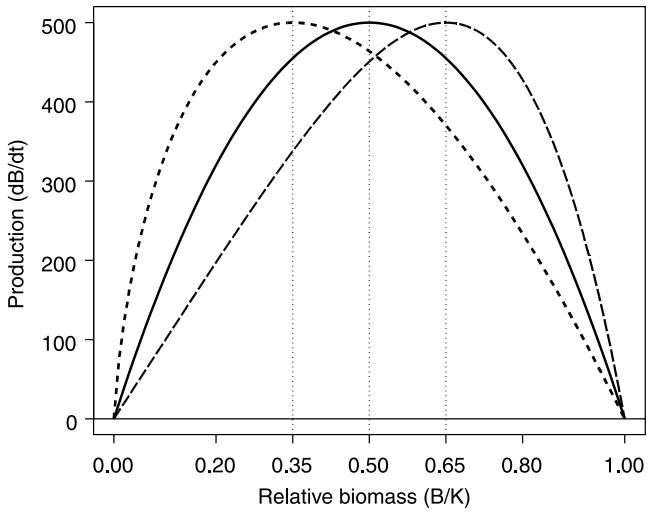
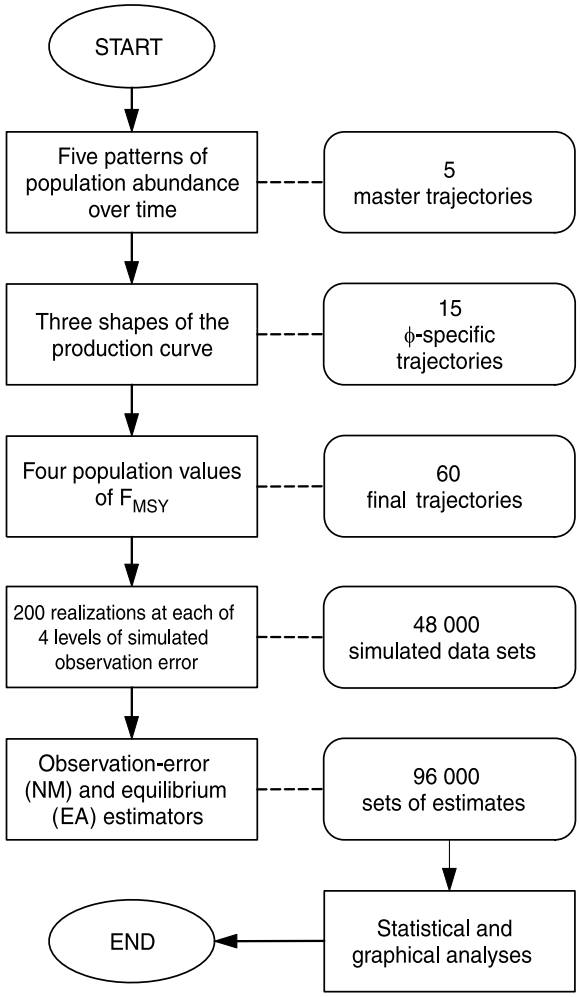


Fig. 2. Schematic of study: generation of simulated data, fitting, and evaluation.



Simulation of data

A simulated population following production-model dynamics is characterized by three basic quantities; in the parameterization chosen (Fletcher 1978; Prager 2002), these are maximum sustainable yield, MSY; carrying capacity, K; and model shape, ϕ. We define $\phi = B_{MSY}/K$, where B_{MSY} is the population biomass at which maximum surplus production occurs and $0 < \phi < 1$ is a dimensionless parameter that defines the shape of the production curve, i.e., the curve of surplus production against population biomass (Fig. 1). To generate a specific population trajectory, it is also necessary to specify the biomass in the first year, B_1, and the catchability coefficient, q, that relates the simulated abundance index I_t to the actual size of the population B_t by $I_t = qB_t$ (where subscript t indicates time in steps of 1 year). Here, each population was characterized by $K = 1\,000\,000$ kg and its abundance index by $q = 1$. We varied three quasi-biological characteristics of these simulated populations: general shape of the historical biomass trajectory, shape of the production curve, and value of F_{MSY}, the fishing mortality rate associated with MSY (Fig. 2). Because we set F_{MSY} and ϕ for each simulated population, MSY was not set directly, but rather by the relationship $MSY = \phi K F_{MSY}$. Starting biomass B_1 was set for each simulated population based on the shape of the biomass time trajectory and the value of ϕ, as described below. Each simulated data set was 40 years long.

Populations were projected using the reformulated Pella–Tomlinson model of Fletcher (1978). Under that model, the biomass rate of change is given by

$$(1)\qquad \frac{dB_t}{dt} = \gamma MSY\frac{B_t}{K} - \gamma MSY\left(\frac{B_t}{K}\right)^n - q f_t B_t$$

where the third term describes removals by fishing and the notational quantity γ is given by

$$\gamma = \frac{n^{n/(n-1)}}{n-1}$$

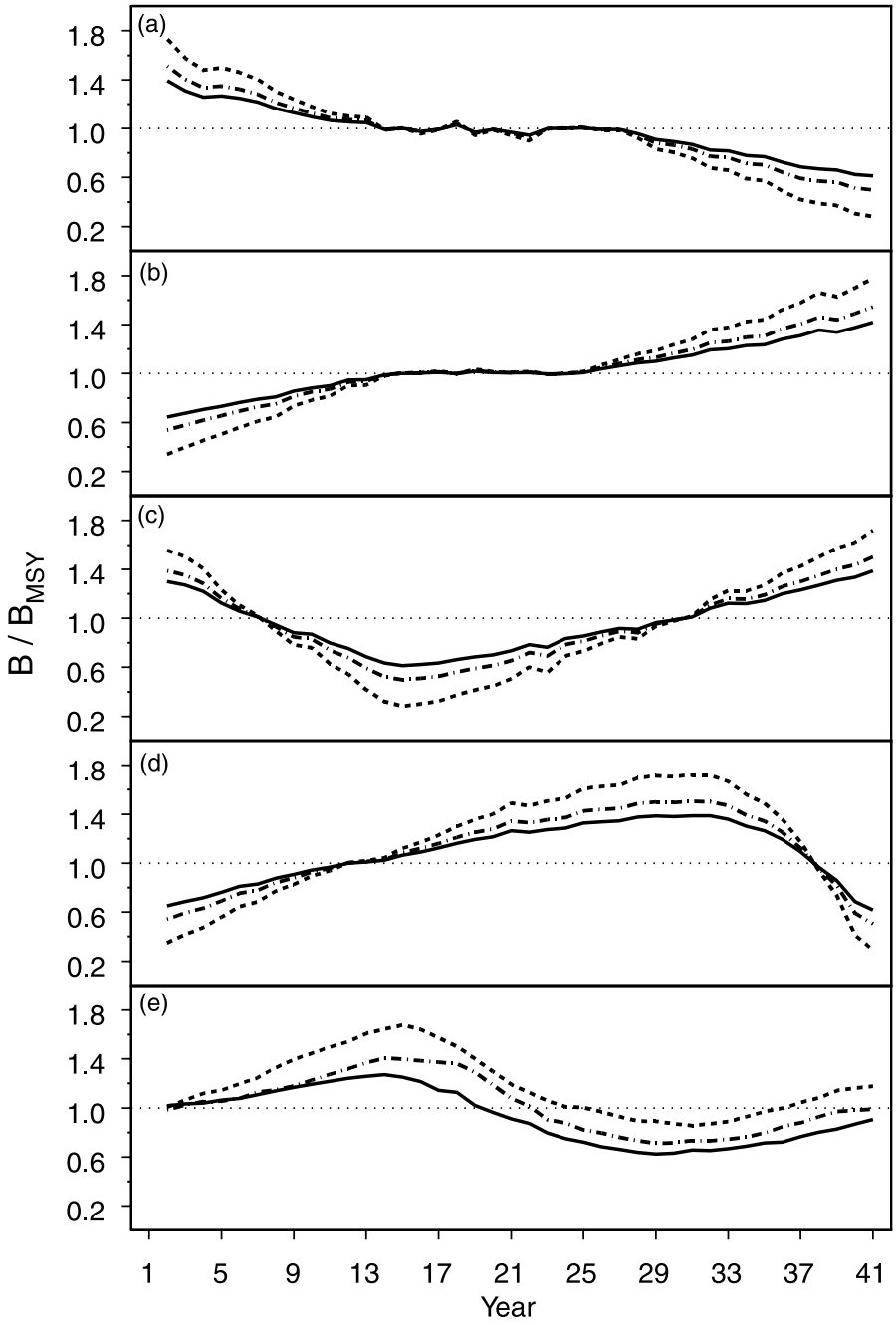
In eq. 1, shape of the production curve is determined by the exponent n, but in this paper, we describe model shape by the dimensionless parameter ϕ. The two are related by $\phi = B_{MSY}/K = (1/n)^{1/(n-1)}$ (Fletcher 1978).

Starting from the simulated value of B_1, each simulated population was projected forward by numerical integration of eq. 1, using 90 time steps per year, which provides a close approximation to a continuous-time model. Simulated annual catches and start-of-year biomass levels were recorded for fitting.

Master trajectories

The sequence of data generation and analysis was simple, yet the number of factors considered and desire for sufficient replication led to a large number of simulated data sets (Fig. 2). To serve as the basis of all simulations, five master biomass trajectories were generated as arbitrarily scaled time series of annual biomass, B_t/B_{MSY}. The five master trajectories can be characterized as decreasing (trajectory 1), increasing (trajectory 2), concave (trajectory 3), convex (trajectory 4), and complex (trajectory 5); together, they represent a wide range of possible stock dynamics (illustrated after scaling to B_{MSY} by the dash–dot lines in Fig. 3).

Fig. 3. Simulated population trajectories used to study competing estimators for generalized production model. Curves correspond to ϕ = 0.35 (short-dashed line), 0.50 (dash–dot line), and 0.65 (solid line).



Shapes of the production curve

The generalized production model is used to accommodate possible skewness of the production curve. Each of the five master biomass trajectories was transformed to each of three shapes of the production curve at which data were simulated, $\phi = \{0.35, 0.5, 0.65\}$ (Fig. 1), to result in 15 ϕ-specific trajectories (Figs. 2, 3). The value $\phi = 0.35$ is very close to the shape of the Fox (1970) exponential yield model, for which $\phi = 1/e \approx 0.37$. The central value $\phi = 0.5$ defines the logistic (Schaefer) model, and $\phi = 0.65$ is symmetrical with $\phi = 0.35$.

