Effects of Mixed-Metric Data on Production Model Estimation: Simulation Study of a Blue-Marlin-Like Stock

MICHAEL H. PRAGER*

*Population Dynamics Team, Center for Coastal Fisheries and Habitat Research,
National Oceanic and Atmospheric Administration,
101 Pivers Island Road, Beaufort, North Carolina 28516, USA*

C. PHILLIP GOODYEAR

*415 Ridgewood Road
Key Biscayne, Florida 33149, USA*

Abstract.—An underlying assumption of surplus-production modeling is that the model's two basic data streams (indices of abundance and records of catch) are expressed in the same metric, either biomass or numbers. For lack of data, that assumption is sometimes violated; recent assessments of blue marlin *Makaira nigricans* and white marlin *Tetrapturus albidus* in the Atlantic Ocean, for example, have used indices of abundance based on numbers with indices of catch based on biomass. We examined the effects of using mixed-metric data in production modeling. Our method was a simulation study based on the life history and fishery characteristics of blue marlin in the Atlantic Ocean. Populations were simulated over a range of growth patterns and with either increasing or declining abundance by the use of a simulation model incorporating sex, size, and age structure; sexually dimorphic growth; variation of size at age; age-varying natural mortality; and deterministic or stochastic recruitment. Simulated abundance data were aggregated across ages and combined with random errors to represent the data sets used in assessments. A logistic surplus-production model was then fitted to those simulated data sets. The resulting estimates of maximum sustainable yield and stock status were surprisingly robust to the use of mixed-metric data. Estimates from consistent data were generally more precise but not necessarily less biased. We thus conclude that this use of mixed-metric data is acceptable for blue marlin and similar species. However, errors in estimation varied strongly by the growth pattern assumed, which indicates that better knowledge of growth patterns would allow future assessments to better define likely biases arising from the use of mixed-metric data.

Obtaining detailed data on widely ranging marine species can be extraordinarily difficult. The difficulty arises both in the political realm (because international cooperation among countries is needed) and in the biological realm (because sampling throughout such a species' range is expensive and challenging). Thus, application of any model to data on such species is likely to involve compromises.

The example considered in this paper is the application of a surplus-production model to blue marlin *Makaira nigricans,* a species whose management in the Atlantic Ocean is coordinated by the International Commission for the Conservation of Atlantic Tunas (ICCAT). In ICCAT assessments of blue marlin, the model of choice has been the surplus-production model. Among the reasons for choosing a production model are the facts that only age-aggregated data are available and ICCAT's

management philosophy involves maximum sustainable yield (MSY).

The theory of production modeling (Lotka 1924; Schaefer 1954, 1957; Pella and Tomlinson 1969; Fox 1970; Schnute 1977; Prager 1994) allows expressing stock size and yield in terms of either biomass or numbers but assumes that a consistent metric is used. In modeling fish populations, the choice has usually been biomass (thus MSY is usually estimated in units of biomass per year). However, in recent assessments of blue and white marlin *Tetrapturus albidus* by ICCAT species groups, catch data have been available in terms of biomass while the abundance indices from important recreational fisheries have necessarily been based on catch per unit effort (CPUE) in numbers (ICCAT 1999). Conversion has not been possible because annual average weights are not available from longline fisheries, which account for over 80% of landings. Thus, mixed-metric data have been used in assessments. It is unknown to what degree such usage may bias estimates of MSY and related quantities.

* Corresponding author: Mike.Prager@noaa.gov

Theoretical considerations predict that bias should arise from such use of mixed-metric data. In an age-structured population, variation in MSY comes both from changes in yield per recruit and from changes in recruitment per spawner as a population is fished down (Goodyear 1996). In a stock of long-lived fish, increased fishing typically reduces yield per recruit by reducing survival to older ages and thus reducing the mean weight of a fish in the catch. An abundance index based solely on the number of individuals cannot capture this effect, and consequently it may underrepresent the reduction in population biomass from increased fishing. It seems, then, that an abundance index based on numbers per unit effort would be at best an imperfect substitute for the biomass-based index that is assumed when fitting a production model in biomass metric.

Given the historical data available on blue marlin and other billfishes in the Atlantic Ocean, we expect that representative indices of abundance in biomass metric will remain unavailable for most if not all major fisheries. We further expect surplus-production models to remain in use for assessment of these stocks, because in general only age-aggregated data are available. The continuing use of a production model with mixed-metric data has prompted this study of the effects of the inconsistency. The problem is complex enough that a general solution does not seem possible; nor is a purely analytical approach likely to yield conclusive results. Here, we explore the problem through simulation of the blue marlin stock and the assessment process.

Methods

The study procedure comprised three steps: (1) repeated simulation of a population similar to blue marlin in the Atlantic Ocean, (2) fitting a logistic surplus-production model to the age-aggregated data extracted from the simulations, and (3) comparison of the production-model estimates with the known true values from the simulations. Because the first step attempted to mimic a specific stock, it required the most choices specific to this study. The second step was done with standard model-fitting methods and the third with graphical analysis tools.

Biological Characteristics of Simulations

The simulations were performed with an age-structured population simulation model (Goodyear 1989) with specific parameter values similar to those of the stock of blue marlin in the Atlantic Ocean. Growth, mortality, recruitment, and fishing were all simulated with age and size structure. We used monthly time steps, 100 ages (to avoid any possible truncation of the age structure), two sexes, and variability in size at age. Estimation of MSY from the simulation model was done by the method of Goodyear (1996), which through a nonlinear search on values of fully recruited fishing mortality (F) determines MSY and the fishing mortality rate at MSY (F_{MSY}) for a given vector of fishing selectivities.

Growth.—The wide variation in estimated growth rates in the literature illustrates that the mean sizes at age and asymptotic maximum size (L_∞) of blue marlin are poorly known. Because different growth patterns cause different responses in mean weight as fishing mortality changes, representative simulation of growth was a major concern in the present analysis. Several authors have fit von Bertalanffy (1938) growth equations to data on blue marlin (Wilson 1984; Cyr 1987; Boggs 1989; Prince et al. 1991; Wilson et al. 1991). In each case, males were smaller than females at the same age. In the studies cited, the estimated von Bertalanffy growth coefficient (K) tended to be larger for females than for males, but the estimated values of K for the two sexes overlapped among studies.

For the present investigation, we chose five growth treatments, characterized by different values of K, labeled K_1 through K_5. Values of K were derived from probability points on the mean of log-transformed estimates of K (pooled separately for males and females) from the five sources cited. The values thus selected (and corresponding probability points) were as follows: $K_1 = 0.0985$ ($p = 0.05$); $K_2 = 0.1426$ ($p = 0.25$); $K_3 = 0.1845$ ($p = 0.50$); $K_4 = 0.2385$ ($p = 0.75$); and $K_5 = 0.3453$ ($p = 0.95$). Following Wilson (1984), the asymptotic maximum sizes (length from lower jaw to fork, cm) for all growth treatments were 207.9 for males and 282.6 for females. In all cases, we set parameter $t_0 = 0$. The resulting growth curves (Figure 1) are intended to bracket plausible true growth patterns for this species. Simulated growth was distinct for each of 101 growth morphs within each cohort. Growth characteristics of morphs were such that size at age by sex exhibited a coefficient of variation (CV = 100·SD/mean) of 12% in the absence of fishing. The length–weight relations used were those of Prager et al. (1996).

