Fitting Surplus Production Models: Comparing Methods and Measuring Uncertainty

Tom Polacheck

CSIRO Division of Fisheries Research, GPO Box 1538, Hobart, Tasmania 7001, Australia

Ray Hilborn

School of Fisheries WH-10, University of Washington, Seattle, WA 98195, USA

and Andre E. Punt

School of Fisheries WH-10, University of Washington, Seattle, WA 98195, USA

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

Three approaches are commonly used to fit surplus production models to observed data: effort-averaging methods; process-error estimators; and observation-error estimators. We compare these approaches using real and simulated data sets, and conclude that they yield substantially different interpretations of productivity. Effort-averaging methods assume the stock is in equilibrium relative to the recent effort; this assumption is rarely satisfied and usually leads to overestimation of potential yield and optimum effort. Effort-averaging methods will almost always produce what appears to be "reasonable" estimates of maximum sustainable yield and optimum effort, and the r^2 statistic used to evaluate the goodness of fit can provide an unrealistic illusion of confidence about the parameter estimates obtained. Process-error estimators produce much less reliable estimates than observation-error estimators. The observation-error estimator provides the lowest estimates of maximum sustainable yield and optimum effort and is the least biased and the most precise (shown in Monte-Carlo trials). We suggest that observation-error estimators be used when fitting surplus production models, that effort-averaging methods be abandoned, and that process-error estimators should only be applied if simulation studies and practical experience suggest that they will be superior to observation-error estimators.

On emploie communément trois méthodes pour ajuster les modèles de production excédentaire aux résultats observés; il y a les méthodes de la moyenne d'effort, les estimateurs des erreurs de traitement ainsi que les estimateurs des erreurs d'observation. Nous comparons ces trois démarches au moyen d'ensembles de données réelles et simulées, et nous parvenons à la conclusion que ces méthodes conduisent à des interprétations largement différentes de la productivité. Les méthodes fondées sur les moyennes d'effort supposent que le stock est en équilibre relativement à l'effort récent; c'est rarement le cas, mais cela conduit ordinairement à une surestimation du rendement potentiel et de l'effort optimal. Ces méthodes produiront presque toujours ce qui semble être des estimations « raisonnables » du rendement soutenable maximal et de l'effort optimal, et la valeur statistique r^2 qui sert à évaluer la validité de l'ajustement peut donner l'illusion non fondée de confiance dans les estimations des paramètres qui sont obtenues. Les estimateurs d'erreurs de traitement donnent des estimations beaucoup moins fiables que les estimateurs des erreurs d'observation. Ces derniers conduisent aux estimations les plus faibles de rendement soutenable maximal et d'effort optimal; ce sont aussi les estimateurs les plus précis et qui comportent le moins d'erreur systématique (selon la méthode Monte-Carlo). Nous proposons d'utiliser des estimateurs d'erreurs d'observation lors de l'ajustement de modèles de production excédentaire, d'abandonner les méthodes de moyenne des efforts, et de n'appliquer les estimateurs d'erreurs de traitement que lorsque des études de simulation et que l'expérience immédiate donnent à penser que ces estimateurs seront supérieurs aux estimateurs d'erreurs d'observation.

Received May 22, 1991
Accepted July 9, 1993
(JB049)

Reçu le 22 mai 1991
Accepté le 9 juillet 1993

The simplest models of fish population dynamics consider only the changes in the exploitable biomass of the fishable stock. These models are generally called surplus production models (Ricker 1975) or more precisely biomass dynamics models (Hilborn and Walters 1992). In their simplest form, only a time series of catch and a relative abundance index (often standardized catch-per-unit effort, CPUE) are needed to estimate the model parameters. From these estimates, two quantities of considerable importance to management can be derived: the maximum sustainable yield (MSY) and the fishing effort at which MSY will be achieved (E_{MSY} — referred to as optimal effort, also sometimes denoted as f_{opt}). The major appeal of biomass dynamics models is that they can be applied in a situation in which the only data are catches and a relative abundance index. The computational simplicity of some of the fitting procedures for biomass dynamics models has meant that

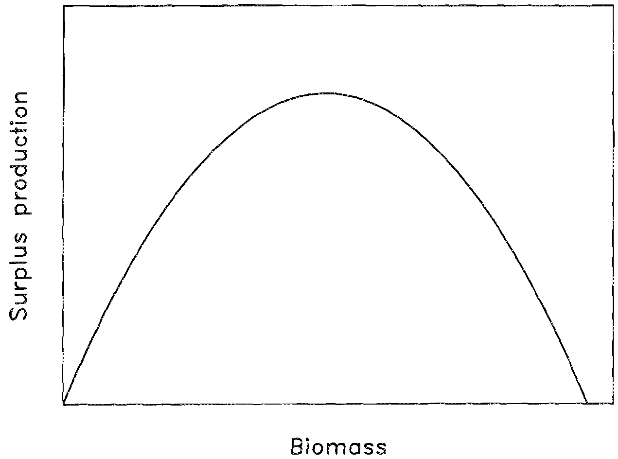

FIG. 1. A generic surplus-production function.

they have been widely used. Paucity of data is often a problem associated with a developing fishery, so that biomass dynamics models are frequently applied in this situation.

While age- and size-structured models (e.g., Deriso et al. 1985; Pope and Shepherd 1985; Gavaris 1988; Bergh and Johnston 1992) are used by many management agencies which have access to a time series of catch-at-age/size data, biomass dynamics models are still used in the management of many fisheries. This may be because the age/size compositions of the historic catches are not available or reliable. Another reason for considering biomass dynamics models is that in some situations they can provide more accurate and precise estimates of management-related quantities than more complex approaches (Ludwig and Walters 1985; Punt 1988, 1992a, b).

The method used to fit a biomass dynamics model to the observed data has been shown to be of much greater importance in terms of the reliability of estimated parameters than the algebraic form of the underlying population dynamics model (Punt 1988, 1992b). There are several ways to fit a biomass dynamics model to a set of observed data. However, only three have been widely used: (a) effort-averaging methods (Gulland 1961; Fox 1975), (b) process-error estimators (Walters and Hilborn 1976; Schnute 1977), and (c) observation-error estimators (Pella and Tomlinson 1969; Butterworth and Andrew 1984; Ludwig and Walters 1985)

Monte-Carlo simulation testing of estimation methods has shown that effort-averaging approaches are frequently highly (positively) biased, and process-error estimators are often very imprecise (Hilborn 1979; Uhler 1980; Punt 1988, 1992b). However, these two approaches continue to be used and even the most recent applications (e.g., Bartoo and Shiohama 1985; Sparre et al. 1989; Yeh et al. 1991) do not mention the existence of observation-error estimators nor do they appear to be aware of the deficiencies of the estimation methods being used. This is perhaps not surprising since the two most commonly-used textbooks in fisheries (Ricker 1975; Gulland 1983) do not mention process- or obervation-error estimators.

The objective of this paper is to review and compare the three methods and to show how the uncertainty associated with the estimates can be evaluated. No such comparison now exists; Hilborn and Walters (1993) provide a brief review of these methods but do not provide direct comparison nor do they show the computational procedures for calculating confidence bounds by likelihood profile. Butterworth and Andrew (1987) compare the methods by evaluating the precision of their estimates of $f_{0.1}$ harvesting strategy TACs for the four hake stocks off southern Africa, but they do not comment on the bias associated with these methods nor do they consider stock dynamics and perturbation histories other than those for hake.