With K held constant (as it was in our simulations), each model shape corresponds to a different value of B_{MSY}, because $B_{MSY} = \phi K$. Transformations made for ϕ were as follows. Master biomass trajectory shapes 1–4 were simply translated up (for $\phi = 0.35$) or down (for $\phi = 0.65$) on the biomass axis so that the corresponding ϕ-specific trajectories passed through B_{MSY} at the same time (Figs. 3*a*–3*d*). The curves in those panels appear to have different shapes for each value of ϕ, because they are shown on a metric of B/B_{MSY}. If plotted in a metric of absolute biomass (or B/K), the three curves within each panel would be parallel.

The adjustments made to trajectory 5 were slightly different. In that case, it proved necessary to further adjust the trajectories to be realizable across a variety of population values of F_{MSY}, as described immediately below. The three trajectories were translated in the same manner as described

for trajectories 1–4, and then heuristic adjustments were made to arrive at realizable trajectories (Fig. 3*e*).

Population values of F_{MSY} and simulation of yield

Four levels of F_{MSY} (0.1, 0.2, 0.4, and 0.8 per year) were chosen to represent a wide range of population productivity. These were applied to the 15 ϕ-specific trajectories to generate 60 final trajectories (Fig. 2). In generating final trajectories derived from master trajectories 1–4, biomass time series were the same for all values of F_{MSY}, but corresponding trajectories of F_t differed. Because MSY = $\phi K F_{MSY}$, the different values of F_{MSY} imply corresponding values of MSY.

Annual fishing mortality (F_t) values corresponding to each of the 60 final trajectories were computed using an iterative algorithm that searched for suitable values under production-model dynamics. The annual values of fishing mortality rate (and the value of B_1/K) were selected to produce the desired final trajectory under the simulated population's characteristic values of ϕ and F_{MSY}. This procedure also provided the simulated yield (catch) for each year.

Observation error

The final step in data simulation was construction of simulated abundance indices by adding lognormal observation error to the simulated biomass in each final trajectory. (Process error was not considered in this study.) In doing this, we generated 800 simulated data sets from each final trajectory (Fig. 2) by adding different random errors 200 times each at four levels of observation error, i.e., at coefficients of variation (CVs) of 2, 5, 15, and 25%. We consider the highest two values the most realistic; the lower values were included to observe results over a wider range. The result was the 48 000 simulated data sets to be used in the fitting part of the study.

Estimators

To implement the EA estimator, we used the Fortran computer program PRODFIT of Fox (1975*b*), slightly modified by us. Modifications included correction of a few obvious coding errors and limiting Fox's user-specified smoothing parameter to a single value, rather than a value for each year; the latter is apparently standard practice among some users of PRODFIT. We also modified the output recorded from PRODFIT to facilitate comparisons between estimators.

Parameters are estimated in the PRODFIT program under a least-squares objective function using a simple pattern-search optimization routine. Before estimation, the program adjusts the recorded annual fishing effort rates in an attempt to approximate equilibrium conditions. This is done by computing annual weighted averages of the fishing effort rate for k years, with k being the number of year-classes "making a significant contribution to the catch" (Fox 1975*a*, 1975*b*). The parameter k must be specified by the user: we set k to the number of year-classes necessary to account for >80% of the catch with a constant total mortality rate set at $Z = 2F_{MSY}$, which corresponds to 9, 5, 3, and 2 year-classes for F_{MSY} = 0.1, 0.2, 0.4, and 0.8 year^{-1}, respectively.

On certain data sets, PRODFIT results indicate a model termed the "asymptotic model" by the program and by Fox (1975*a*), but otherwise not documented. This appears to be a model in which ϕ is asymptotically close to zero; under such a model, maximum sustainable yield would be taken at fishing mortality rates approaching infinity and population sizes approaching zero. Such a model implies that for all $B_t > 0$, $B_t/B_{MSY} = \infty$ and for all $F_t < \infty$, $F/F_{MSY} = 0$. Such results were considered completely unrealistic and thus considered estimation failures.

To implement the NM method, we used a modified version of the ASPIC software of Prager (1995), which employs a nonequilibrium approach to fitting the parameters of the generalized production model. The method includes estimation conditional on observed catch, assuming lognormal observation error in the abundance index. In the ASPIC program, parameters are estimated by minimization of a lognormal objective function using the simplex or "polytope" algorithm of Nelder and Mead (1965). Numerous restarts of the algorithm are used to avoid local minima. Parameters are MSY (maximum sustainable yield), K (carrying capacity), q (catchability coefficient), B_1 (biomass at the start of the first year), and ϕ (shape of the production curve). The major modification to the program described by Prager (1995) was the addition of an algorithm to integrate the generalized model numerically so that its parameters could be estimated. Starting guesses for parameters were obtained from a fit of the logistic production model (ϕ = 0.5). The fitting procedure is described in more detail in Prager (1994) and Prager (2002).

The GENPROD computer program of Pella and Tomlinson (1969) also provides nonequilibrium estimates for the generalized production model. Our estimation approach differs from that of Pella and Tomlinson (1969) mainly in (*i*) conditioning on yield, rather than fishing effort, (*ii*) finer resolution in the numerical integration, (*iii*) use of a logarithmic objective function, (*iv*) use of Fletcher's (1978) restructuring of the generalized model, (*v*) characterization of model shape by the ratio $\phi = B_{MSY}/K$, rather than the exponent, (*vi*) use of a different optimizer (Nelder and Mead 1965), and (*vii*) direct estimation of the shape parameter, rather than use of a coarse grid of values. Many of these changes were made possible by the enormous improvements in computers since the work of Pella and Tomlinson (1969).

Results

Results presented here focus on estimation of four quantities of management interest: $B_./B_{MSY}$, $F_./F_{MSY}$, MSY, and ϕ. We use the notation $B_./B_{MSY}$ to mean the biomass at the end of the final year, relative to B_{MSY}, and $F_./F_{MSY}$ to mean the fishing mortality rate during the final year, relative to F_{MSY}. Box–whisker plots are used to allow visual assessment of accuracy and precision of estimates. For further summarization, results are presented as median proportional unsigned error (MPUE, defined below) in comparable groups of estimates.

Realized values of the status indicators ($B_./B_{MSY}$ and $F_./F_{MSY}$) for each simulated population trajectory are given (Table 1) to aid interpretation of estimation results. The values vary by trajectory, and within trajectory, the true value of $B_./B_{MSY}$ varies by model shape ϕ, whereas the true value of $F_./F_{MSY}$ varies by ϕ and population value of F_{MSY}.

Graphical presentation of results

Similar graphical presentations are used for results in each

Table 1. Underlying values of management quantities in simulated populations following generalized production model.

Simulated trajectory	$B./B_{MSY}$			$F./F_{MSY}$		
	$\phi = 0.35$	$\phi = 0.5$	$\phi = 0.65$	$\phi = 0.35$	$\phi = 0.5$	$\phi = 0.65$
1[a]	0.28	0.50	0.61	3.13–2.42	1.83–1.54	1.44–1.26
2[a]	1.78	1.54	1.42	0.02–0.40	0.13–0.43	0.07–0.31
3[a]	1.72	1.50	1.45	0.001–0.42	0.07–0.47	0.04–0.23
4[b]	0.29	0.50	0.62	2.59–5.70	1.66–3.06	1.35–2.26
5[a]	1.18	0.99	0.91	0.72–0.83	0.90–1.00	0.68–1.04

Note: Range in population F_{MSY} was 0.1–0.8 year^{-1}. Definitions: B_{MSY}, biomass at which maximum sustainable yield can be obtained; F_{MSY}, corresponding instantaneous rate of fishing mortality; $B.$, biomass at end of final year of simulation; $F.$, fishing mortality rate during final year of simulation; ϕ, shape parameter of generalized model expressed as ratio of B_{MSY} to carrying capacity K.

[a]In ranges shown, higher $F./F_{MSY}$ corresponds to population with higher F_{MSY}.

[b]In ranges shown, higher $F./F_{MSY}$ corresponds to population with lower F_{MSY}.

quantity for each trajectory (Figs. 4–13); that presentation is explained in detail here. A set of three panels displays estimation results for a single quantity for a single trajectory (e.g., Figs. 4*a*–4*c* for estimates of $B./B_{MSY}$ from trajectory 1). The top panel (*a* or *d*) of the three shows results for $\phi = 0.65$; the center panel (*b* or *e*), for $\phi = 0.5$; the bottom panel (*c* or *f*), for $\phi = 0.35$. Within each panel, estimates from the NM estimator are in the upper row and estimates from the EA estimator are in the lower row. Within each row are three subpanels with results from the four cases of F_{MSY}. The subpanel labeled "Low" corresponds to $F_{MSY} = 0.1$, "Med" (medium) to aggregated results from $F_{MSY} = 0.2$ and 0.4, and "High" to $F_{MSY} = 0.8$. Within each subpanel, three box–whisker plots (Tukey 1977) correspond to the three levels of simulated observation error noted at the left of each panel. (Results on simulations with 5% observation error are intermediate to those for 2% and 15% and are omitted to save space.) Therefore, each box–whisker plot represents a unique combination of trajectory, model shape (ϕ), quantity estimated, estimator, level of F_{MSY}, and amount of simulated observation error. For clarity, graphics vary by the quantity estimated. In plots of $B./B_{MSY}$ and $F./F_{MSY}$, estimation error is displayed; in plots of ϕ, the estimates themselves are displayed; in plots of MSY, which can vary widely in scale, estimation error is displayed as a percentage of the true value.

Within each box–whisker plot, the solid diamond is the median result from fitting 200 (or for medium F_{MSY}, 400) data sets; the box (open rectangle) marks the quartiles; the whiskers mark the range of most other observations; and open circles mark "outside values", defined as those further from the corresponding quartile than twice the interquartile range, a nonstandard value chosen here to reduce the number of points labeled as outside. Each plot includes a vertical broken line, representing a perfectly accurate estimate. Each plot also contains a number from 0 to 400, indicating the number of successful estimates obtained from that estimator in that specific case. (Those numbers are identical for the four estimated management quantities.) A value of 200 for low or high F_{MSY} or 400 for medium F_{MSY} indicates 100% success in estimation. Estimation failures included three cases: (1) inability to converge on parameter estimates, (2) parameter estimates at the relatively wide constraints set on parameters, or (3) a PRODFIT result specifying the asymptotic model described above.