Fecundity.—Reproduction in Atlantic blue marlin has been studied by Erdman (1968), Cyr (1987), and De Sylva and Breder (1997). By one
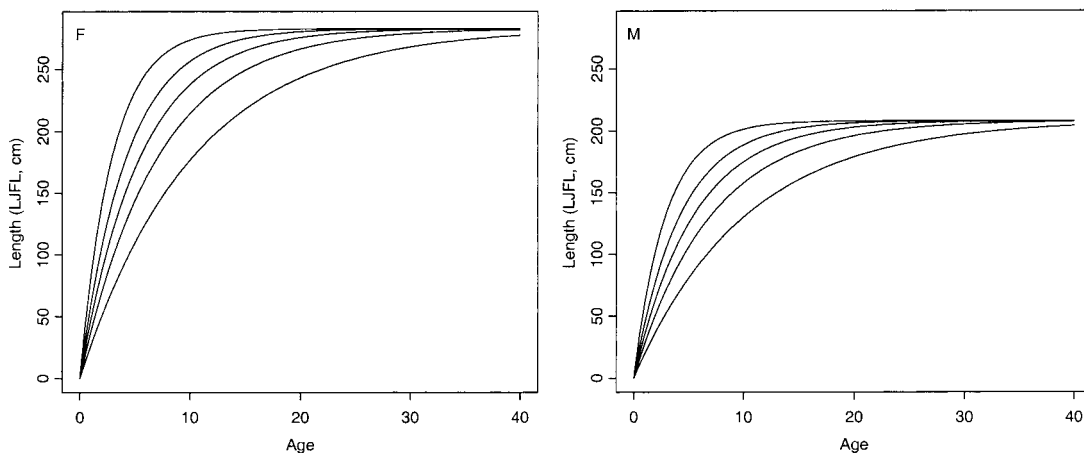
FIGURE 1.—Growth curves used in simulating blue marlin populations. Female (left) and male (right) sizes at age (length from lower jaw to fork, cm) are shown for growth treatments K_1 (lowest curve) through K_5 (highest curve). Curves are truncated at 40 years to show detail.

report, female blue marlin reach maturity at 120 kg (perhaps less) and are capable of spawning up to four times per reproductive year, which in the North Atlantic includes July and August (De Sylva and Breder 1997). Erdman (1968) also concluded that the spawning season was July and August but reported female minimum size at spawning of about 45 kg.

We concluded that the data available from the literature were too sparse to characterize a detailed model of blue marlin fecundity, and we thus modeled individual fecundity as a simple linear multiple of body weight for females above 45 kg; smaller fish were assumed to be immature. This fecundity model may overestimate the reproductive importance of newly maturing females, but, lacking an estimated maturity ogive, this simple model seemed the best choice. We assumed that abundance of males never limits reproduction.

Natural mortality rate.—The instantaneous rate of natural mortality, *M,* is notoriously difficult to estimate and especially difficult to separate from the total mortality rate, *Z*. A few data sets are available to estimate *Z* from the oldest fish in a sample, using the method of Hoenig (1983) as corrected for sample size. The oldest blue marlin reported by Hill et al. (1989) from a sample of 48 females in the Pacific was 27 years old. Assuming age of full recruitment to the sampling protocol of 2 years, the resulting estimate is $Z = 0.18$/year. In a separate study, Wilson et al. (1991) found the oldest blue marlin in a Pacific sample of 18 females to be 21 years old; the resulting estimate is $Z = 0.19$/year. As some part of total mortality is from

fishing, corresponding estimates of *M* must be lower than *Z*. Assuming a minimal fishing mortality rate of $F \sim 0.03$/year implies 0.15/year as a likely upper bound on *M*.

Wilson et al. (1991) also sampled blue marlin in the Atlantic Ocean; the oldest of a sample of 104 females was 17 years old, giving the estimate $Z = 0.36$/year. This estimate can be coupled with surplus-production model estimates of *F* from the last ICCAT assessment of blue marlin, which averaged 0.31/year for years preceding the 1981–1984 data collections of Wilson et al. (1991); this gives the estimate $M = 0.36 - 0.31 = 0.05$/year. However, estimates of absolute *F* from surplus-production models are often highly imprecise (Prager 1994), and based on biological considerations, we believe this one to be unrealistically high.

For simulation purposes, we assumed that *M* declines continuously from an initial value of 0.5/year for the youngest recruits to 0.10/year for the oldest fish in the population (Figure 2). The assumption of a relatively low adult mortality rate of 0.10/year was chosen in part because of the relationship between *M* and the mean weight of an individual in the stock. A relatively low *M* increases the number of age-classes in the unfished population and thus causes heavier unfished mean weight. As a consequence, fishing mortality tends to have a greater proportional effect on mean weight than with a higher value of *M*. In this study, the relatively low value of *M* should accentuate the differences between abundance indices based on weight and those based on numbers.

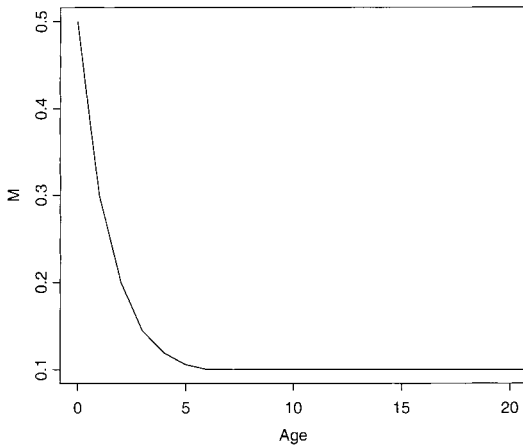*Recruitment*.—Recruitment was defined in this

FIGURE 2.—Instantaneous rates of natural mortality (M) used in the simulation of a blue-marlin-like population. For ages older than 20 years, $M = 0.1$/year.
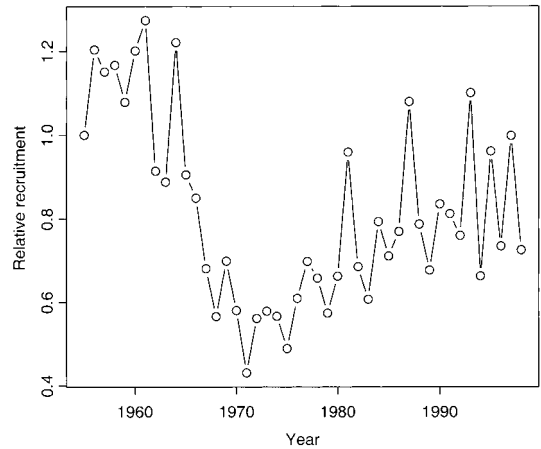

FIGURE 3.—Recruitment relative to that in the first simulation year for a typical simulation of a blue-marlin-like population, illustrating the stochasticity of the recruitment function.

study as the number of fish attaining 3 months of age. Annual recruitment in simulated populations was determined from population fecundity at the beginning of each year using a Beverton–Holt stock–recruitment function. The shape of the stock–recruitment curve helps determine the population's response to fishing. In the present context, the main effect is an increase in recruitment per spawner as fishing mortality increases from zero. For any particular fishing mortality rate, the equilibrium mean weight of a fish in an ideal age-structured population is a joint function of growth and mortality, but it is not influenced by recruitment. Hence, at least under equilibrium conditions, the stock–recruitment curve determines only average recruitment and does not affect the relationship between CPUE in numbers and CPUE in weight. Consequently, we limited the analysis to a single Beverton–Holt stock–recruitment curve whose slope at the origin was 10 recruits per spawner. Absolute recruitment levels were not a concern here; mean recruitment at MSY was set at 1,000,000 fish at 3 months of age.