The comparison includes two steps. First, the three estimation methods are applied to the actual data for three major marine resources [the Cape hake (*Merluccius capensis* and *M. paradoxus*) stock off northern Namibia, New Zealand rock lobster (*Jasus edwardsii*), and the south Atlantic albacore (*Thunnus alalunga*)]. These three data sets were selected because surplus-production models have been recently applied to them. Annual age-composition data do not exist for the rock lobster and albacore stocks so it is not possible to apply an age-based stock assessment procedure to them. Age-composition data for the years 1965 to 1989 are available for the hake stock, but are of questionable value. Punt and Butterworth (1989) applied a Laurec-Shepherd ad hoc tuned VPA to the data for this stock and found the relationship between fishing effort and fishing mortality to be extremely poor — to such an extent that the estimates of the quantities of importance to management were essentially worthless. The scientific basis for the TACs set by the International Commission for the South East Atlantic (ICSEAF) for this resource was thus based on fits of surplus-production models to catch and catch-rate data. The data sets for two of these three stocks (albacore and rock lobster) consist of a time series of increasing effort and decreasing index of abundance, which (Hilborn 1979) called a "one-way-trip" and showed were uninformative. The data set for the Cape hake stock is somewhat more informative as the index of abundance initially decreases, but then increases when effort is reduced.

The second part of the comparison involves a Monte-Carlo simulation evaluation of the estimation ability of the three methods. Simulations are based on each of the three stocks in order to detect the effects of the length of the data series and the extent of data contrast on estimation ability.

Models and Estimators

Biomass Dynamics Models

The essential feature of a biomass dynamics model-estimation procedure is the parameterization of the relationship between fishing intensity and long-term sustainable yield based on an assumed relationship between stock biomass and production. A number of alternative functional forms for this relationship exist, but all have the general shape shown in Fig. 1 [see the summary given in Punt (1988)]. Deterministically, all biomass dynamics models are of the form

$$(1)\quad B_{y+1} = B_y + g(B_y) - C_y$$

$$(2)\quad I_y = qB_y$$

where B_y is the (exploitable) biomass at the start of year y, $g(B)$ is surplus production as a function of biomass, for example either $g(B) = rB(1 - B/K)$ (Schaefer 1954 form), $g(B) = rB(1 - \log(B)/\log(K))$ (Fox 1970 form), or $g(B) = \frac{r}{p} \times B(1 - (B/K)^p)$ (Pella–Tomlinson 1969 form), r is the intrinsic growth rate parameter, K is the average biomass level prior to exploitation, q is the catchability coefficient, C_y is the catch during year y, and I_y is a index of relative abundance for year y.

Note that in the Pella–Tomlinson form, the r/p term is often omitted when the formula is presented. The parameter p in the Pella–Tomlinson form controls the asymmetry of the sustainable yield versus stock biomass relationship. It can be shown that the Schaefer form is equivalent to the Pella–Tomlinson form with $p = 1$ and that the Fox form is the limit of the Pella–Tomlinson form as $p \to 0$.

The various approaches to fitting biomass dynamics models to observed data considered in this paper (effort-averaging, observation-error, and process-error), differ in how error is introduced into Equations (1) and (2). The following discussion considers the case in which only one index of abundance is available.

Equilibrium Estimators

MSY and E_{MSY} can be estimated by assuming that the rate of change of biomass (i.e., dB/dt) is zero for all years (i.e., $B_{t+1} = B_t$ = Constant), and assuming that Equation (2) is exact. Now, it is often the case that the index of abundance $\{I_y\}$ is a time series of catch-rates (i.e., $qB_y = I_y = C_y / E_y$, where E_y is the fishing effort during year y). Solving Equation (1) for C_y after assuming that $B_{y+1} = B_y$, and after substituting $C_y / (qE_y) = B_y$, gives

$$(3)\quad C_y = \frac{rC_y}{pqE_y}\left(1 - \left(\frac{C_y}{qE_y K}\right)^p\right).$$

Equation (3) can then be solved for C_y / E_y:

$$(4)\quad C_y/E_y = \left((qK)^p - \frac{pq^{p+1} K^p E_y}{r}\right)^{1/p}.$$

If we define the first term $(qK)^p$ to be a new parameter a and the second term $\frac{pq^{p+1}K^p}{r}$ to be a new parameter b, then sum of squares estimates of the parameters a, b, and p can then be obtained by minimizing the quantity

$$(5)\quad \sum_y \left((C/E)_y - (\widehat{C/E})_y\right)^2$$

where $(C/E)_y$ is the observed catch-rate for year y and $(\widehat{C/E})_y$ is the model-predicted catch-rate for year y.

If the value of p is assumed to be one, estimates for the parameters a and b can be obtained using standard linear regression techniques. Estimates of MSY and E_{MSY} are obtained by means of the formulae

$$(6)\quad E_{MSY} = \frac{pa}{b\,(p+1)}$$

$$MSY = \frac{p}{b}\left(\frac{a}{p+1}\right)^{\frac{p+1}{p}}.$$

Effort-Averaging Approaches

It is recognized that fish stocks are rarely, if ever, in equilibrium. The most common approach to overcome this difficulty is to replace effort of the right hand side of Equation (4) by a weighted average fishing effort (Gulland 1961; Fox 1975). The common way in which to obtain the weighted average fishing effort for year y is by means of the formula suggested by Fox(1975), i.e.:

$$(7)\quad \bar{E}_y = \frac{kE_y + (k-1)\,E_{y-1} + \ldots + E_{y-k+1}}{k + (k-1) + \ldots + 1}$$

where k is the number of age classes being fished.

Combining Equation (4) with Equation (7) assumes recruitment is independent of spawner stock size; the stock size depends only on historical effort and not the size of the stock in the recent past (no spawner recruit effect). This assumption is not generally acknowledged by users of the method and effort-averaging methods are in practice an ad hoc approach for dealing with non-equilibrium conditions.

Process-Error Estimators

A process-error estimator is based on the assumption that the observations are made without error and that all of the error occurs in the change in population size [i.e., Equation (2) is assumed to be exact and Equation (1) is assumed to be subject to error]. For the Pella–Tomlinson form of the biomass dynamics function, substituting I_y for qB_y in Equation (1) and simplifying gives

$$(8)\quad I_{y+1} = I_y + \frac{r}{p} I_y \left(1 - (I_y/qK)^p\right) - qC_y.$$

For the case in which $p = 1$ (the Schaefer form), Equation (8) can be written in a form which is linear in its parameters:

$$(9)\quad I_{y+1} = (1+r)\,I_y - \frac{r}{qK}(I_y)^2 - qC_y.$$

In this form, estimates of r, q, and K can be obtained by multiple linear regression.

A large number of variants of this general approach exist. For example, Schnute (1977) provides the variant of Equation (9) which results from replacing Equation (1) with a continuous equation, while Lleonart et al. (1985) and Lleonart and Salat (1989) present a process-error estimator based on the concept of the inertia of a stock.