To increase legibility, some plots do not display the entire range of percent error in MSY estimates; in those plots, a few far-outside values were repositioned close to the highest value shown (100% error in most cases).

Trajectory 1

Estimates of $B./B_{MSY}$ from the NM estimator were generally more accurate and precise than those from the EA estimator (Figs. 4*a*–4*c*), sometimes by a considerable margin. The EA estimator generally showed a patterned variation in median error with observation error, showing least accuracy with $F_{MSY} = 0.1$ (low) and 2% observation error. In contrast, results from NM were quite accurate at all levels of observation error and for all ϕ. Both estimators were usually able to provide estimates, EA slightly more frequently than NM.

Estimates of $F./F_{MSY}$ showed a slightly different pattern (Figs. 4*d*–4*f*). Precision of the EA estimator was similar to that of the NM estimator for most cases, with the notable exception of high F_{MSY} and low ϕ (lower right subpanel of Fig. 4*f*), in which case EA was extremely imprecise. Accuracy also appears similar between estimators, except for simulated populations with $F_{MSY} = 0.1$ (low); in those cases, EA estimates were precise but inaccurate.

Estimates of MSY from both estimators were extraordinarily precise, with EA estimates from $\phi = 0.35$ and $F_{MSY} = 0.8$ (high) being the exception, especially at higher levels of observation error (Figs. 5*a*–5*c*). The NM estimator, and to a lesser degree the EA estimator, exhibited greater precision in estimating MSY as the value of F_{MSY} increased. Accuracy was another question: although NM estimates of MSY were generally accurate, EA tended to overestimate MSY at lower values of F_{MSY}, although the inaccuracy was reduced at higher values of F_{MSY}. On this trajectory, the NM estimator slightly overestimated ϕ when $\phi < 0.5$ and underestimated it when $\phi > 0.5$, a pattern more pronounced at higher levels of observation error. The EA estimates of ϕ were generally more precise than the corresponding NM estimates at low levels of observation error but less precise than NM at higher levels of observation error. Accuracy of EA estimates

of ϕ was lower overall and varied more with level of observation error than did NM estimates of ϕ.

Trajectory 2

As for trajectory 1, estimates of $B./B_{MSY}$ from the NM estimator were generally more accurate, and often more precise, than those from the EA estimator (Figs. 6*a*–6*c*). Indeed, NM estimates were in general well centered on the true values (more accurate), with slight inaccuracy appearing with the highest level of observation error. The precision of the NM estimates appears to increase with increasing values of F_{MSY}. On this trajectory, the EA estimator performed poorly when $F_{MSY} = 0.1$ (low): EA failed to provide estimates in many cases, and its results were inaccurate and imprecise in other cases. Surprisingly, with this value of F_{MSY}, the ability of EA to provide an estimate in our procedure was best with higher levels of observation error; apparently, its tendency to choose the asymptotic model on data with low F_{MSY} was diminished by greater noise content. The importance of noise in arriving at estimates may explain the poor accuracy and precision in such cases.

Estimates of $F./F_{MSY}$ from the EA estimator were of poor accuracy on most combinations of F_{MSY}, ϕ, and observation error, with results strongly patterned with those variables (Figs. 6*d*–6*f*). Estimates of $F./F_{MSY}$ from EA generally were too high, often quite substantially. Estimates from NM generally were well centered on the true values and reasonably precise, with both accuracy and precision apparently varying systematically with ϕ and F_{MSY}.

In estimating MSY on this trajectory, EA performed relatively well when F_{MSY} was medium or high ($F_{MSY} > 0.1$). Estimates were precise, though slightly less accurate than those from NM (Figs. 7*a*–7*c*), particularly under low levels of observation error. In estimating model shape, however, EA was markedly less successful than NM: EA estimates of ϕ were much less accurate than those from NM and often less precise, although the estimator with better precision varied. Each estimator exhibited some patterning of estimates with level of observation error (mild patterning for NM, more marked patterning for EA), suggesting that observation error contributes to inaccuracy in estimating ϕ. The long left-hand tails of some distributions from EA with $F_{MSY} = 0.2$ (aggregated into the medium group in the figure) suggest that the estimator tends toward the asymptotic model in those cases.

Trajectory 3

This trajectory proved a more difficult challenge, particularly for the EA estimator. Interestingly, the two estimators most often failed to provide results on different sets of cases: NM tended to fail with $\phi = 0.65$ and the highest levels of observation error (but for all values of F_{MSY}), whereas EA tended to fail with lower ϕ, lower F_{MSY}, and less observation error (Fig. 8). Overall, estimates of $B./B_{MSY}$ from NM were usually more accurate and precise than those from EA, the exception being at $\phi = 0.65$, higher F_{MSY}, and higher levels of observation error; there, EA performed better (Figs. 8*a*–8*c*).

In estimating $F./F_{MSY}$, NM also performed better in most cases (Figs. 8*d*–8*f*). The exception, again, was the conjunction of high F_{MSY} and high levels of observation error; in many such cases, NM was more accurate but less precise. However, performance of EA in such cases was not particularly good, either. Both methods estimated $F./F_{MSY}$ relatively well with $\phi = 0.5$ and $\phi = 0.35$, although EA was less successful at low values of F_{MSY}.

Similar patterns were found in estimating MSY; NM performed better in most cases (Figs. 9*a*–9*c*) and offered more consistent performance. In estimating ϕ, NM tended toward the logistic, especially with higher levels of observation error, but EA tended to be less precise and usually less accurate as well, with results for a particular combination of ϕ and F_{MSY} also depending strongly on the level of observation error (Figs. 9*d*–9*f*).

Trajectory 4

On this trajectory, the NM estimator outperformed EA in estimating $B./B_{MSY}$ (Figs. 10*a*–10*c*) and $F./F_{MSY}$ (Figs. 10*d*–10*f*) in every scenario. The EA estimator suffered from many estimation failures, severe inaccuracy, and poor precision. In several cases, EA displayed the counter-intuitive pattern of improved accuracy with increasing level of observation error. In contrast, NM estimates were well centered in all scenarios, and estimation precision was high, systematically decreasing as the level of observation error in the data increased.

As one might expect, those patterns also were found in estimates of MSY. The NM estimator performed well generally, whereas EA performed well only when F_{MSY} was high and $\phi \geq 0.5$ (Figs. 11*a*–11*c*). In estimating model shape, NM results were generally well centered about the true values, whereas EA results showed systematic and severe inaccuracies and were often less precise, in some cases exhibiting slight bimodality (Figs. 11*d*–11*f*).

Trajectory 5

In analyzing trajectory 5, the EA estimator more frequently failed to provide estimates than did NM. Comparative performance in estimating $B./B_{MSY}$ and $F./F_{MSY}$ was similar to that on trajectory 4: NM markedly outperformed EA in both accuracy and precision (Fig. 12). The EA estimator offered competitive performance only in a few scenarios with higher values of F_{MSY}.

In estimating MSY on this trajectory, both estimators performed well when $F_{MSY} \geq 0.4$, with NM appearing slightly more accurate. For lower F_{MSY}, NM performed far better than EA, which (as noted) failed in many cases and exhibited poor accuracy in others (Figs. 13*a*–13*c*). In estimating model shape, neither model was outstanding (Figs. 13*d*–13*f*), but NM performed better in most cases. As in previous trajectories, NM showed some lack of accuracy, tending toward the logistic in cases with $\phi = 0.35$ and higher observation errors. However, in most cases, EA results were worse.

Summary results

Relative performance of the two estimators is summarized in a plot of median percent unsigned error (MPUE) of one estimator against the other (Fig. 14). Each point represents MPUE computed on 200 runs of each estimator, all on the same master trajectory, shape of production curve ϕ, value of F_{MSY}, and amount of simulated observation error. Because some runs did not succeed, the corresponding points repre-