In nature, recruitment exhibits stochastic annual variability due to environmental factors unrelated to adult stock size. We acknowledged this characteristic by adding a lognormal random component to mimic natural variability in egg-to-recruitment survival. Specifically, in some simulations the predicted (mean) recruitment from the stock–recruitment relationship was multiplied by $\epsilon = \exp(\nu \cdot \text{CV} - [\text{CV}^2/2])$, where ν is a random normal deviate with mean 0 and variance 1 and CV is the coefficient of variation of the logarithm

of the random multiplier ϵ, which was set at 0.15 for this analysis. An example of the magnitude of this variation taken from one of the simulations for growth treatment K_1 is illustrated in Figure 3. To isolate possible estimation bias or variability caused by recruitment stochasticity, we performed separate analyses of estimation results from simulated data with deterministic recruitment and from simulated data with stochastic recruitment.

Other Aspects of Simulations

Initial conditions.—The initial year of each simulation was designated as 1955 and began with the population at its unfished stable age distribution, a property determined by M and the stock–recruitment curve. The population was then simulated through 1998, for a total of 44 years.

Fishing mortality rates and population trajectory.—The fishing mortality rate (F) was adjusted annually for each growth treatment so that the resulting pattern of relative biomass over time closely approximated the pattern estimated in a recent assessment of blue marlin (ICCAT 1999); this is a pattern of increasing depletion over time. Because the results might be sensitive to the particular population trajectory used, we also constructed simulations in which annual F was adjusted to provide a set of alternative trajectories simulating a population recovering from overfishing (Figure 4). In each of those simulations, the stock was initially set at the stable age distribution associated with the high fishing mortality rate responsible for its overfished state.
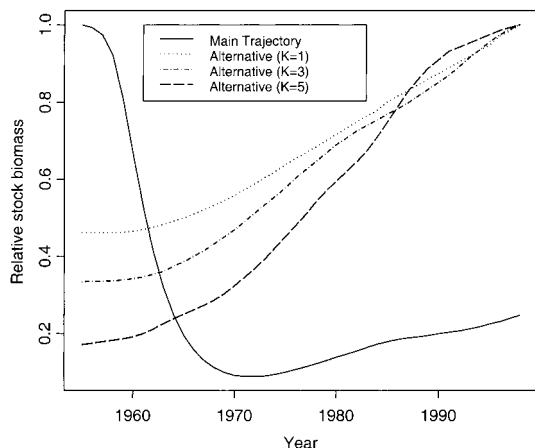
FIGURE 4.—Main and alternative biomass trajectories approximated by simulated blue-marlin-like populations. The main trajectory resembles estimates for this species in ICCAT (1999); the three alternative (hypothetical) trajectories that are shown vary by growth treatment and were devised to examine the generality of results. Here, each trajectory is standardized to its maximum; simulated observation error was added to all trajectories before production modeling.

Fishing selectivities were the same across all years and simulations. Fish first recruited to the fishery at age 1 and at all ages older than age 0 were equally susceptible to fishing. Simulated annual catch data were saved for production modeling.

Abundance indices.—Two abundance indices were constructed from each simulation, one in units of numbers of fish and the second in units of biomass. The first was derived by multiplying an arbitrary catchability coefficient ($q = 10^{-6}$) by the average number of individuals alive at ages 1 and older each year. The second was derived by multiplying q by the average total biomass of ages 1 and older (averaging was within each year). Resulting values (e.g., Figure 5) were saved for use in production modeling. We refer to these indices throughout as indices of NPUE (numbers per unit effort) and WPUE (weight per unit effort), respectively. When an abundance index without regard to metric is considered, we refer to CPUE (catch per unit effort). Each input data set for production modeling contained annual simulated data on catch in weight and either NPUE or WPUE.

Observation error.—Simulated observation error (measurement error) was added to catch and CPUE data before production modeling. Separate simulated data sets were made with four levels of error, at CVs of 0.01, 0.05, 0.125, and 0.25. In all
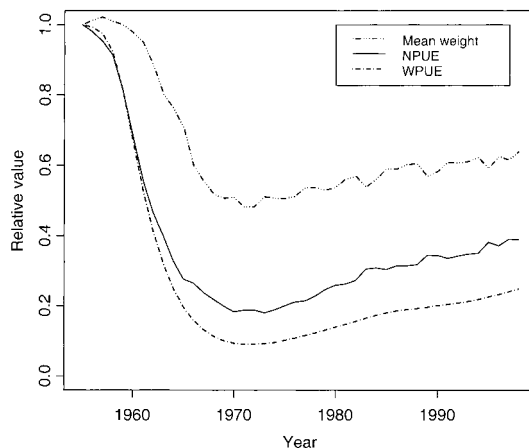

FIGURE 5.—Comparison of relative-abundance indices based on numbers (NPUE) and biomass (WPUE), shown with the corresponding change in mean weight of a fish (all from same simulation in Figure 3). Each quantity is standardized to its value in the first simulation year, designated as 1955.

cases, error was constrained to remain within 2 SD, a procedure that reduces the realized CV to about 88% of its original value. The same level of error was applied to both catch and CPUE in a given data set, but the error in catch was normally distributed while the error in CPUE was lognormally distributed.

Extent of Simulations and Production Modeling

The number of simulations and production-model fits was large. For the main (declining) population trajectory, five growth treatments and the choice of stochastic or deterministic first-year survival resulted in 10 combinations; for each, 100 simulations were performed, resulting in 1,000 simulations. In addition, four levels of measurement error and two types of CPUE (NPUE or WPUE) were modeled, giving 8,000 different simulated data sets for production modeling. For the alternative (increasing) population trajectories, only simulations with stochastic (and not deterministic) recruitment were made; thus 4,000 data sets were generated. For both trajectories, then, a total of 12,000 simulated data sets were fitted by the production model.

Production modeling was accomplished with the ASPIC software of Prager (1995). That program implements an analytical solution of the logistic production model (Pella 1967; Prager 1994), fitted with an observation-error estimator and assuming multiplicative errors in CPUE. We fitted the models conditioned on observed (simulated) yields.