Observation-Error Estimators

An observation-error estimator (Pella and Tomlinson 1969; Butterworth and Andrew 1984; Ludwig and Walters 1985; Ludwig et al. 1988) is constructed by assuming that the population dynamics equation (Equation 1) is deterministic and that all of the error occurs in the relationship between stock biomass and the index of abundance. The stock biomass time series is estimated by projecting the biomass at the start of the catch series ($B_{initial}$) forward under the historic annual catches.

Assuming that the error in Equation (2) is multiplicative and log-normal with a constant coefficient of variation (i.e., $I_y = qB_y e^{\varepsilon}$, $\varepsilon \sim N(0;\sigma^2)$), the estimates of the model parameters ($B_{initial}$, r, q, and K) are obtained by maximizing the appropriate likelihood function:

$$(10)\quad L = \prod \exp\left\{-\hat{v}_y^2/(2\hat{\sigma}_v^2)\right\}/(\sqrt{2\pi}\,\hat{\sigma}_v)$$

where the product is over all years (y) for which CPUE data are available:

$$\hat{v}_y = \log(C/E)_y - \log(\widehat{C/E})_y,$$

$$\hat{\sigma}_v^2 = \sum \hat{v}_y^2/n$$

where n is the number of data points.

Equation (10) is an expression which is non-linear in its four parameters and so a non-linear minimization approach (e.g., Press et al. 1986) must be applied to obtain values for these parameters. It can be shown that the value of q which minimizes Equation (10) is given by the formula

$$(11) \quad \hat{q} = \exp\left\{\frac{1}{n}\sum_y \log(I_y/\hat{B}_y)\right\}.$$

Two other common assumptions regarding observation-error noise are that it is either additive and normal with a constant standard deviation (i.e., $I_y = qB_y + \varepsilon$, $\varepsilon \sim N(0; \sigma^2)$), or additive and normal with a constant coefficient of variation (i.e., $I_y = qB_y + \varepsilon$, $\varepsilon \sim N(0; (\sigma q B_y)^2)$).The maximum likelihood estimates of q for these choices of error structures are for an additive normal

$$(12) \quad \hat{q} = \sum_y I_y B_y / \sum_y B_y^2$$

and for a constant CV

$$(13) \quad \hat{q} = \frac{-\sum_y (I_y/B_y) + \sqrt{\left(\sum_y (I_y/B_y)\right)^2 + 4n\sigma^2 \sum_y (I_y/B_y)^2}}{2n\sigma^2}.$$

In practice, the likelihood surface is usually relatively flat near its maximum value (e.g., Rivard and Bledsoe 1978) and it is often necessary to introduce additional constraints (such as that $B_{\text{initial}} = K$). Fletcher (1978) discusses the source of this flatness while Punt (1990) found by simulation that for the Cape hake stock off northern Namibia, even in situations in which B_{initial}/K is substantially different from unity, better estimation performance is achieved by fixing B_{initial}/K at unity than by estimating it.

Uncertainty of Estimates and Confidence Bounds

One way to quantify the uncertainty associated with an estimate is to compute its confidence bounds. The three most frequently used methods are: asymptotic methods which assume that the likelihood function is quadratic near its minimum (Draper and Smith 1966), bootstrap approaches (e.g., Efron 1982; Punt and Butterworth 1993), and likelihood profile (Venzon and Moolgavkor 1988). Likelihood profile is used here because of its generality and computational ease.

The basic principles of likelihood and likelihood ratio can be used to define the confidence bounds for estimated parameters of any model for which the likelihood can be determined (Press et al. 1986, p. 532; Venzon and Moolgavkor 1988). For a single parameter p the confidence interval is defined as all values of parameter p that satisfy the inequality

$$(14) \quad 2[L(Y \mid p) - L(Y \mid p_{\text{best}})] \leq \chi^2_{1,1-\alpha}$$

where $L(Y|p_{\text{best}})$ is the log likelihood of the most likely value of p and $\chi^2_{1,1-\alpha}$ is value of the chi-squared distribution with 1 df at confidence level $1-\alpha$.

Thus, the 95% confidence interval for p encompasses all values of p for which twice the difference between the log likelihood and the log likelihood of the best estimate of p is less than 3.84.

Likelihood profiles can be used to determine confidence bounds for the parameters either jointly or individually. The confidence region for n parameters is estimated based on the χ^2 distribution with n degrees of freedom. The likelihood profile method is preferred because it is computationally more efficient than bootstrapping and because many confidence regions are asymmetric (i.e., in two dimensions are often banana-shaped) rather than symmetric ellipses as assumed by asymptotic methods.

Application to Real Data

To ensure that comparisons made in this section are comparable, all data were fit to the Schaefer form of the surplus production function (i.e., $p = 1$), and for the effort-averaging method k (see Equation 7) was set to 3.

The results for each stock are given in the same format. The catch-rate data used in the assessments are presented in Table 1 and the estimates of a number of management-related quantities for each assessment method in Table 2. Note that the effort-averaging method does not provide estimates of biomass, so the estimates of quantities related to biomass are left out of Table 2. For each stock, a figure (Fig. 2, 4, and 6) is provided showing two plots. The first plot (panel a) shows the actual catches plotted against the corresponding efforts, with the equilibrium catch-effort curves obtained from each estimation method superimposed. The second plot (panel b) shows the catch-rate time series with the fits of the process- and observation-error estimators superimposed. Finally, for each stock, a figure (Fig. 3, 5, and 7) shows the point estimates of MSY and E_{MSY} and the 95% confidence bounds for these parameters computed using the method of likelihood profile.

South Atlantic Albacore

Up to 1990, the assessment of this stock by ICCAT (the International Commission for the Conservation of Atlantic Tuna) was based solely on the application of the effort-averaging method of Fox (1975). The value of p used in these assessments was 0.001 (Yeh et al. 1991).

The results of the three assessments are different in several respects, particularly regarding productivity (Table 2, Fig. 2 and 3). The effort-averaging method produces the most optimistic appraisal of the situation. According to this assessment, MSY is just larger than 28 000 t and the optimal effort is over 100 000 000 hooks. In only 6 yr is the annual catch larger than the estimated MSY (Table 1) which makes it rather difficult to see why the catch-rate should have dropped to roughly 30% of its initial size. The observation-error estimator is the least optimistic of the three assessments suggesting an MSY of only 19 650 t and an optical effort level roughly half that suggested by the effort-averaging method. The estimates of MSY and E_{MSY} obtained from the application of the process-error estimator are intermediate between the estimates obtained from the other methods (Table 2).

One of the reasons for the high estimates of MSY and E_{MSY} provided by the effort-averaging method is that there are a number of very high (E_{MSY}, MSY) data points (see Fig. 2a). Since the effort-averaging method assumes that the stock is relatively close to equilibrium, it attempts to pass the equilibrium catch-effort curve through these points. The other two methods assume instead that the catch taken reflects surplus production and reduction of standing stock. It is possible to infer from the trend in the catch-rate data that the stock has been declining over

TABLE 1. Catch and catch-rate data for the three stocks considered in this paper.