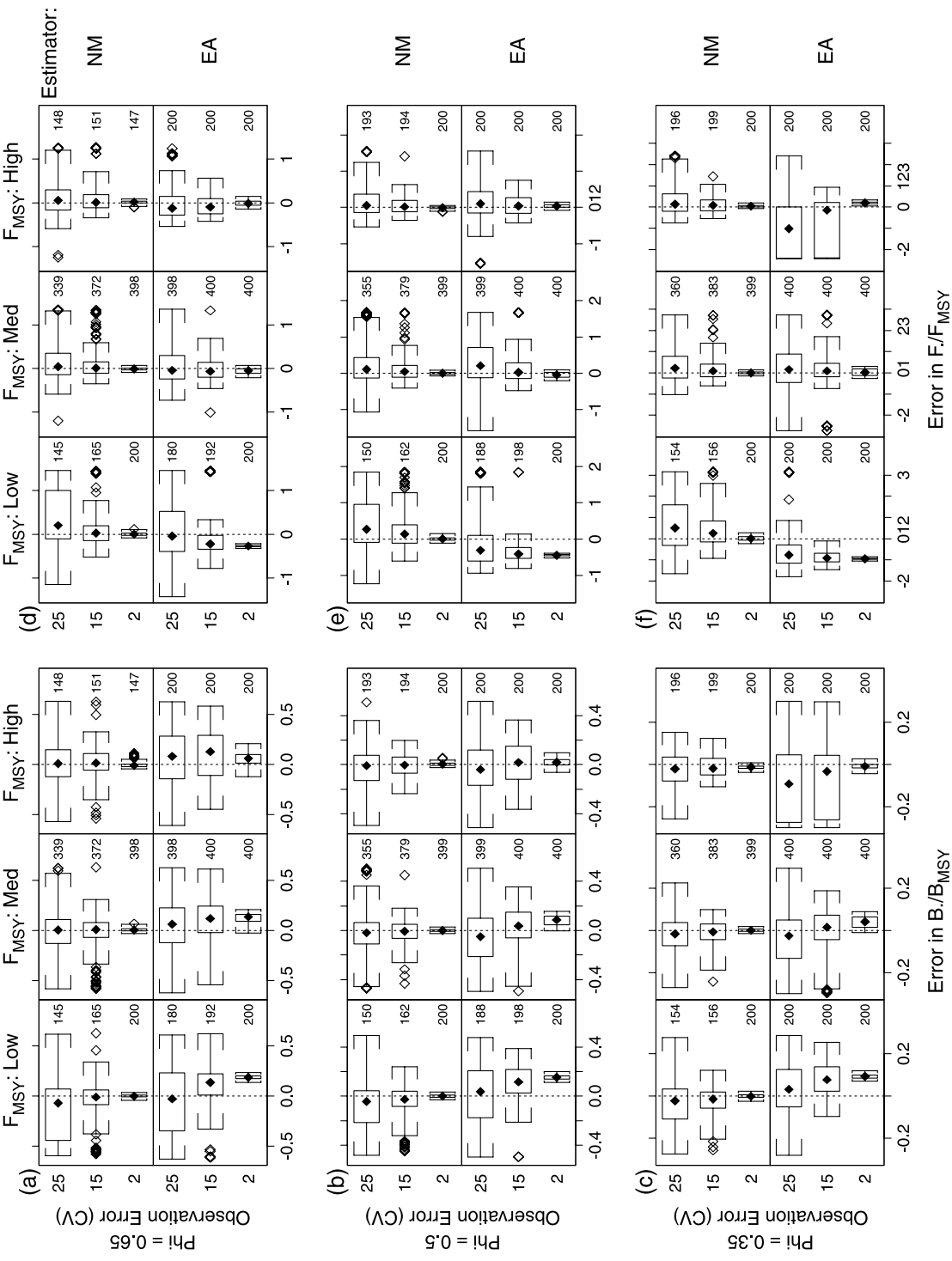


Fig. 4. Results for master trajectory 1. Errors in estimates of final-year relative biomass ($B./B_{MSY}$) and relative fishing mortality rate ($F./F_{MSY}$) from two estimators of generalized production model. See text for detailed explanation of plot and simulations.

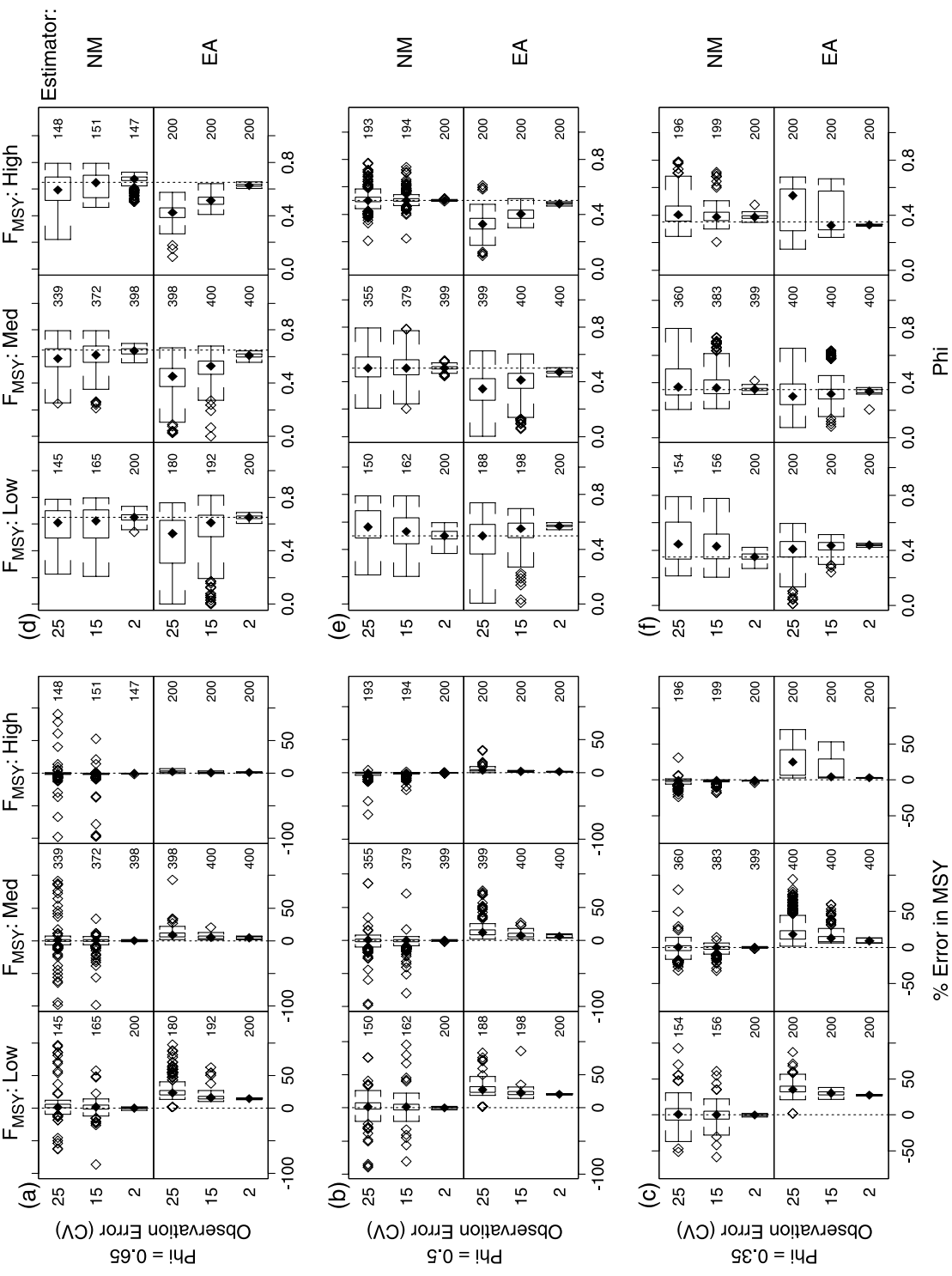


Fig. 5. Results for master trajectory 1. Errors in estimates of maximum sustainable yield (MSY) production-curve shape (ϕ) from two estimators of generalized production model. See text for detailed explanation of plot and simulations.

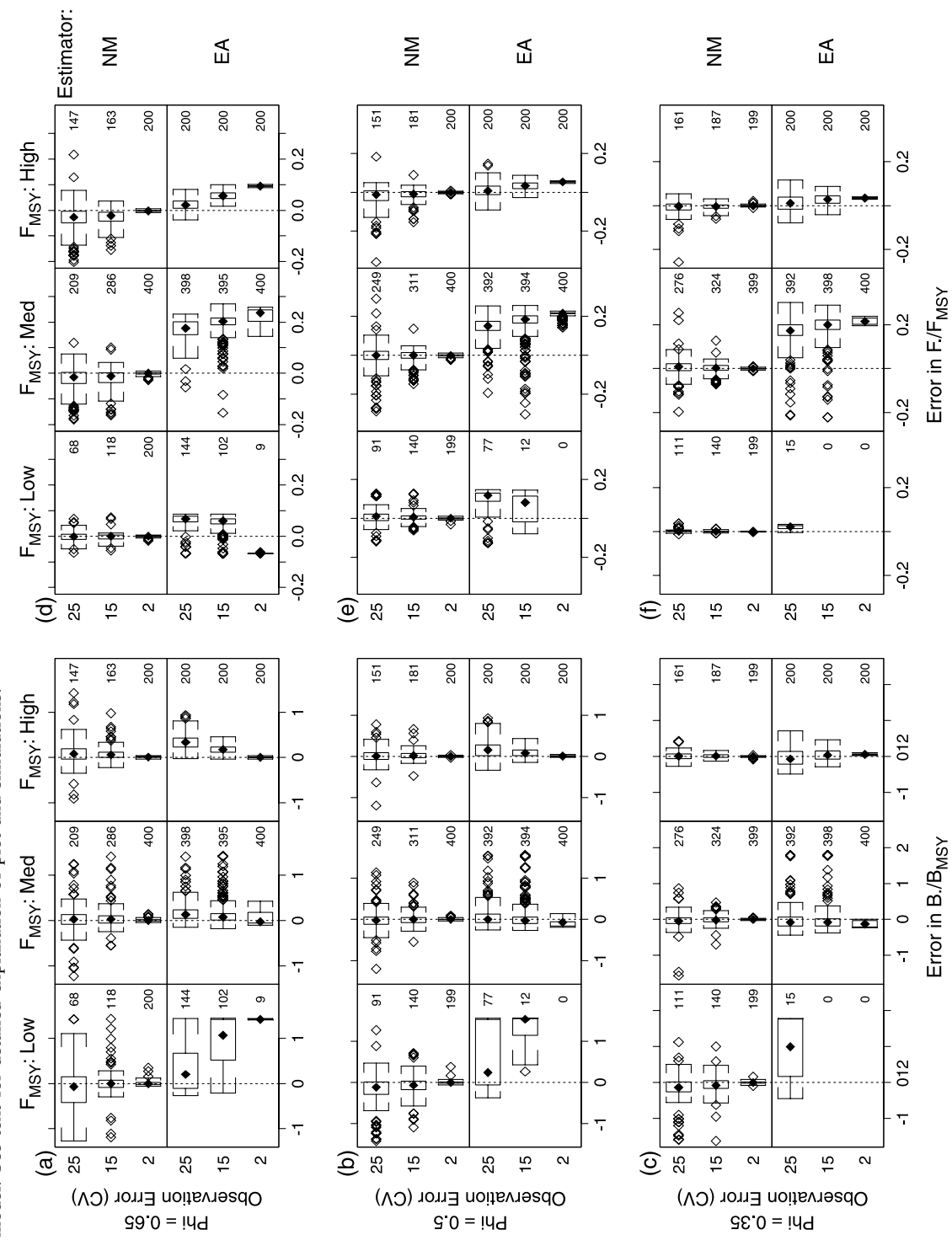


Fig. 6. Results for master trajectory 2. Errors in estimates of final-year relative biomass ($B./B_{MSY}$) and relative fishing mortality rate ($F./F_{MSY}$) from two estimators of generalized production model. See text for detailed explanation of plot and simulations.