Evaluation of Estimates

To evaluate the bias and precision of estimation, we compared production-model estimates of three quantities of management interest to their known values from the simulations. The three quantities were maximum sustainable yield (MSY), final-year fishing mortality rate relative to fishing mortality rate at MSY (denoted $F./F_{MSY}$), and final-year biomass relative to biomass at MSY (denoted $B./B_{MSY}$). Estimation quality was represented by percent error (PE) in each estimate. If the known value of an estimated quantity is x and its estimate is $\hat{x}$, the percent error in the estimate is computed as $PE(x) = 100\% \cdot (\hat{x} - x)/x$. Thus, a PE of zero indicates an unbiased estimate. Results were analyzed graphically through box-and-whisker diagrams (Tukey 1977) using S-Plus software (Math-soft 1999).

Results

Properties of Simulated Populations

In each simulated data set, known values of the three management quantities depended on the growth treatment. For the main trajectory and growth treatments K_1 through K_5, known values of MSY were approximately 3,893, 6,629, 8,765, 11,450, and 15,860 metric tons /year, respectively. The corresponding known values of final-year relative biomass $B./B_{MSY}$ in simulations with deterministic recruitment were 0.81, 1.30, 0.68, 0.43, and 0.25, respectively. Values of this quantity for simulations with stochastic recruitment varied, but they were distributed tightly about those values (Figure 6a). The corresponding known values of final-year relative fishing mortality rate $F./F_{MSY}$ were 0.75, 0.53, 0.85, 1.87, and 2.44.

The known values of MSY from the alternative trajectories were the same as those from the main trajectory because MSY is determined by biology, not fishing history. The corresponding mean known values of $B./B_{MSY}$ ranged from 1.37 to 1.85 (Figure 6b), and the corresponding known values of $F./F_{MSY}$ ranged from 0.25 to 0.57. Considering both trajectories together, the realized ranges in stock status measures represented a wide variety of final-year stock conditions in which some stocks were overexploited and some were not; also, some were being fished at unsustainable rates and some were not.

Effects of Observation Error

Observation error had a minor to moderate effect on the degree of relative bias experienced (Figure 7). These effects were often more pronounced when NPUE, rather than WPUE, was used in modeling. Differences in the median relative bias between using NPUE and WPUE were relatively constant across levels of observation error in the increasing trajectory and, when estimating $F./F_{MSY}$, in the main (declining) trajectory. The median bias of estimates from the main trajectory did not depend appreciably on whether simulated recruitment was stochastic or deterministic (bottom two rows of Figure 7).

The level and form of observation error in real fishery data are unknown. For that reason, and to control the number of figures required, results from all levels of measurement error were pooled within each combination of growth treatment, CPUE type, and recruitment type for further presentation. Because two of our four levels of observation error were quite small (CV = 0.01 and 0.05), pooling may underrepresent the effects of observation errors (e.g., in estimating $B./B_{MSY}$ for the main trajectory). Nonetheless, figures (not shown) based on just the two higher levels of observation error displayed patterns quite similar to those displayed for all levels of error in Figures 8–13; we thus believe that the results presented here are representative.

Presentation Format of Results

The same graphical format is used to display the distribution of PE in the estimates of each benchmark (Figures 8–13); that format is explained in detail here. A pair of panels is used to display the error in estimates of a single benchmark (e.g., Figure 8 for PE in MSY): within this pair, simulations with deterministic recruitment are on the left and simulations with stochastic recruitment are on the right. All results from the alternative trajectories reflect stochastic recruitment (Figures 11–13).

Each of the two panels is divided horizontally into two subpanels, with the results of fitting to NPUE on the left and those of fitting to WPUE on the right. The results for growth treatments K_1 through K_5 are arranged vertically; one box–whisker plot represents each combination of growth treatment and CPUE type. Within each box–whisker plot, the filled circle is the median PE from 400 estimates. The rectangle of each box–whisker plot encloses the interquartile range, the outer whiskers denote the range of most other observations, and the open circles are ''outside values'' by the criteria of Tukey (1977:44). Tukey's definition of outside is based on the data's empirical distribution; for normally distributed data, it
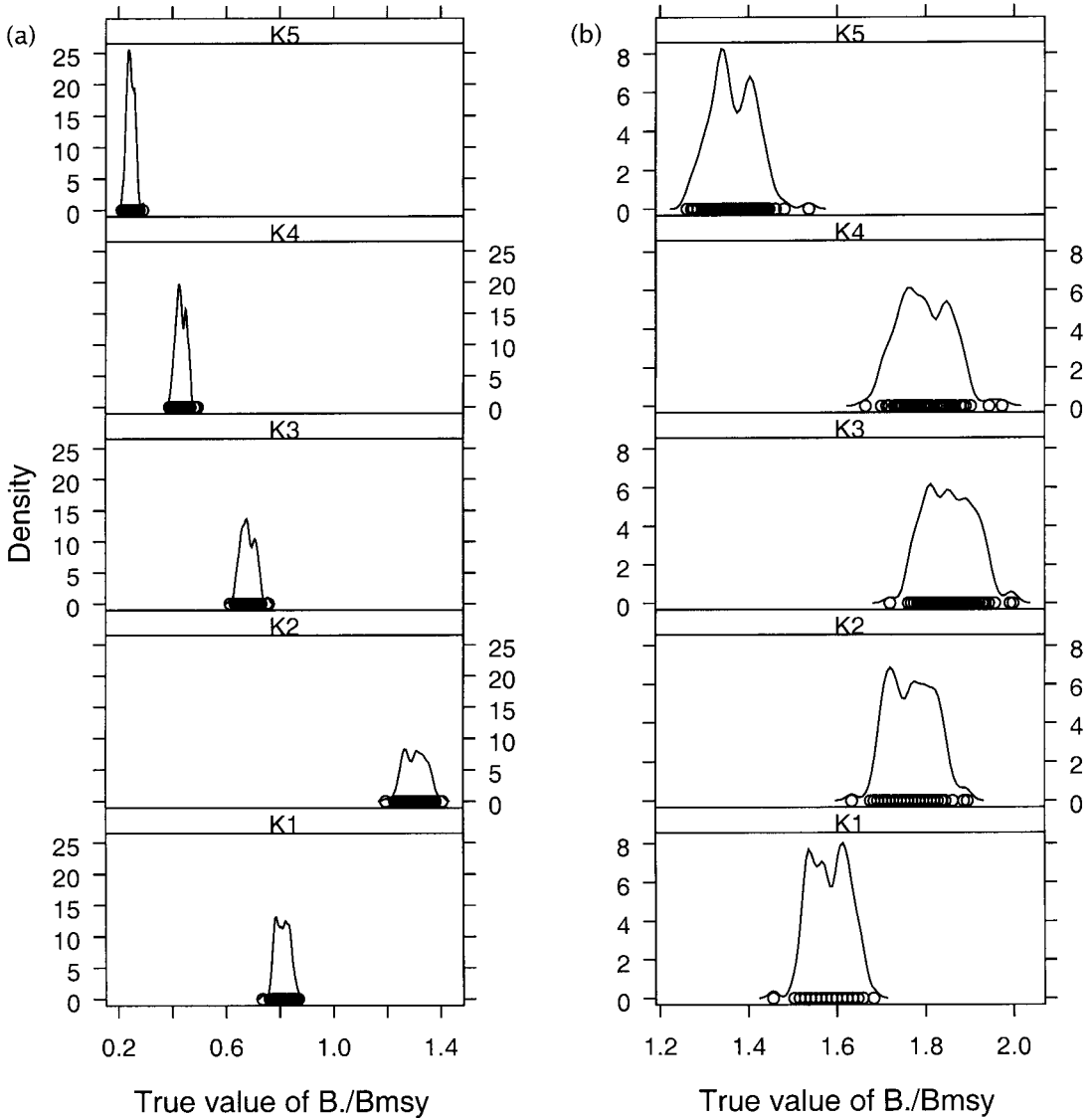
FIGURE 6.—Distributions of final-year biomass relative to that at maximum sustainable yield ($B./B_{MSY}$) realized in simulated populations based on **(a)** a trajectory of population decline and **(b)** a trajectory of population recovery. Both populations have stochastic recruitment. Distributions are shown separately by growth treatment, from K_1 (slowest growth) through K_5 (fastest growth).