Year	New Zealand rock lobster Catch (t)	New Zealand rock lobster CPUE	Northern Namibian hake Catch ('000 t)	Northern Namibian hake CPUE	South Atlantic albacore Catch ('000 t)	South Atlantic albacore CPUE (kg/100 hooks)
1945	809	3.49				
1946	854	3.38				
1947	919	3.18				
1948	1360	3.56				
1949	1872	1.79				
1950	2672	4.35				
1951	2834	2.33				
1952	3324	2.57				
1953	4160	2.88				
1954	5541	3.85				
1955	5909	4.16				
1956	6547	4.34				
1957	5049	3.70				
1958	4447	2.37				
1959	4018	2.46				
1960	3762	2.06				
1961	4042	2.21				
1962	4583	2.19				
1963	4554	2.44				
1964	4597	2.14				
1965	4984	2.18	93.510	1.78		
1966	5295	2.13	212.444	1.31		
1967	4782	1.86	195.032	0.91	15.9	61.89
1968	4975	1.53	382.712	0.96	25.7	78.98
1969	4786	1.32	320.430	0.88	28.5	55.59
1970	4699	1.45	402.467	0.90	23.7	44.61
1971	4478	1.40	365.557	0.87	25.0	56.89
1972	3495	1.09	606.084	0.72	33.3	38.27
1973	3784	1.23	377.642	0.57	28.2	33.84
1974	3643	1.12	318.836	0.45	19.7	36.13
1975	2987	0.92	309.374	0.42	17.5	41.95
1976	3311	1.02	389.020	0.42	19.3	36.63
1977	3237	1.00	276.901	0.49	21.6	36.33
1978	3418	1.05	254.251	0.43	23.1	38.82
1979	4050	1.09	170.006	0.40	22.5	34.32
1980	4190	1.03	97.181	0.45	22.5	37.64
1981	4058	1.01	90.523	0.55	23.6	34.01
1982	4331	0.98	176.532	0.53	29.1	32.16
1983	4385	0.88	216.181	0.58	14.4	26.88
1984	4911	0.85	228.672	0.64	13.2	36.61
1985	4856	0.84	212.177	0.66	28.4	30.07
1986	4657	0.81	231.179	0.65	34.6	30.75
1987	4500	0.84	136.942	0.61	37.5	23.36
1988	3128	0.68	212.000	0.63	25.9	22.36
1989	3318	0.62			25.3	21.91
1990	2770	0.54				

most of the period and so almost all of the catches reflect surplus production plus some fraction of the standing stock. This is why the equilibrium catch-effort curves for the process-error estimator and particularly the observation-error estimator lie below most of the observed data points (see Fig. 2a).

The estimates of current depletion obtained from the observation- and the process-error estimators are quite similar at roughly 32%. This result is not particularly surprising since the catch-rate series essentially determines the value of this quantity. However, there are major discrepancies between the estimates of current biomass provided by these two approaches. The observation-error estimator suggests that the stock was initially large with low productivity while the process-error estimator suggests exactly the opposite.

The apparent conflict between the process-error and observation-error point estimates of MSY and E_{MSY} are minor when the likelihood profile 95% confidence bounds are considered (Fig. 3). It is clear from this figure that the process-error estimator is extremely imprecise. Surprisingly, the results of the effort-averaging method appear to be quite precise and do not overlap with those of the (very precise) observation-error estimator. Hilborn (1979) showed that when the data are a one-way trip (constantly declining index of abundance, constantly increasing effort), a process-error estimator attempts

TABLE 2. Estimates of a number of management quantities obtained by applying three model-estimation procedures to the data in Table 1. Effort units are as in Table 1.

Quantity	Effort-averaging ($k = 3, p = 1$)	Observation-error estimator	Process-error estimator
		South Atlantic albacore	
r	—	0.328	0.620
q ($\times 10^4$)	—	26.71	43.72
K ('000 t)	—	239.6	153.4
MSY ('000 t)	28.09	19.65	23.78
B_{initial}/K	—	1.000	0.923
B_{current} ('000 t)	—	75.51	50.04
B_{current}/K	—	0.315	0.326
E_{MSY}	102.5	61.4	70.9
σ	—	0.111	0.736
		New Zealand rock lobster	
r	—	0.0659	0.448
q ($\times 10^5$)	—	2.461	6.184
K ('000 t)	—	129.0	52.6
MSY (t)	5278.5	2133.74	5897.7
B_{initial}/K	—	1.000	1.074
B_{current} (t)	—	21.15	0.87
B_{current}/K	—	0.163	0.166
E_{MSY}	2994.9	1337.9	3629.0
σ	—	0.207	0.568
		Northern Namibian hake	
r	—	0.379	0.304
q ($\times 10^4$)	—	4.360	2.701
K ('000 t)	—	2772.6	3448.2
MSY ('000 t)	365.8	263.2	262.3
B_{initial}/K	—	1.000	1.911
B_{current} ('000 t)	—	1646.3	2332.66
B_{current}/K	—	0.594	0.676
E_{MSY}	871.2	435.5	563.2
σ	—	0.124	0.662

to fit a 3-dimensional plane to a 2-dimensional line and will therefore be very imprecise.

Note from the formula given earlier that the observation-error estimation can use data from years where catch, but not catch-rate are known, and can use only occasional catch-rate data, while the process-error method can only use pairs of catch and catch-rate data.

New Zealand Rock Lobster

This stock (data from Breen 1991) has been assessed using both effort-averaging methods (Fox 1975) and process-error estimators. Qualitatively, the results for this stock (Table 2, Fig. 4 and 5) are the same as for South Atlantic albacore. The observation-error estimator provides the least optimistic appraisal once again, although this time it is the process-error estimator which is the most optimistic (if only slightly). Both the process-error and the observation-error estimators suggest that the stock is very highly depleted (to roughly 16–17% of K) although there are marked differences between the estimates of K and hence current biomass produced by these two estimators. The process-error estimator indicates a highly productive but small resource whereas the observation-error estimator suggests a highly unproductive resource which was initially quite large (Table 2).

Once again the results of the process-error estimator are very imprecise (Fig. 5). The extremely high precision associated with the effort-averaging method is due to the fact that the method is essentially performing a regression on the catch-rate versus effort data, which, being a one-way trip, are highly correlated. Users of effort averaging methods have often used this high correlation (r^2 is often >0.95) as an indication that the model is correct. Note that in this case all the catch-effort data points lie above the catch-effort curve estimated from the observation-effort estimator (Fig. 4a). This is because with the very low productivity rate estimated (0.0659, see Table 2), the catches are estimated to have been comprised almost entirely of standing stock. Surplus-production is estimated to have been almost nothing over the period considered.