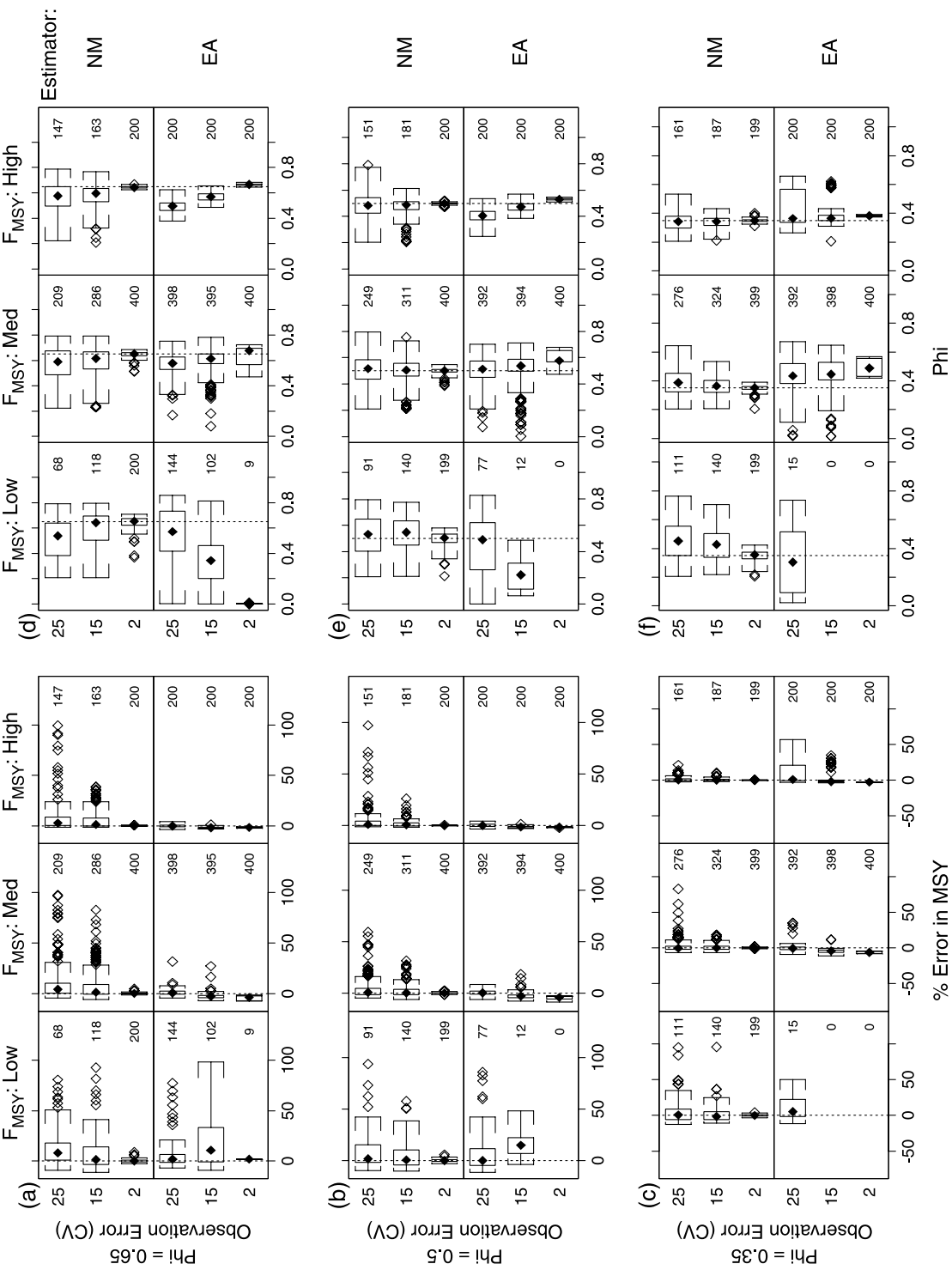


Fig. 7. Results for master trajectory 2. Errors in estimates of maximum sustainable yield (MSY) production-curve shape (ϕ) from two estimators of generalized production model. See text for detailed explanation of plot and simulations.

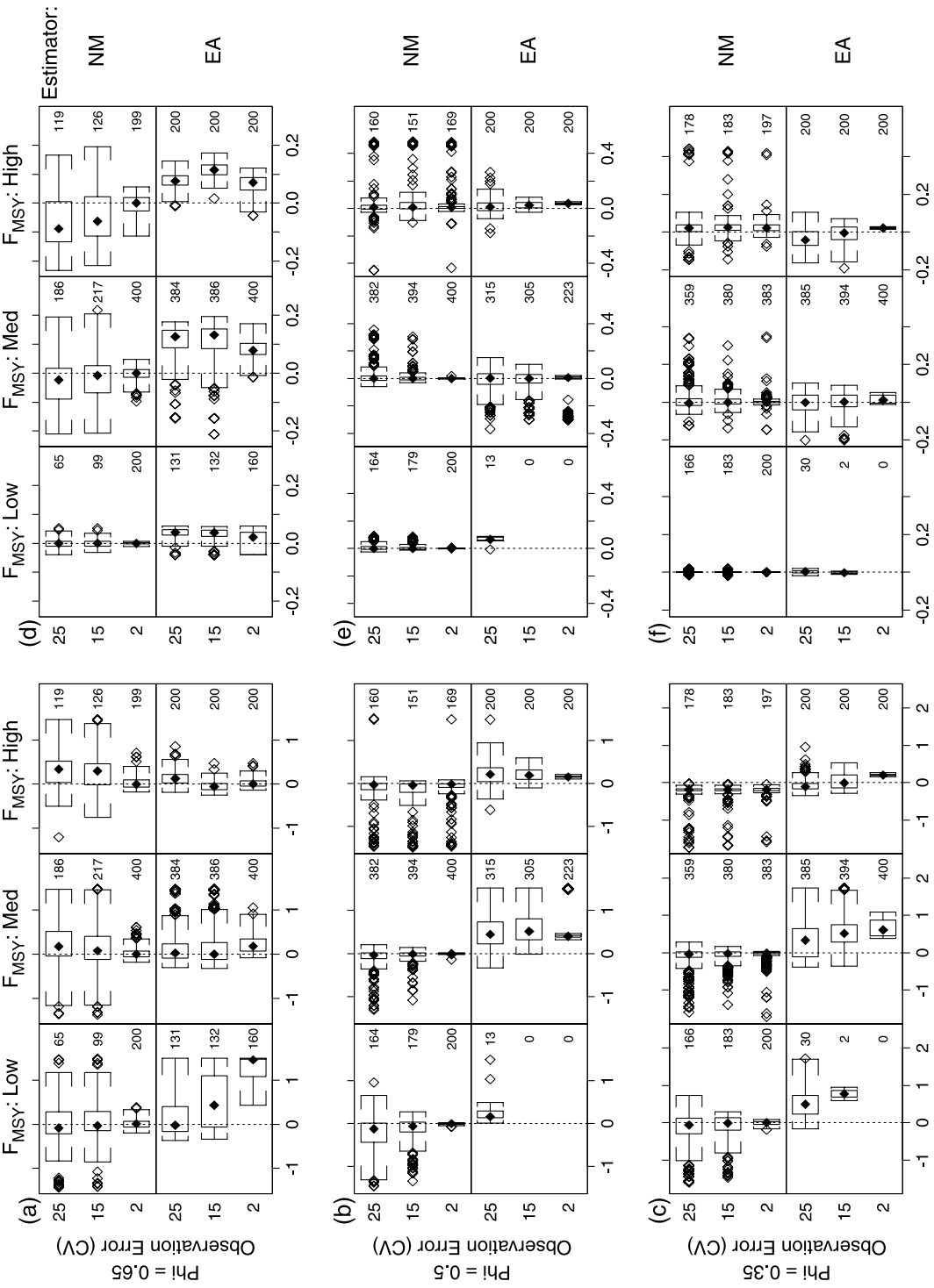


Fig. 8. Results for master trajectory 3. Errors in estimates of final-year relative biomass ($B./B_{MSY}$) and relative fishing mortality rate ($F./F_{MSY}$) from two estimators of generalized production model. See text for detailed explanation of plot and simulations.