would begin about 2.6 SD from the mean. To increase legibility, we have not always displayed the entire range of PE in estimates; in some cases, a few far-outside values have been repositioned very near 100% PE, the highest value of PE shown in most cases.

Estimates on Main Trajectory

Maximum sustainable yield.—Estimates of MSY based on NPUE were moderately more variable than estimates based on WPUE (Figure 8). Adding stochasticity to recruitment increased the variability in the estimates, but it did not change the median bias appreciably. Most estimates of MSY from either index were positively biased: median PE ranged from about +10% to more than +50%, depending upon the growth treatment. There was no clear pattern in the magnitude of the relative bias (as reflected in median PE) with respect to the use of NPUE versus WPUE; in some cases,
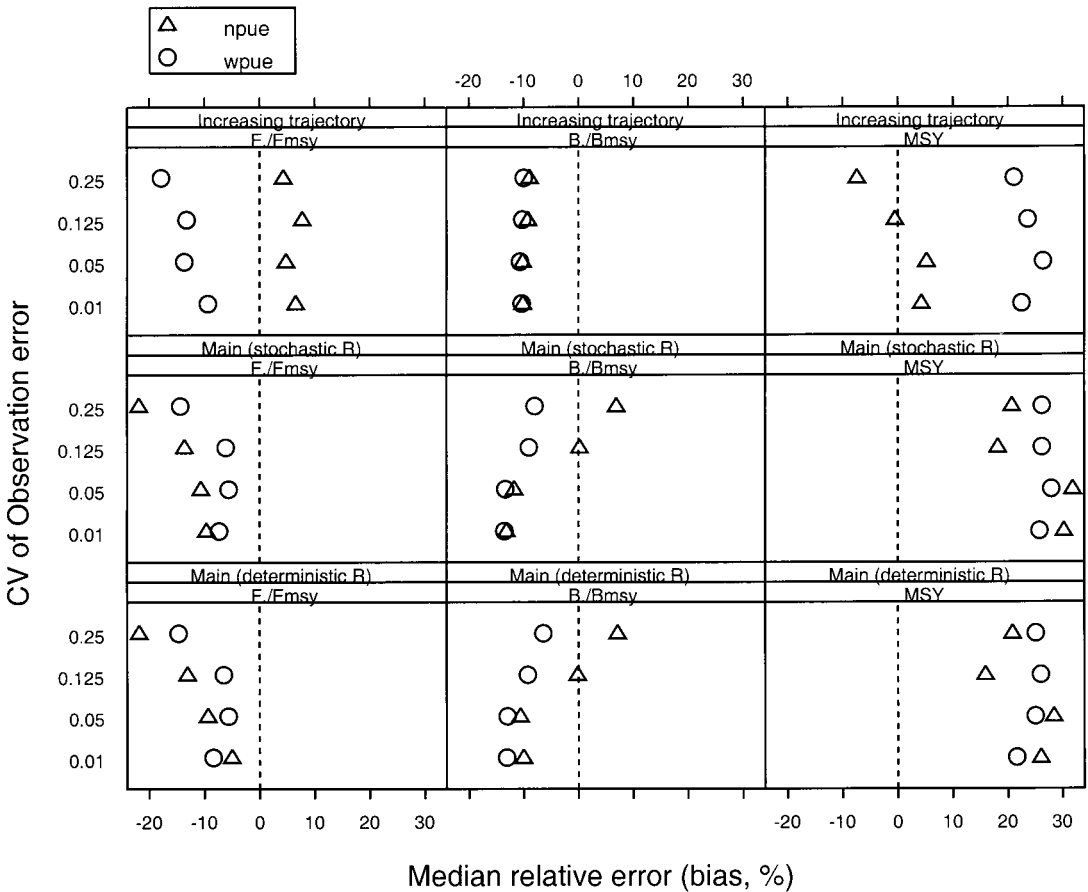
FIGURE 7.—Median percent error in management estimates as a function of the level of observation error (lognormal in CPUE and normal in catch) in data used to fit a production model. Data were on a simulated population whose parameters resemble those of the actual blue marlin population in the Atlantic Ocean. See text for a detailed explanation.

NPUE gave less biased results, while in others WPUE gave less biased results. Bias in estimating MSY generally increased with increasing values of the von Bertalanffy growth coefficient K.

Stock status after final year.—Median estimates of relative biomass $B./B_{MSY}$ after the final year were within about 20% of the true value and generally exhibited negative bias (Figure 9). As with MSY, estimates of $B./B_{MSY}$ were more variable when NPUE was used to index abundance than when WPUE was used. The bias does not appear consistently larger in either WPUE-based estimates or those based on NPUE. Adding stochasticity to recruitment increased the dispersion of the estimates, but no other effect was apparent.

Fishing status in final year.—Median estimates of the final-year relative fishing mortality rate $F./F_{MSY}$ tended to be within about 20% of the true

values (Figure 10). Estimation variability appeared to lessen with increasing growth rate. Because estimates were particularly variable with growth treatment K_1 and stochastic recruitment, the decrease of variability with increasing growth rate appears stronger with stochastic recruitment. Also, there is tendency towards greater negative bias with increasing growth rate (Figure 10). As in estimating $B./B_{MSY}$, estimation with NPUE may slightly outperform estimation with WPUE in terms of relative bias at the cost of greater variance, but systematic differences between the two indices are small.