Northern Namibian Hake

The Namibian hake data set shows more data contrast, since the catch-rate starts to increase towards the end of the time-series (Fig. 6). This results in the observation-error approach providing very precise estimates, while the process-error and the effort-averaging estimator are quite imprecise (Fig. 7). Based on the arguments of Hilborn (1979) discussed above, the process-error estimator is expected to be more precise for this series than it is for a one-way trip. While this appears to be true compared to the lobster data, it is hard to see that the hake estimates are

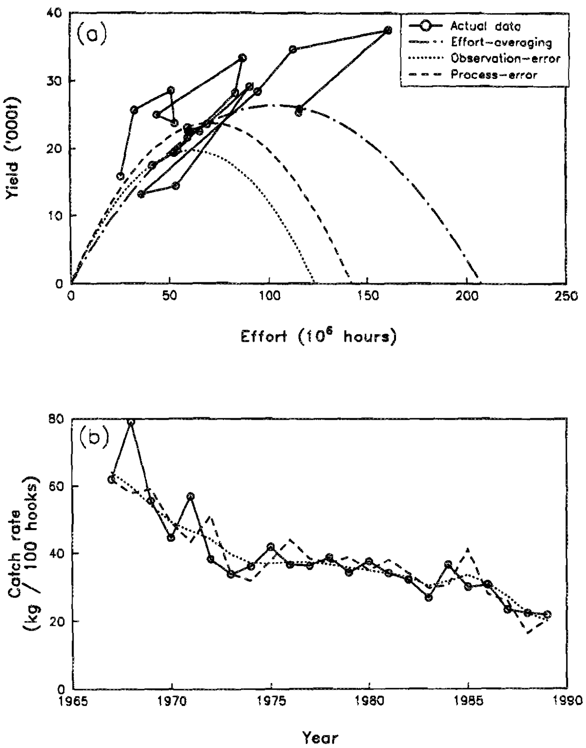

FIG. 2. (a) Catch and effort data for South Atlantic albacore with the equilibrium catch–effort curves obtained from the three estimation methods superimposed. (b) Catch-rate time series for South Atlantic albacore with the fits of the process-error and observation-error estimators superimposed.

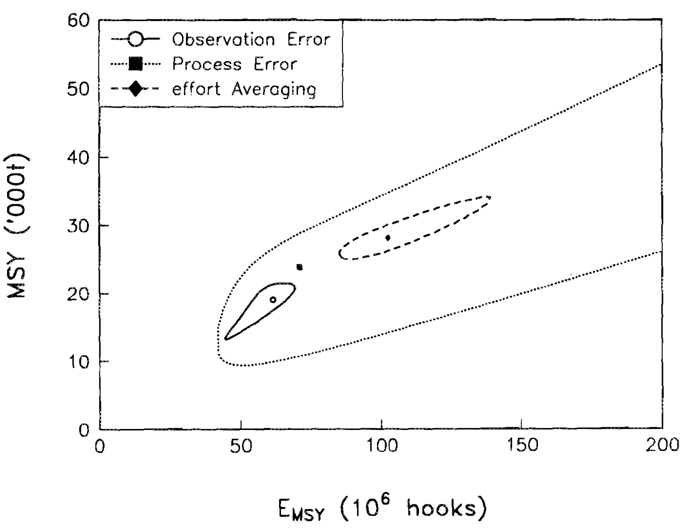

FIG. 3. Point estimates and 95% likelihood profile confidence bounds for the estimates of MSY and E_{MSY} obtained for South Atlantic albacore using the three alternative estimation methods.

more precise than the albacore estimates. The effort-averaging method has poorer precision because catch-rate is not correlated as closely with effort; the upturn in catch-rate reduces the correlation.

Qualitatively, the results for this data set are as before. The effort-averaging approach is much more optimistic than the other two approaches and the observation-error estimator is the most pessimistic. One disconcerting feature of the results for the

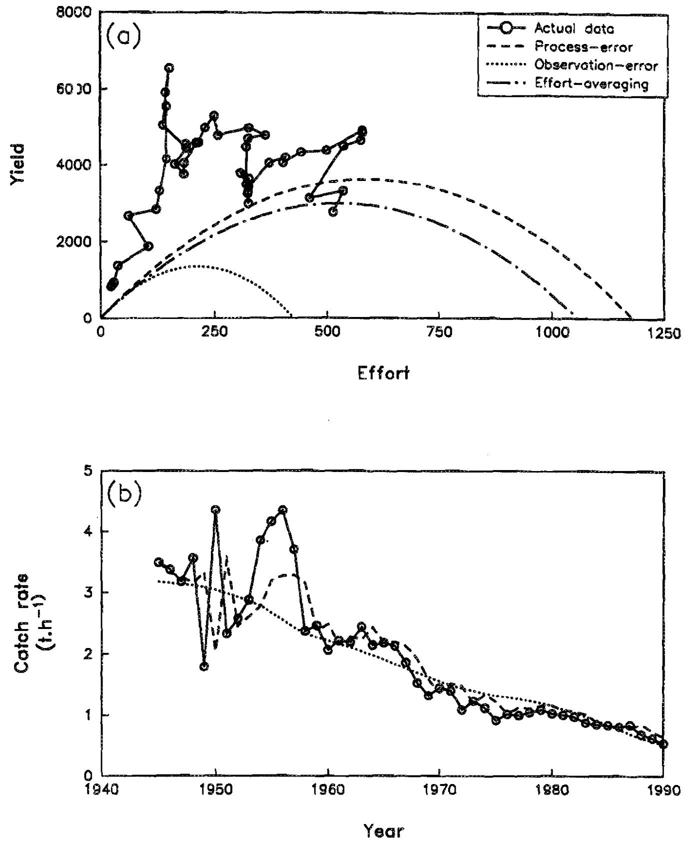

FIG. 4. (a) Catch and effort data for New Zealand rock lobster with the equilibrium catch–effort curves obtained from the three estimation methods superimposed. (b) Catch-rate time series for New Zealand rock lobster with the fits of the process-error and observation-error estimators superimposed.

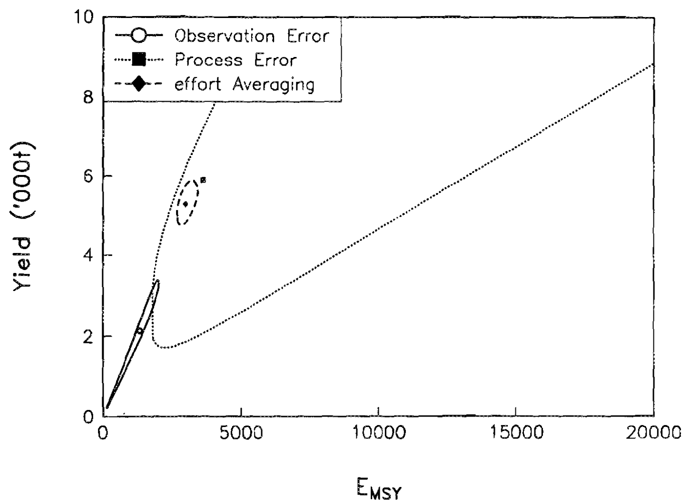

FIG. 5. Point estimates and 95% likelihood profile confidence bounds for the estimates of MSY and E_{MSY} obtained for New Zealand rock lobster using the three alternative estimation methods.

process-error estimator is that it estimates that, at the start of exploitation (1965), the resource was almost twice its carrying capacity. While it is certainly possible that due to recruitment fluctuations, the biomass at the start of exploitation may have differed from the unexploited equilibrium biomass, it seems