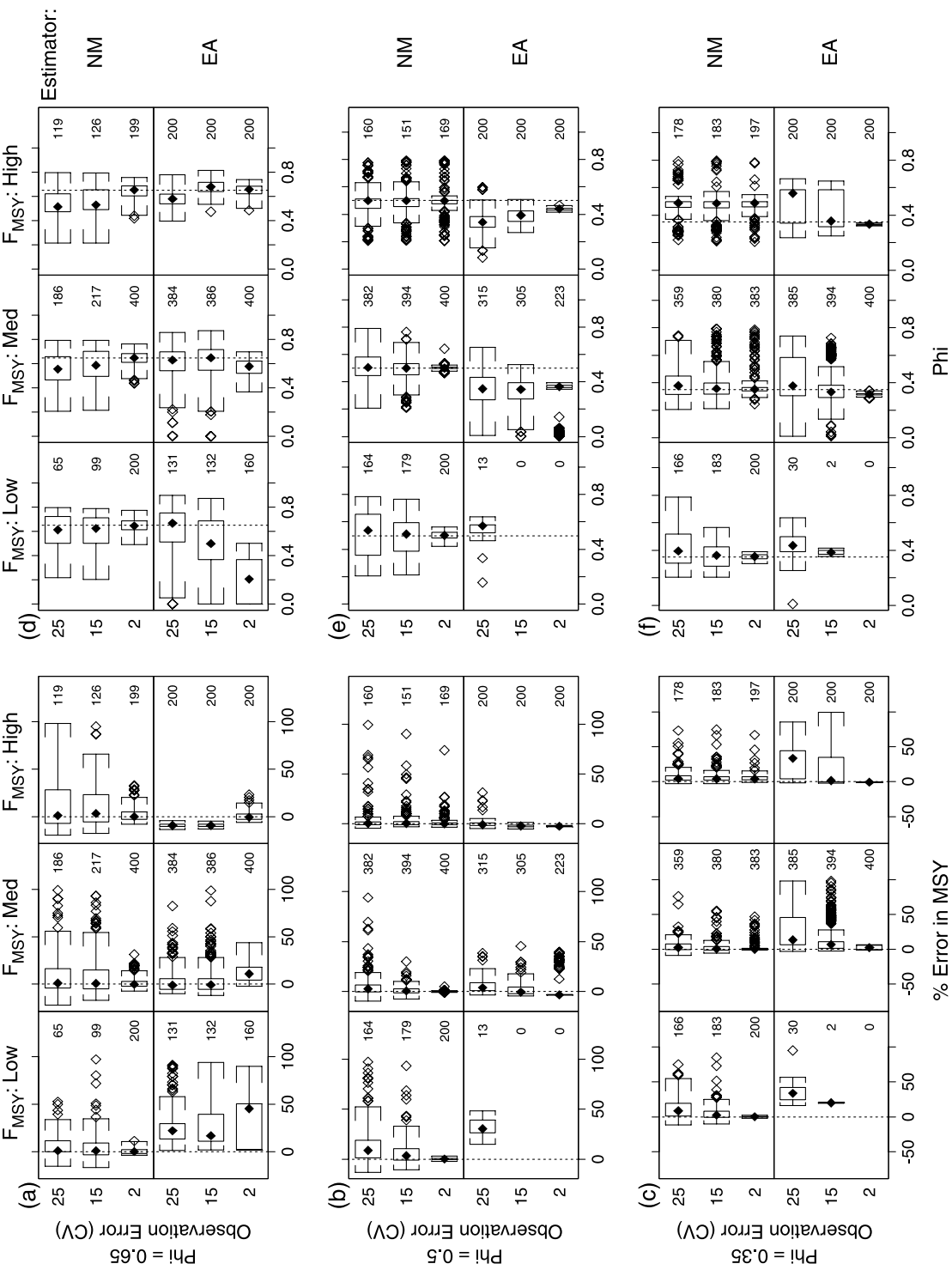


Fig. 9. Results for master trajectory 3. Errors in estimates of maximum sustainable yield (MSY) production-curve shape (ϕ) from two estimators of generalized production model. See text for detailed explanation of plot and simulations.

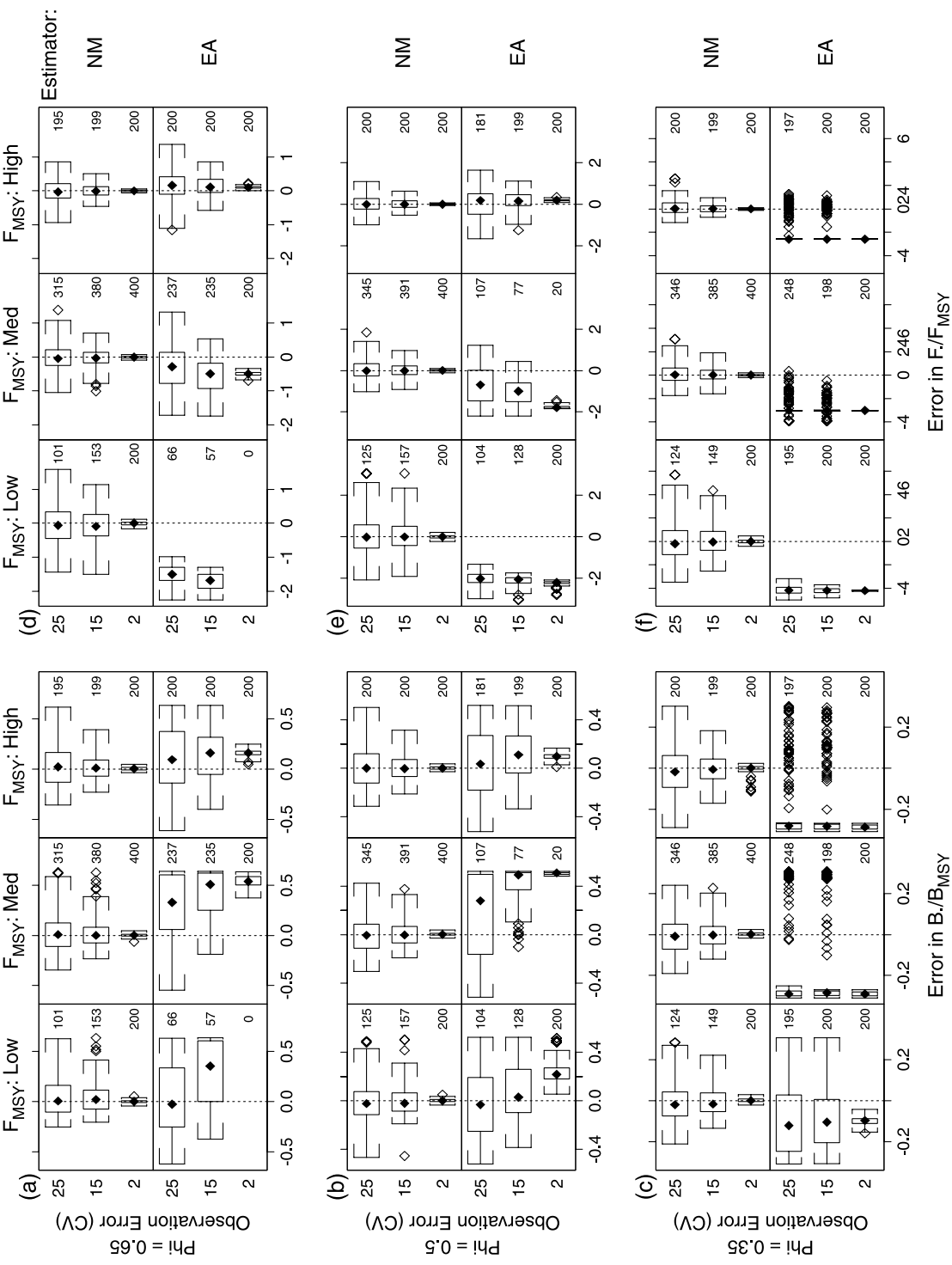


Fig. 10. Results for master trajectory 4. Errors in estimates of final-year relative biomass (B_t/B_{MSY}) and relative fishing mortality rate (F_t/F_{MSY}) from two estimators of generalized production model. See text for detailed explanation of plot and simulations.

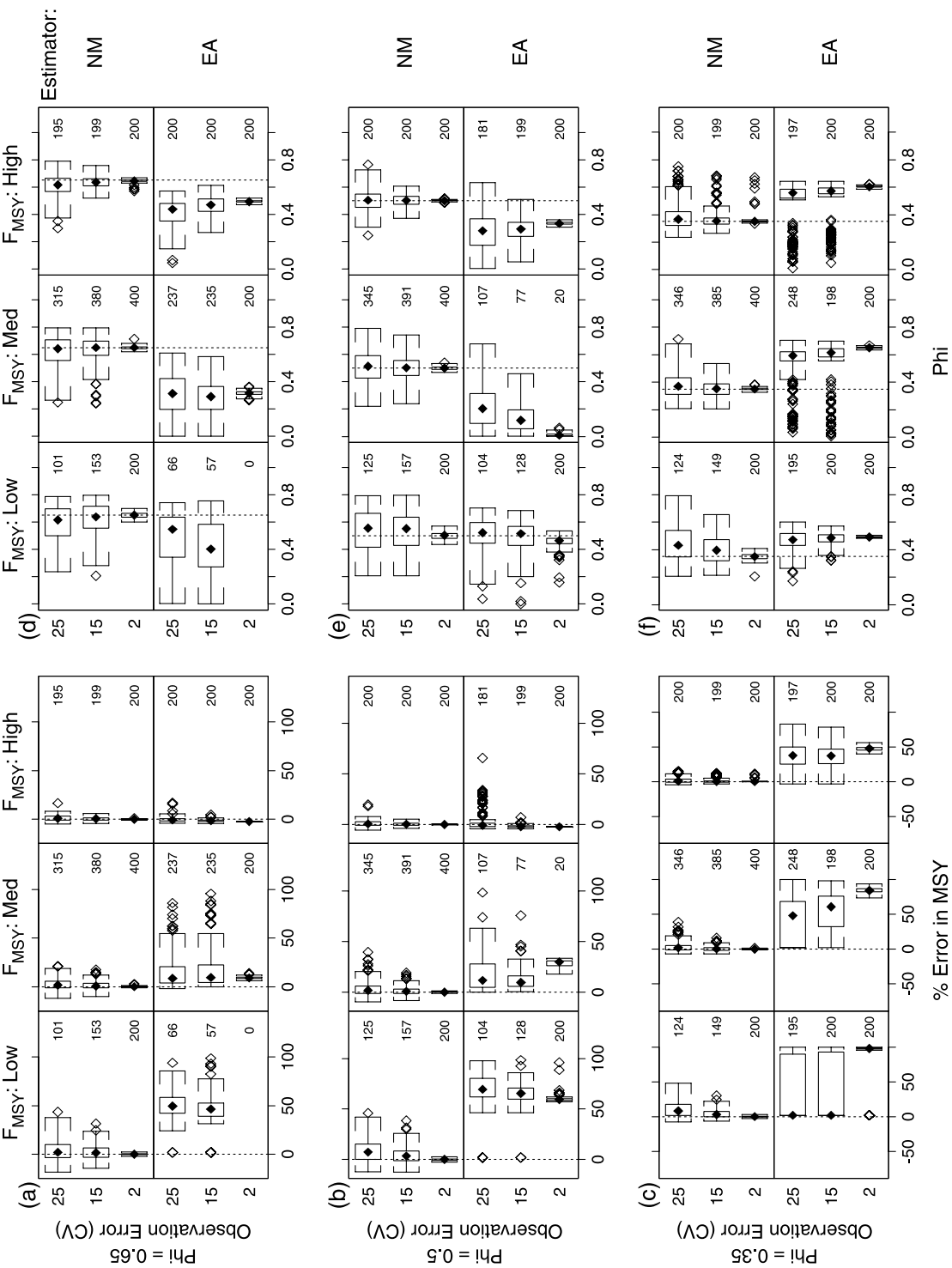


Fig. 11. Results for master trajectory 4. Errors in estimates of maximum sustainable yield (MSY) production-curve shape (ϕ) from two estimators of generalized production model. See text for detailed explanation of plot and simulations.