Estimates on Alternative Trajectories

Maximum sustainable yield.—Estimates of MSY (Figure 11) based on NPUE were slightly negatively biased for the slower-growth treatments (K_1
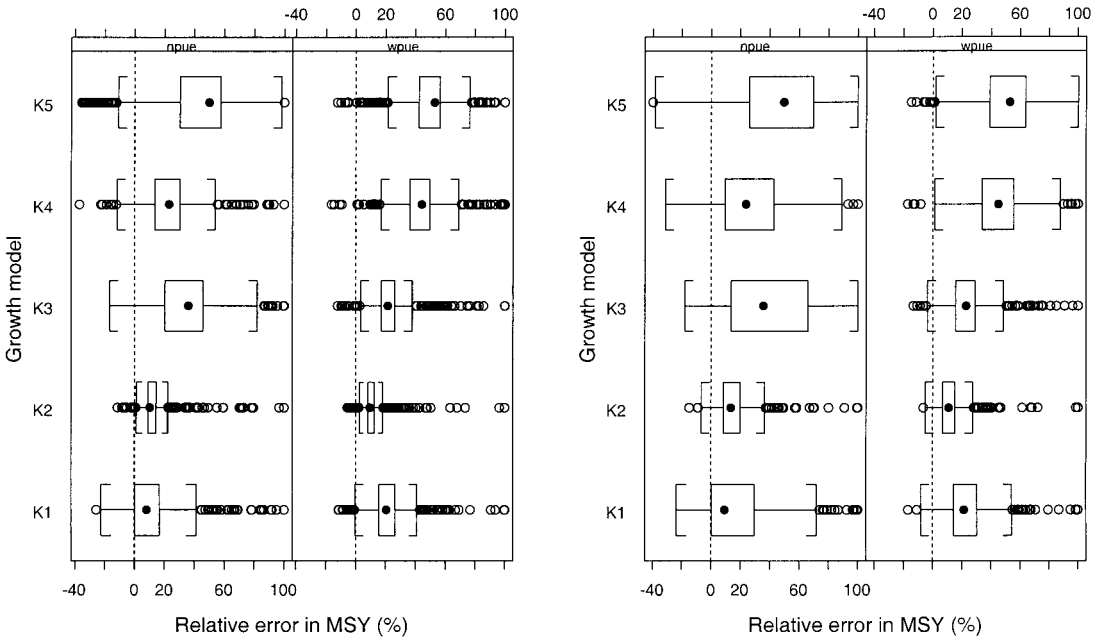
FIGURE 8.—Distribution of percent error in estimates of maximum sustainable yield (MSY) from simulated populations resembling blue marlin in the Atlantic Ocean. The left panel is based on simulated data with deterministic recruitment, the right panel on simulated data with stochastic recruitment. See text for a detailed explanation.
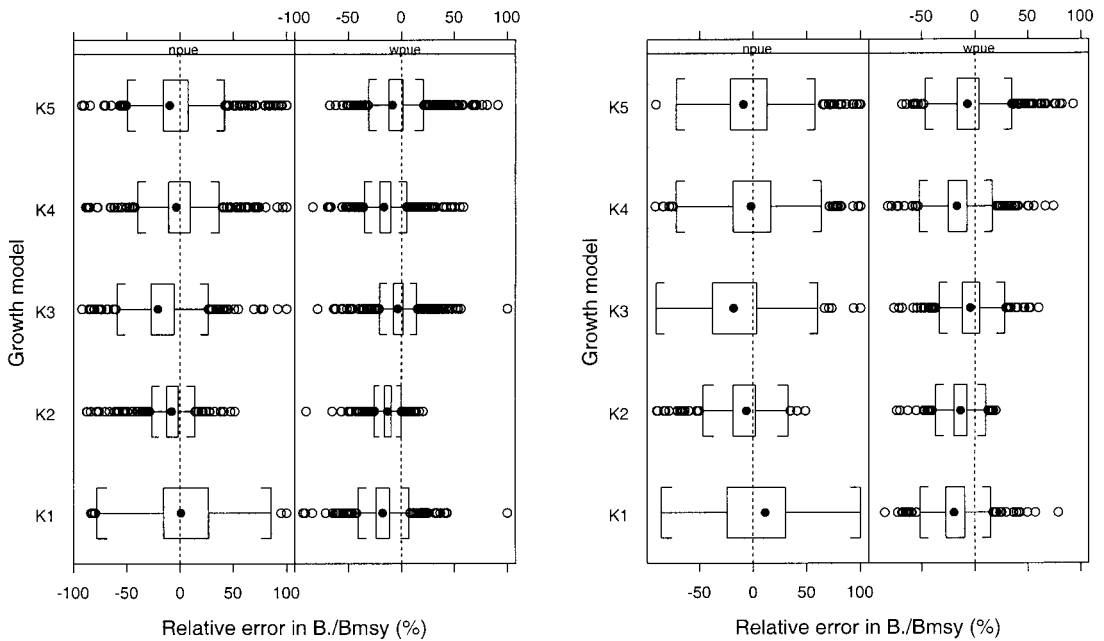

FIGURE 9.—Distribution of percent error in estimates of final-year biomass relative to that at maximum sustainable yield ($B./B_{MSY}$) estimated from simulated populations resembling blue marlin in the Atlantic Ocean. The left panel is based on simulations with deterministic recruitment, the right panel on simulations with stochastic recruitment.
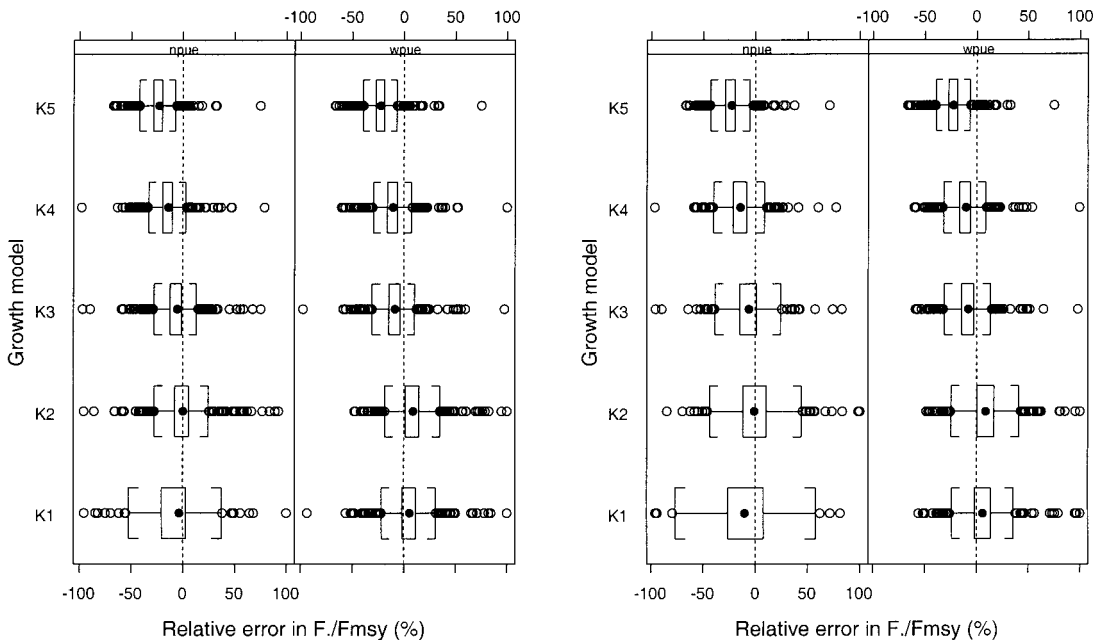
FIGURE 10.—Distribution of percent error in estimates of final-year fishing mortality rate relative to that at maximum sustainable yield ($F./F_{MSY}$), estimated from simulated populations resembling blue marlin in the Atlantic Ocean. The left panel is based on simulations with deterministic recruitment, the right panel on simulations with stochastic recruitment.

and K_2), had little bias for the central growth treatment K_3, and were slightly positively biased for the faster-growth treatments (K_4 and K_5). Compared with the results from the main trajectory (Figure 8), these results exhibited less median bias and somewhat less variability. Estimates of MSY based on WPUE exhibited consistent positive median bias (Figure 11) and about the same level of variability as those based on NPUE. This is in contrast to results from the main trajectory; there, MSY estimates from WPUE were slightly less variable than those from NPUE and of about the same median bias (Figure 8).