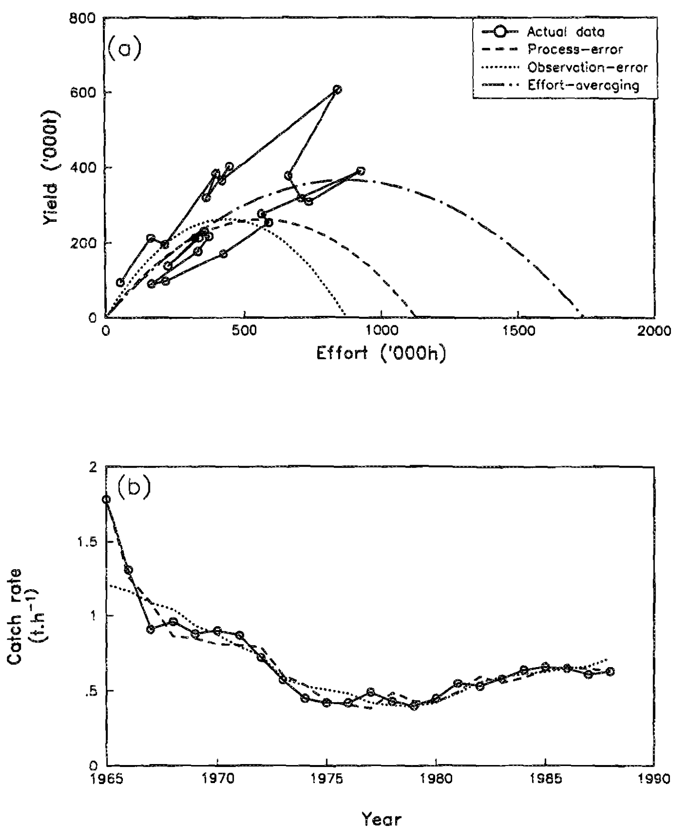

FIG. 6. (a) Catch and effort data for northern Namibian hake with the equilibrium catch-effort curves obtained from the three estimation methods superimposed. (b) Catch-rate time series for northern Namibian hake with the fits of the process-error and observation-error estimators superimposed.

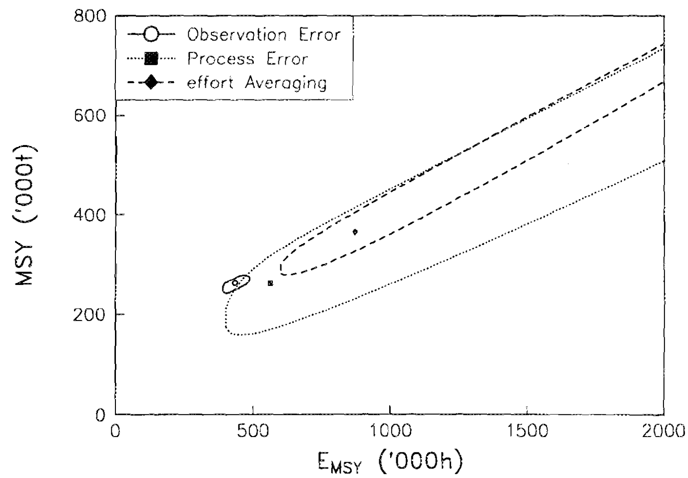

FIG. 7. Point estimates and 95% likelihood profile confidence bounds for the estimates of MSY and E_{MSY} obtained for northern Namibian hake using the three alternative estimation methods.

difficult to accept that the biomass could have been as different from K as is suggested by this assessment. The reason for this problem is, of course, the extremely high catch-rates for the initial years (see Fig. 6b). Whereas the observation-error estimator interprets this as observation-error, the process-error estimator assumes that the change in catch-rate between 1965 and 1967 reflects a real change (caused by a large process-error).

Summary of Results from Three Data Sets

It is not possible from the results of Table 2 and Fig. 2–7 to conclude which of the three estimators is best, because we do not know the true population parameters. However, certain general features can be concluded. First, the observation-error estimator provides the least optimistic appraisal while the effort-averaging method is the most optimistic. The process-error estimator is extremely imprecise. If the results of these assessments were to be pooled using, say, inverse variance weighting, the results for the process-error estimator would be given effectively no weight at all.

The extremely high precision associated with the observation-error estimator may be fallacious. The estimates of confidence bounds are naturally contingent on the model being correct. Punt and Butterworth (1993) using Monte-Carlo simulation, found that if the actual situation includes both process- and observation-error, estimates of variance obtained from the variance estimation procedures they considered were biased low by roughly 30%. Nevertheless, even if 30% is added to the confidence bounds, the results of observation-error estimator are still much tighter than those of the process-error estimator.

Monte-Carlo Testing

We were unable to conclude from the previous exercise which of the various methods was "best," primarily because it is unknown what the correct values for the management quantities of interest are. One way to overcome this problem is by testing the estimation performance of each method by means of Monte-Carlo simulation (Hilborn 1979; Uhler 1980; Ludwig and Walters 1985; Ludwig et al. 1988; Punt 1988).

Five hundred data sets were generated for each of the stocks based on the results of two of the estimation methods (process-error and observation-error estimators). It is not possible to generate data sets based on the results of the effort-averaging method because it does not estimate the biomass time-series. For the observation-error estimator, the data sets were generated by adding noise to the model-predicted catch-rate series using the equation

$$(15)\quad (C/E)_y^U = \hat{q}\hat{B}_y e^{\eta_y^U} \qquad \eta_y^U \sim N(0;\hat{\sigma}^2)$$

where $(C/E)_y^U$ is the catch-rate for year y in data set U, $\hat{q}$ is the estimate of q obtained by applying the observation-error estimator to the actual data, $\hat{B}_y$ is the estimate of biomass at the start of year y obtained by applying the observation-error estimator to the actual data, and $\hat{\sigma}$ is the standard deviation of the residuals of the fit of the observation-error estimator to the actual data.

For the process-error estimator, it would be desirable to apply the same basic procedure, namely:

$$(16)\quad I_{y+1}^U = I_y^U + \hat{r}I_y^U\left(1-\frac{I_y^U}{\hat{q}\hat{K}}\right) - \hat{q}C_y + \eta_y^U$$

$$\eta_y^U \sim N(0;\hat{\sigma}^2)$$

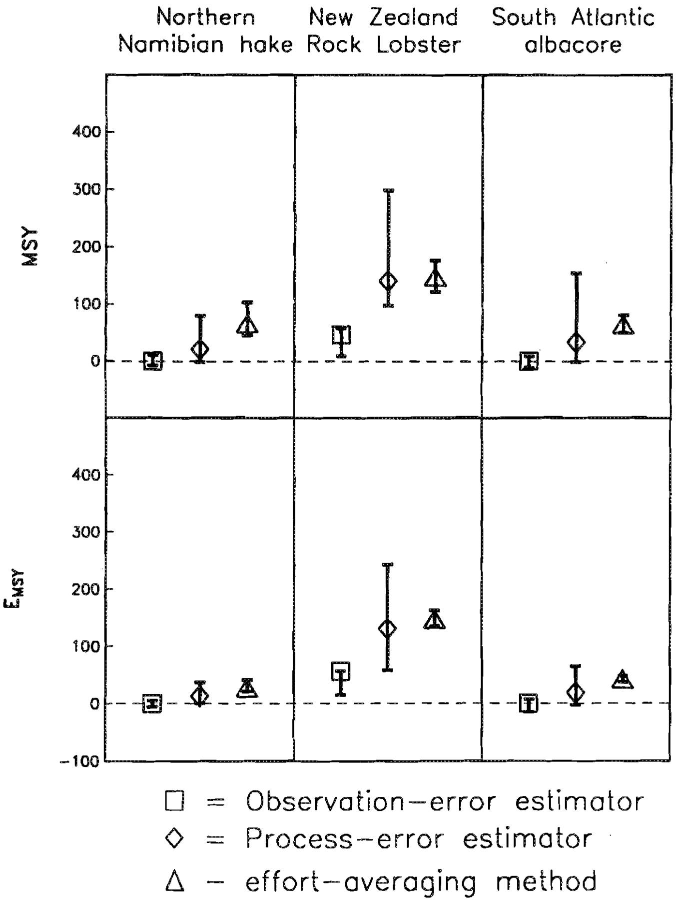

FIG. 8. Relative error distributions for MSY and E_{MSY} for the three estimation methods, for each of the three stocks, The data sets used in the calculation of this Figure were generated using the fit and assumptions of the observation-error estimator.