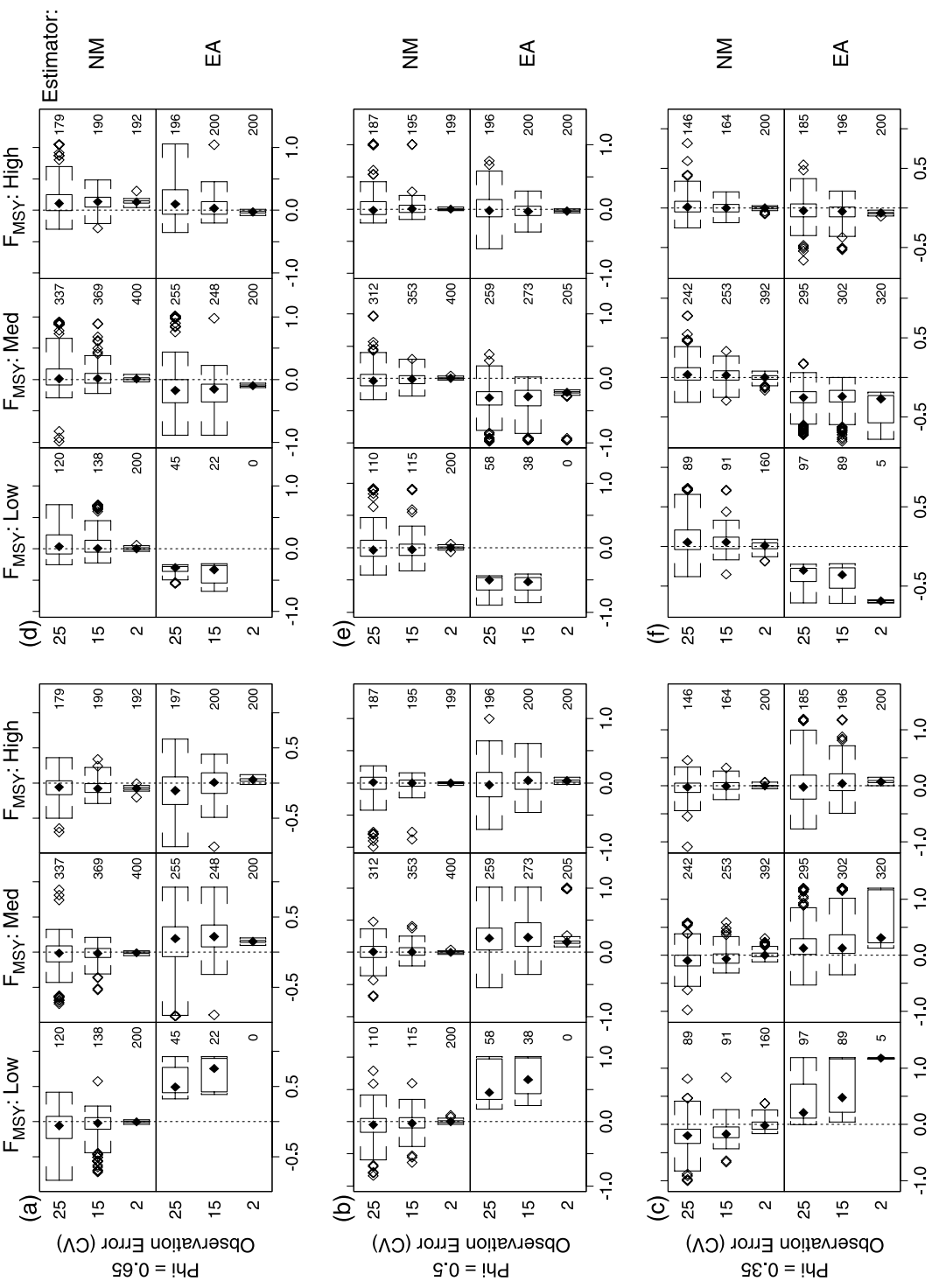


Fig. 12. Results for master trajectory 5. Errors in estimates of final-year relative biomass (B_t/B_{MSY}) and relative fishing mortality rate (F_t/F_{MSY}) from two estimators of generalized production model. See text for detailed explanation of plot and simulations.

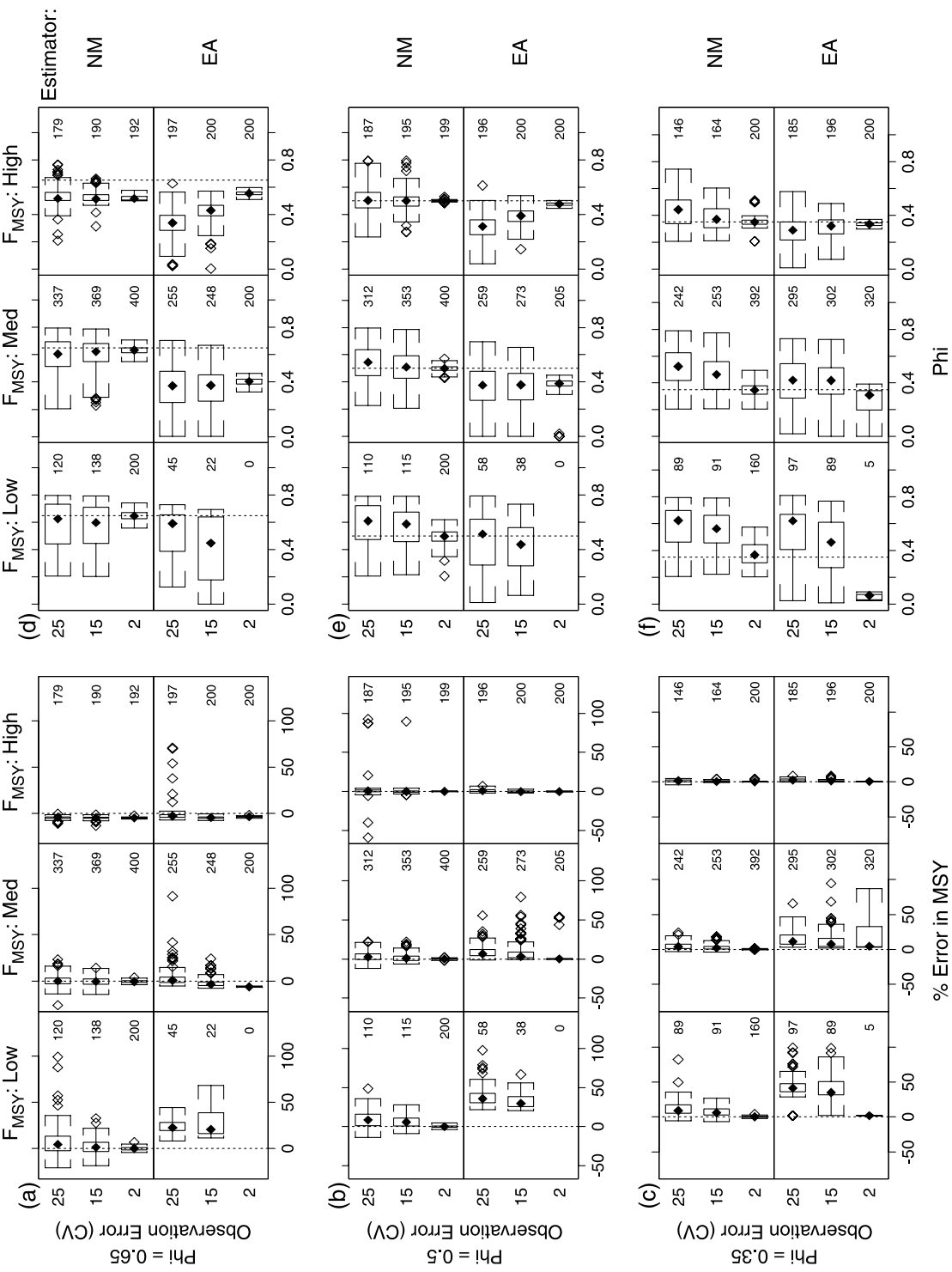
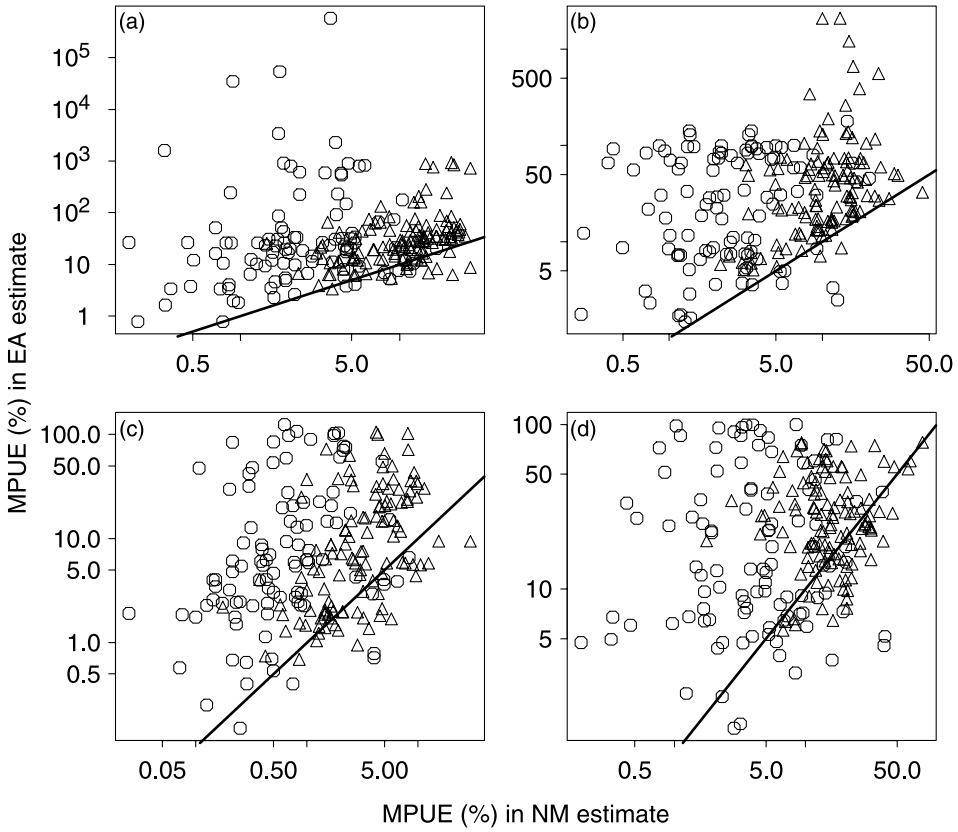


Fig. 13. Results for master trajectory 5. Errors in estimates of maximum sustainable yield (MSY) production-curve shape (ϕ) from two estimators of generalized production model. See text for detailed explanation of plot and simulations.