Stock status after final year.—Estimates of $B./B_{MSY}$ based on NPUE and WPUE (Figure 12) were less variable than those from the main trajectory (Figure 9; note the scale change between the two figures.) In both cases, biases were small, and the estimates based on WPUE exhibited slightly more negative median bias than those based on NPUE (Figures 9, 12).

Fishing status in final year.—Estimates of $F./F_{MSY}$ based on NPUE (Figure 13) were considerably more variable than those based on WPUE. Bias differed between indices, but neither was consistently better than the other (Figure 13). The variability of estimation appeared larger than from

the main trajectory (Figure 10); however, the alternative trajectories had on average much lower underlying ("true") values of this benchmark, so the errors appear larger when expressed as PE. The patterns of median bias were slightly different from those obtained on the main trajectory, but bias was not markedly different in scale.

Discussion

Our results indicate few systematic problems arising from the use of a numbers-based abundance index (NPUE) as a surrogate for the biomass-based abundance (WPUE) index assumed when fitting a surplus-production model to removals in weight. Thus, the use of mixed-metric data for stocks similar to our simulated blue marlin stock would seem to be acceptable. Surprisingly, we noticed a possible slight improvement in median relative bias when using NPUE to estimate stock and fishery status, as represented by the ratios $B./B_{MSY}$ and $F./F_{MSY}$. However, the precision of estimates was often reduced slightly when using mixed-metric data (particularly in estimates of $B./B_{MSY}$), and the improvement in bias was most pronounced when observation error was at higher levels (Figure 7). Increased imprecision counteracts reduced bias, in
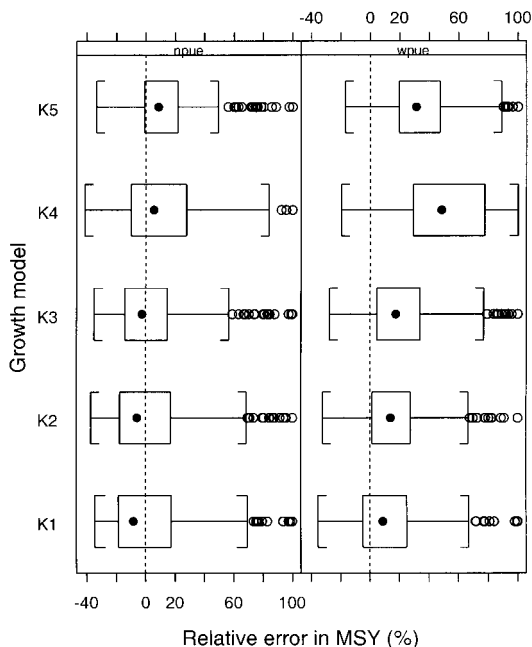
FIGURE 11.—Distribution of percent error in maximum sustainable yield (MSY) estimates from simulated populations resembling blue marlin in the Atlantic Ocean but with increasing rather than decreasing abundance. Simulated recruitment was stochastic.
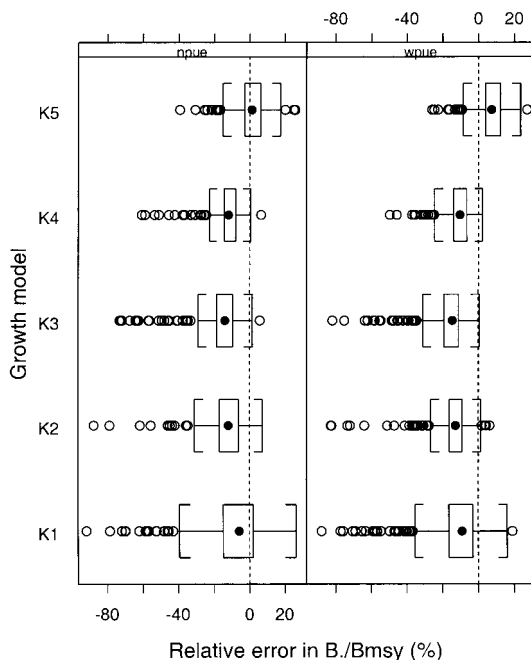

FIGURE 12.—Distribution of percent error in estimates of final-year biomass relative to that at maximum sustainable yield ($B./B_{MSY}$), estimated from simulated populations resembling blue marlin in the Atlantic Ocean but with increasing rather than decreasing population size. Simulated recruitment was stochastic.

that it increases the probability of a sizable estimation error from any particular set of data.

The success of estimation using mixed-metric data is surprising and raises two immediate questions: (1) Why should using mixed-metric data result in successful production-model estimation?; and (2) Why should using mixed-metric data in some cases be superior to using consistent units? At present, answering these questions requires some speculation. The answer to (1) appears to be that NPUE and WPUE are highly correlated despite changes in mean weight (Figure 5), so that the two indices provide similar information. For example, the correlation between the NPUE and WPUE indices in Figure 5 exceeds 0.99, and the change in time slope of each index from negative to positive occurs at about the same year. The correlation between indices is undoubtedly due in part to both indices' reflecting only the limited size range of the stock available to fishing gear. Mean weight, the link between NPUE and WPUE, depends mainly on a stock's growth pattern (which is often assumed to be constant or nearly so) and its age composition (which varies over time). In exploited stocks, the latter's expected value is changed primarily by fishing. It is difficult to imag-

ine many combinations of fishing pattern and biology that would cause NPUE and WPUE to become negatively, or even poorly, correlated over time.