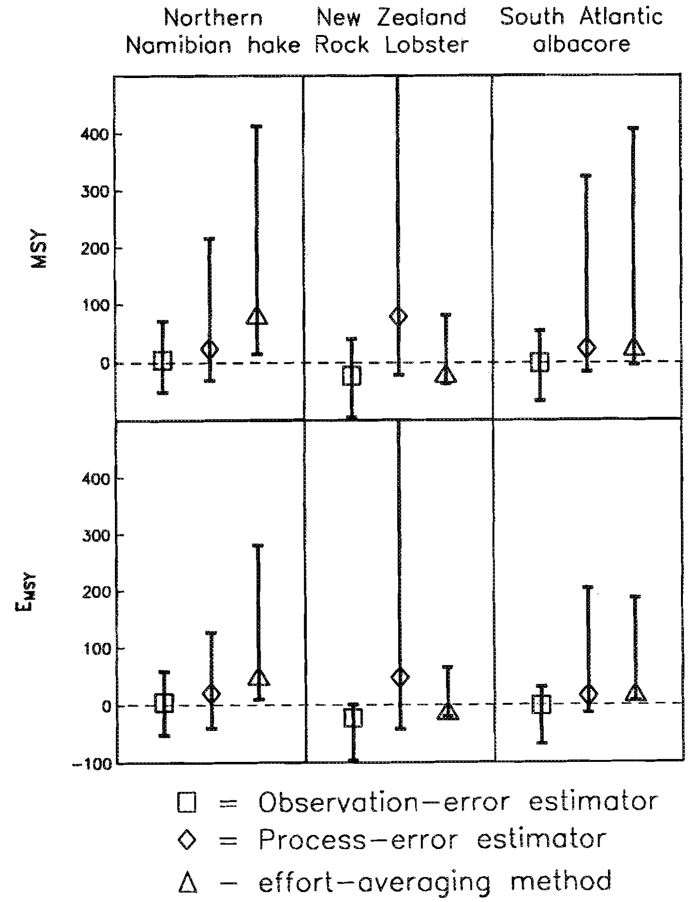

FIG. 9. Relative error distributions for MSY and E_{MSY} for the three estimation methods, for each of the three stocks. The data sets used in the calculation of this Figure were generated using the fit and assumptions of the process-error estimator. Only those data sets which corresponded to a population which was not extinct at the end of the catch time series were considered.

where $\hat{r}$, $\hat{q}$, and $\hat{K}$ are the estimates of the model parameters, and $\hat{\sigma}$ is the standard deviation of the residuals of the fit of the observation-error estimator to the actual data.

This prescription can result in population trajectories which are very different from those produced by the original assessment. This in principle, is not a problem. However, cases in which the population is rendered extinct before 1991 seem unrealistic. The obvious solution is to constrain the trajectories to those which are "similar" to the original data. This approach was used by Punt (1992b). Two cases have been considered, one in which all data sets that correspond to populations that did not go extinct before 1991 are considered, and one in which only those data sets that resulted in a current depletion which was within 10% of that reported in Table 2 are considered.

The results of this exercise (Fig. 8–10) are presented in the form of relative error distributions. The distributions are represented by their lower 5 percentiles, medians, and upper 5 percentiles.

For data sets based on the fits of the observation-error estimator (as described above), the results (Fig. 8) provide clear evidence that if the noise in Equations (1) and (2) is actually pure observation-error, then an observation-error estimator is to be preferred. The bias of the estimates provided by the observation-error estimator is only substantially larger than zero for the lobster example. The reasons for this are probably the fact that the estimate of "r" is very low for this stock (see Table 2) so that positive bias is more likely. The variance of the estimates of E_{MSY} and MSY are highest for the rock lobster example. This is a consequence of the higher value for σ used in this case (see Table 2).

As might have been anticipated, the effort-averaging approach is positively biased. This bias is lowest for Cape hake and largest for lobster. In all three cases, the estimates are fairly tight about the mean. This result suggests that it is not sufficient to perform analyses which suggest than an estimate is precise — it may nevertheless still be substantially biased. The process-error estimator is, in some respects, the poorest of the three approaches. It can be substantially positively biased (for rock lobster) and is the least precise of the three methods (as might have been expected from the results presented earlier).

The results for the process-error estimator suggest that when trajectories that go extinct are excluded (Fig. 9), we again see that the observation estimator is generally the least biased and most precise. The only exception is the lobster data set where the bias and precision of the effort-averaging method appear to be comparable to that of the observation-error estimator. In all cases the process-error estimator is both more biased and much less precise than the observation-error estimator.

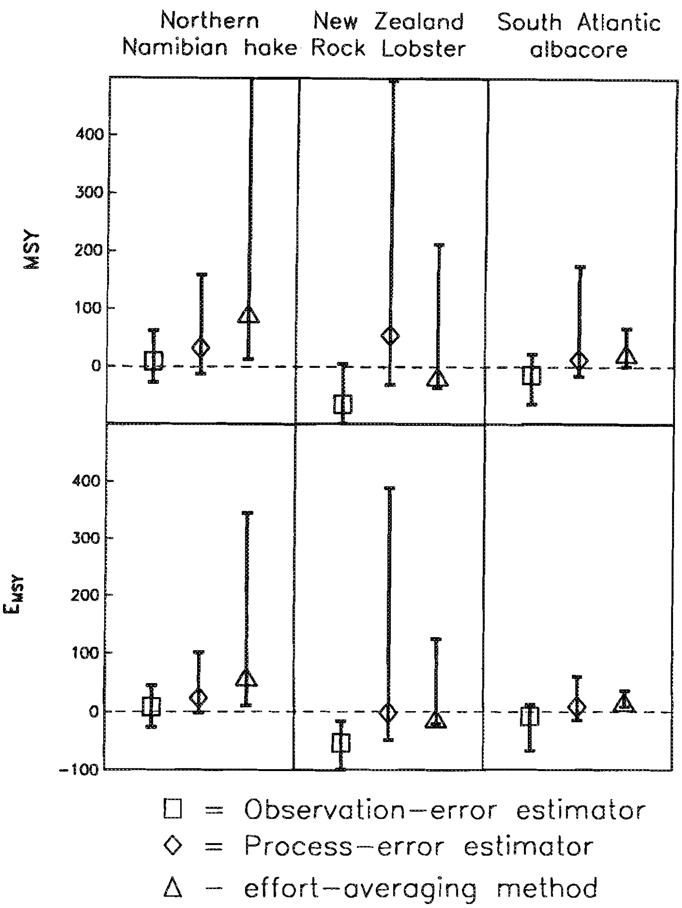

FIG. 10. Relative error distributions for MSY and E_{MSY} for the three estimation methods, for each of the three stocks. The data sets used in the calculation of this Figure were generated using the fit and assumptions of the process-error estimator. Only those data sets which corresponded to a current depletion which differed by less than 10% from the estimate provided by the process-error estimator were considered.