Fig. 14. Summary comparison of estimators for generalized production model. Each point summarizes 200 pairs of estimates (fewer where estimation failures occurred) with median unsigned percent error statistic (MPUE). *Y* coordinate is MPUE from equilibrium estimator (EA); *X* coordinate, from nonequilibrium estimator (NM); line is $Y = X$. Circles, low observation error (2 or 5%) in simulated data; triangles, high observation error (15 or 25%). Quantities estimated (defined in text) are (*a*) $B./B_{MSY}$, (*b*) $F./F_{MSY}$, (*c*) MSY, (*d*) model shape ϕ. Points above or to the left of the line indicate better performance (lower MPUE) by NM estimator than by EA estimator.



sent fewer than 200 estimates, as per the sample sizes given (Figs. 4–13). We use MPUE as a resistant measure of estimation merit that accounts for both accuracy and precision (Prager 2002). For a set of N estimates of quantity x with true value = x^*, MPUE is defined as

$$\text{MPUE}(x) = \text{median}_{n=1}^{N}\left|\frac{\hat{x}_n - x^*}{x^*}\right| \times 100 \qquad (2)$$

The interpretation of MPUE is that, on similar data, estimation error without respect to sign will be worse than MPUE about half of the time. This examination of results indicates far better performance by the NM estimator overall. In addition to giving lower MPUE than EA in most cases, NM estimates were characterized by low MPUE with low-noise data and MPUE increasing with the noise level in the data (symbol shapes in Fig. 14). In contrast, EA often provided high MPUE even from low-noise data. Superiority of NM was particularly striking in calculation of fishery status ($F./F_{MSY}$); in 231 of 240 cases, the NM estimator offered lower MPUE.

Discussion

In any estimation problem, one expects the information content of the data to influence the precision, and possibly the accuracy, of results. Here, the simulated stocks with higher F_{MSY} are by definition more dynamic than those with lower F_{MSY}. (Under the logistic production model, the definition $r = 2F_{MSY}$, where r is the population's intrinsic rate of increase, illustrates this clearly.) It follows that simulated data sets representing the former should provide more information on response to exploitation than simulated data on the latter and thus should support better estimates. That view seems to be confirmed by the systematic increase of precision from series of equivalent simulated data with increasing values of F_{MSY}, i.e., the increased precision as one moves to the right in most rows of Figs. 4*a*–13*f*.

The overall pattern of each master trajectory provides the core information used in parameter estimation. In this study, trajectories 3 and 4 provided estimates with the worst overall accuracy and precision. That seemed odd to us at first, as each of those trajectories covers a wide range of population size and each crosses $B = B_{MSY}$ twice; they would seem to be information-rich. However, trajectories 1 and 2, which provided better accuracy and precision, also included sustained time periods (roughly years 13–25) of biomass levels near B_{MSY}. We conclude that not only contrast in the data, but also sufficient sampling intensity at each region of population size, is important in parameter estimation for production models.

A referee viewed our study as a case history in nonlinear regression based experimental design. From that perspective,

only certain designs (patterns of exploitation rates and abundances over time) can produce accurate and precise parameter estimates for our nonlinear model, and description of such designs may not have much in common with classification of trajectories as increasing, decreasing, and so on. That is an interesting thought and one explanation for the complexity of our results.

An unexpected pattern occasionally observed in EA estimates was the tendency toward improved accuracy with increasing observation error (e.g., Fig. 4*a*, lower left subpanel). We note that the degree and direction of inaccuracy in EA estimates often seemed to depend systematically on the combination of F_{MSY} and the observation-error level, but beyond that we can say little to explain the observed pattern.

The remark of Sissenwine (1978) that equilibrium estimators can readily produce answers, albeit ill-founded ones, was directed to estimation practices in the 1970s for the logistic production model. If instead the Fox (1975*a*) equilibrium estimator for the generalized production model is used, and if one rejects the asymptotic model (as here), success in estimation is no longer guaranteed. However, our results agree with Sissenwine's impression that the quality of estimates from equilibrium methods can be poor.

Equilibrium methods repeatedly have been found to be biased; the fisheries literature includes recommendations against them. We examined the question again in the first study to compare estimators for the generalized production model, and with the most comprehensive set of simulations to date. Our simulated data sets covered wide ranges of stock productivity, shape of the production curve, and observation error, while also considering a broad range of possible time trajectories of stock biomass. We found it impossible to discern a priori from a data set whether the equilibrium estimator would perform acceptably, and in general, the equilibrium estimator performed poorly in estimating parameters of the generalized production model and, as a special case, parameters of the logistic production model. In contrast, a more modern observation-error estimator, although not perfect, showed better accuracy and precision. Equilibrium methods played an important role when computer power was far less available, but there is no reason to use them today. We conclude that equilibrium methods should be abandoned and abandoned completely. We hope that the scope of our simulations provides sufficient evidence for even former skeptics to adopt that recommendation.

Acknowledgments

We thank D. Vaughan for reviewing the manuscript and K. Shertzer for numerous valuable suggestions for improvement. Supercomputer time and assistance were generously provided by the NOAA Forecast Systems Laboratory in Boulder, Colorado. This work was supported by the U.S. National Marine Fisheries Service, Southeast Fisheries Science Center, through the NOAA Center for Coastal Fisheries and Habitat Research.

References

Fletcher, R.I. 1978. On the restructuring of the Pella–Tomlinson system. U.S. Fish. Bull. **76**: 515–521.

Fox, W.W., Jr. 1970. An exponential yield model for optimizing exploited fish populations. Trans. Am. Fish. Soc. **99**: 80–88.

Fox, W.W., Jr. 1971. Random variability and parameter estimation for the generalized production model. U.S. Fish. Bull. **69**: 569–580.

Fox, W.W., Jr. 1975*a*. Fitting the generalized stock production model by least-squares and equilibrium approximation. U.S. Fish. Bull. **73**: 23–37.

Fox, W.W., Jr. 1975*b*. User's manual for program PRODFIT. Report, NMFS Southwest Fisheries Science Center, La Jolla, Calif.

Graham, M. 1935. Modern theory of exploiting a fishery, and application to North Sea trawling. J. Cons. Int. Explor. Mer, **10**: 264–274.

Gulland, J.A. 1961. Fishing and the stocks of Iceland. Fish. Invest. Ser. II, Mar. Fish. G.B. Minist. Agric. Fish. Food, XXIII(4): 1–52.

Haddon, M. 2001. Modeling and quantitative methods in fisheries. Chapman & Hall, Boca Raton, Fla.

Hilborn, R., and Walters, C.J. 1992. Quantitative fisheries stock assessment. Chapman and Hall, New York.

ICCAT (International Commission for the Conservation of Atlantic Tunas). 2000. Report of the Standing Committee on Research and Statistics (SCRS). ICCAT, Madrid, Spain.

Ludwig, D., Walters, C.J., and Cooke, J. 1988. Comparison of two models and two estimation methods for catch and effort data. Nat. Res. Model. **2**: 457–498.

Methot, R.M. 1989. Synthetic estimates of historical abundance and mortality for northern anchovy. Am. Fish. Soc. Symp. **6**: 66–82.

Mohn, R.K. 1980. Bias and error propagation in logistic production models. Can. J. Fish. Aquat. Sci. **37**: 1276–1283.

Nelder, J.A., and Mead, R. 1965. A simplex method for function minimization. Comp. J. **7**: 308–313.

Pella, J.J. 1967. A study of methods to estimate the Schaefer model parameters with special reference to the yellowfin tuna fishery in the eastern tropical Pacific Ocean. Dissertation, University of Washington, Seattle, Wash.

Pella, J.J., and Tomlinson, P.K. 1969. A generalized stock production model. Inter-Am. Trop. Tuna Comm. Bull. **13**(3): 419–496.

Polacheck, T., Hilborn, R., and Punt, A.E. 1993. Fitting surplus production models: comparing methods and measuring uncertainty. Can. J. Fish. Aquat. Sci. **50**: 2597–2607.

Prager, M.H. 1994. A suite of extensions to a nonequilibrium surplus-production model. U.S. Fish. Bull. **92**: 374–389.

Prager, M.H. 1995. User's manual for ASPIC: a stock-production model incorporating covariates. Version 3.6x. NMFS Southeast Fisheries Science Center, Miami Laboratory Document MIA-2/93-55. 4th ed.

Prager, M.H. 2002. Comparison of logistic and generalized surplus-production models applied to swordfish, *Xiphias gladius*, in the north Atlantic Ocean. Fish. Res. **58**: 41–57.

Quinn, T.J., and Deriso, R.B. 1999. Quantitative fish dynamics. Oxford University Press, New York.

Roff, D.A., and Fairbairn, D.J. 1980. An evaluation of Gulland's method for fitting the Schaefer model. Can. J. Fish. Aquat. Sci. **37**: 1229–1235.

Schaefer, M.B. 1954. Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. Inter-Am. Trop. Tuna Comm. Bull. **1**(2): 27–56.

Schaefer, M.B. 1957. A study of the dynamics of the fishery for yellowfin tuna in the eastern tropical Pacific Ocean. Inter-Am. Trop. Tuna Comm. Bull. **2**: 247–268.

Schnute, J. 1989. The influence of statistical error in stock assessment: illustrations from Schaefer's model. *In* Effects of ocean variability on recruitment and an evaluation of parameters used in stock assessment models. *Edited by* R.J. Beamish and G.A. McFarlane. Can. Spec. Pub. Fish. Aquat. Sci. No. 108. pp. 101–109.

Sissenwine, M.P. 1978. Is MSY an adequate foundation for optimum yield? Fisheries, **3**(6): 22–42.

Tukey, J.W. 1977. Exploratory data analysis. Addison–Wesley, Reading, Mass.

Xiao, Y. 1997. Subtleties in, and practical problems with, the use of production models in fish stock assessment. Fish. Res. **33**: 17–36.

Xiao, Y. 2000. A general theory of fish stock assessment models. Ecol. Model. **128**: 165–180.

Zhang, C.I., Gunderson, D.R., and Sullivan, P.J. 1991. Using data on biomass and fishing mortality in stock production modelling of flatfish. Neth. J. Sea Res. **27**: 459–467.