Understanding why less consistent data might give a better answer is more difficult, but the cause may be offsetting biases. Specification error (mismatch between a system and the model used to represent it) is a well-studied cause of bias in parameter estimates (Kotz et al. 1982:230), and specification error is likely in fitting a surplus-production model of a complex and highly age-structured population. Use of an incorrect metric in the CPUE index would constitute a second specification error that would also be expected to bias estimation. Observation error was introduced into our simulated data in both the catch and CPUE series, which is a violation of the fitting method's assumption of error only in CPUE and which thus is a third source of bias (e. g., Figure 7). In our simulated populations, and perhaps in real populations similar to them, these biases can combine algebraically to reduce overall estimation bias.
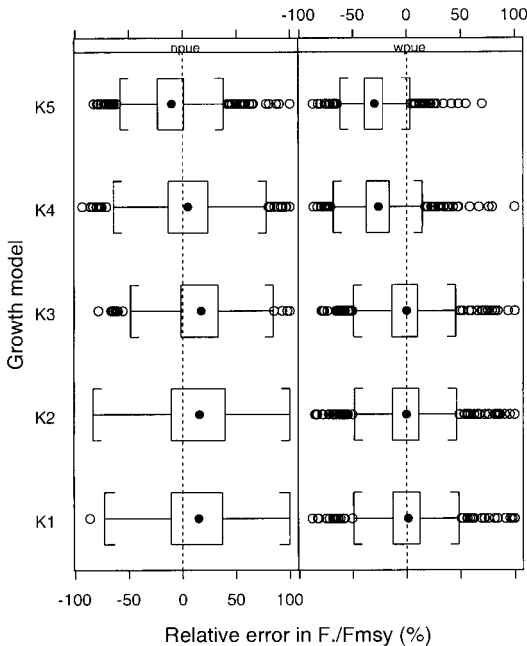
In summary, use of NPUE in place of WPUE

FIGURE 13.—Distribution of percent error in estimates of final-year fishing mortality rate relative to that at maximum sustainable yield ($F./F_{MSY}$), estimated from simulated populations resembling blue marlin in the Atlantic Ocean but with increasing rather than decreasing population size. Simulated recruitment was stochastic.

seems to be acceptable for populations similar to blue marlin in the Atlantic Ocean. This conclusion did not depend on any of the factors examined, namely, the level of observation error, presence of moderate stochasticity in recruitment, rate of individual growth, or general trend of population time trajectory. In addition, we believe it should be independent of the stock–recruitment relationship and the natural mortality rate, neither of which affects the dynamic relationship between population size in numbers and in weight. Our results should be generalizable to other stocks with similar life histories, such as white marlin and sailfish *Istiophorus platypterus*. Our results may be least applicable to stocks in which one or both of the indices reflect very young or unrecruited stock components or in which the two indices reflect different stock components. The conscientious analyst will conduct simulation studies tailored to the particular stock being assessed.

Acknowledgments

Note added in proof: A colleague has reminded us that bias is formally defined as a statistical expectation. Our terminology in this paper was thus imprecise, as the study focused on median percent error (MPE), and at times we used the word "bias." Bias itself cannot accurately be judged from the graphical presentations used here. Although bias and MPE both measure central tendancy of an estimator's errors, the two coincide (apart from scaling) only if estimation errors are symmetrically distributed. We prefer to use MPE in this context for several reasons, an important one being that it is less prone to influence by a few large estimation errors.

References

Boggs, C. H. 1989. Vital rate statistics for billfish stock assessment. Pages 225–233 *in* R. H. Stroud, editor. Planning the future of billfishes: research, and management in the 90's, and beyond. Marine Recreational Fisheries 13, Part 1. National Coalition For Marine Conservation, Savannah, Georgia.

Cyr, E. C. 1987. Age, growth, and reproduction of blue marlin, *Makaira nigricans,* from South Carolina billfish tournament collections. Master's thesis. University of South Carolina, Columbia.

De Sylva, D. P., and P. R. Breder. 1997. Reproduction, gonad histology, and spawning cycles of north Atlantic billfishes (Istiophoridae). Bulletin of Marine Science 60:668–697.

Erdman, D. S. 1968. Spawning cycle, sex ratio, and weights of blue marlin off Puerto Rico and the Virgin Islands. Transactions of the American Fisheries Society 97:131–137.

Fox, W. W. 1970. An exponential yield model for optimizing exploited fish populations. Transactions of the American Fisheries Society 99:80–88.

Goodyear, C. P. 1989. LSIM: A length-based fish population simulation model. National Oceanic and Atmospheric Administration, Technical Memorandum NMFS-SEFC-219, Washington, D.C.

Goodyear, C. P. 1996. Variability of fishing mortality by age: Consequences for MSY. North American Journal of Fisheries Management 16:8–13.

Hill, K. T., G. M. Cailliet, and K. T. Hill. 1989. A comparative analysis of growth zones in four calcified-structures of Pacific blue marlin, *Makaira nigricans.*

U.S. National Marine Fisheries Service Fishery Bulletin 87:829–843.

Hoenig, J. M. 1983. Empirical use of longevity data to estimate mortality rates. U.S. National Marine Fisheries Service Fishery Bulletin 82:898–902.

ICCAT (International Commission for the Conservation of Atlantic Tunas). 1999. Report of the third ICCAT billfish workshop. ICCAT, Madrid.

Kotz, S., N. L. Johnson, and C. B. Read, editors. 1982. Encyclopedia of statistical sciences, volume 1. Wiley, New York.

Lotka, A. J. 1924. Elements of physical biology. Reprinted 1956 as Elements of mathematical biology by Dover Press, New York.

Mathsoft, Inc. 1999. S-Plus 2000 programmer's guide. Data Analysis Products Division, Mathsoft, Seattle, Washington.

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. Doctoral dissertation. University of Washington, Seattle.

Pella, J. J., and P. K. Tomlinson. 1969. A generalized stock production model. Bulletin of the Inter-American Tropical Tuna Commission 13:419–496.

Prager, M. H. 1994. A suite of extensions to a non-equilibrium surplus-production model. U.S. National Marine Fisheries Service Fishery Bulletin 92:374–389.

Prager, M. H. 1995. User's manual for ASPIC: a stock-production model incorporating covariates, program version 3.6x. U.S. National Marine Fisheries Service, Southeast Fisheries Science Center, Miami Laboratory, 4th edition, Document MIA-2/93–55, Miami.

Prager, M. H., C. P. Goodyear, and G. P. Scott. 1996. Application of a surplus production model to a swordfish-like simulated stock with time-changing gear selectivity. Transactions of the American Fisheries Society 125:729–740.

Prince, E. D., D. W. Lee., J. R. Zweifel, and E. B. Brothers. 1991. Estimating age and growth of young Atlantic blue marlin *Makaira nigricans* from otolith microstructure. U.S. National Marine Fisheries Service Fishery Bulletin 89:441–459.

Schaefer, M. B. 1954. Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. Bulletin of the Inter-American Tropical Tuna Commission 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. Bulletin of the Inter-American Tropical Tuna Commission 2:247–268.

Schnute, J. 1977. Improved estimates from the Schaefer production model: theoretical considerations. Journal of the Fisheries Research Board of Canada 34:583–603.

Tukey, J. W. 1977. Exploratory data analysis. Addison-Wesley Publishing, Reading, Massachusetts.

von Bertalanffy, L. 1938. A quantitative theory of organic growth. Human Biology 10:181–213.

Wilson, C. A. 1984. Age and growth aspects of the life history of billfishes. Doctoral dissertation. University of South Carolina, Columbia.

Wilson, C. A., J. M. Dean, E. D. Prince, and D. W. Lee. 1991. An examination of sexual dimorphism in Atlantic and Pacific blue marlin using body weight, sagittae weight, and age estimates. Journal of Experimental Marine Biology and Ecology 151:209–225.