If we exclude all trajectories that are not similar to those for the real data (Fig. 10), the results do not change very much. The bias of the observation-error method for the rock lobster example is increased, however, and the process-error and effort-averaging methods are in fact less biased for the lobster example, although much less precise.

Discussion

The bias of the effort-averaging method, and the high variance of the process-error estimates make both of these methods extremely suspect. We believe that it is clear that the observation-error method should be expected to provide more precise and more accurate estimates of parameters, and if a single method is to be used, it should be the observation-error estimator. A rigorous assessment would include Monte-Carlo testing of alternative estimators as we have done in this paper. Such testing takes only a few hours on a desktop computer and should become standard practice.

Under no circumstances should agency staff, conference organizers, reviewers, managers or journal editors accept assessments or publications that are based on effort-averaging or process-error estimators only. The evidence presented in this paper, combined with previously published papers (Punt 1988; Punt 1992b; Hilborn and Walters 1992 etc.) make it clear that for most fisheries data sets observation-error estimators are superior.

Acknowledgments

This research was supported in part by the South Pacific Commission, University of Washington, the U.S. National Marine Fisheries Service, the University of Cape Town, and CSIRO Division of Fisheries Research. We thank several anonymous reviewers for comments, and Ulrike Hilborn for editorial assistance in the preparation of the manuscript.

References

BARTOO, N., AND T. SHIOHAMA. 1985. A production model analysis of North Pacific albacore population including estimates of the sensitivity of results to measurement errors in input data. Bull. Far Seas Res. Lab. No. 22:109–116.

BERGH, M.O., AND S.J. JOHNSTON. 1992. A size-structured model for renewable resource management with application to resources, of rock lobster in the south-east Atlantic, p. 1005–1016. *In* A.I.L. Payne, K.H. Brink, K.H. Mann, and R. Hilborn [ed.] Benguela Tropic Functioning. S. Afr. J. Mar. Sci. 12.

BREEN, P.A. 1991. Assessment of the red rock lobster (*Jasus edwardsii*) North and South Island stock, November 1991. New Zealand Fisheries Assessment Research Document 91/16. 36 p. Document available at Ministry of Agriculture and Fisheries Library, Greta Point, Wellington, N.Z.

BUTTERWORTH, D.S., AND P.A. ANDREW. 1984. Dynamic catch–effort models for the hake stocks in ICSEAF Divisions 1.3 to 2.2. Collect. Sci. Pap. Int. Comm. SE Atl. Fish. 11(1): 29–58.

BUTTERWORTH, D.S., AND P.A. ANDREW. 1987. On the appropriateness of approaches used at ICSEAF to obtain TAC estimates from catch–effort data for hake. Collect. Sci. Pap. Int. Comm. SE Atl. Fish. 14(1): 161–191.

DERISO, R.B., T.J. QUINN, AND P.R. NEAL. 1985. Catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. 42: 815–824.

DRAPER, N.R., AND H. SMITH. 1966. Applied regression analysis. John-Wiley, New York, N.Y. 709 p.

EFRON, B. 1982. The jackknife, the bootstrap and other resampling plans. Society for Industrial and Applied Mathematics, Philadelphia, PA. 92 p.

FLETCHER, R.I. 1978. Time-dependent solutions and efficient parameters for stock-production models. U.S. Fish. Bull. 76: 377–388.

FOX, W.W. 1970. An exponential surplus-yield model for optimizing exploited fish populations. Trans. Am. Fish. Soc. 99: 80–88.

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

GAVARIS, S. 1988. An adaptive framework for the estimation of population size. Can. Atl. Fish Sci. Adv. Comm. Res. Doc 88/29: 12 p.

GULLAND, J.A. 1961. Fishing and the stocks of fish at Iceland. Fish. Invest. Minst. Agric. Fish. Food, U.K. (Ser. 2) 23(4): 52 p.

GULLAND, J.A. 1983. Fish stock assessment: a manual of basic methods. John Wiley, New York, N.Y. 223 p.

HILBORN, R. 1979. Comparison of fisheries control systems that utilize catch and effort data. J. Fish. Res. Board Can. 36: 1477–1489.

HILBORN, R., AND C.J. WALTERS. 1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Chapman and Hall, New York, N.Y. 572 p.

LLEONART, J., AND J. SALAT. 1989. A dynamic approach to catch-effort data analysis using a new concept: the inertia of the stock. Can. J. Fish Aquat. Sci. 46: 1367–1374.

LLEONART, J., J. SALAT, AND B. ROEL. 1985. A dynamic production model. Collect. Sci. Pap. Int. Comm. SE Atl. Fish 12(1): 119–146.

LUDWIG, D., AND C.J. WALTERS. 1985. Are age-structured models appropriate for catch-effort data? Can. J. Fish. Aquat. Sci. 42: 1066–1072.

LUDWIG, D., C.J. WALTERS, AND J.G. COOKE. 1988. Comparison of two models and two estimation methods for catch and effort data. Natural Resource Modelling 2(3): 457–498.

PELLA, J.J., AND P.K. TOMLINSON. 1969. A generalized stock production model. Bull. Inter-Am. Trop. Tuna Comm. 13: 421–458.

POPE, J.G., AND J.G. SHEPHERD. 1985. A comparison of the performance of various methods for tuning VPAs using effort data. J. Cons. Perm. Int. Explor. Mer 42: 129–151.

PRESS, W.H., B.P. FLANNERY, S.A. TEUKOLSKY, AND W.T. VETTERLING. 1986. Numerical recipes: the art of scientific computing. Cambridge University Press, Cambridge. 818 p.

PUNT, A.E. 1988. Model selection for the dynamics of Southern African hake resources. Rep. Benguela Ecol. Progm. S. Afr. 15: [vii] + 395 p.

PUNT, A.E. 1990. Is $B_1 = K$ an appropriate assumption when applying an observation-error production model estimator to catch-effort data? S. Afr. J. Mar. Sci. 9: 249–259.

PUNT, A.E. 1992a. Selecting management methodologies for marine resources, with an illustration for southern African hake. S. Afr. J. Mar. Sci. 12: 943–958.

PUNT, A.E. 1992b. Management procedures for Cape hake and baleen whale resources. Rep. Benguela Ecol. Progm. S. Afr. 23: [viii] + 689 p.

PUNT, A.E., AND D.S. BUTTERWORTH. 1989. Application of an ad hoc tuned VPA assessment procedure to the Cape hake stocks in the ICSEAF Convention area. ICSEAF Document SAC/89/S.P./28: 54 p.

PUNT, A.E., AND D.S. BUTTERWORTH. 1993. Variance estimates for fisheries assessment: their importance and how best to evaluate them, p. 145–162. *In* S.J. Smith, J.J. Hunt, and D. Rivard [ed.] Risk evaluation and biological reference points for fisheries management. Can. J. Fish. Aquat. Sci. Spec. Publ. 120.

RICKER, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Bull. Fish. Res. Board Can. 191: 382 p.

RIVARD, D., AND J. BLEDSOE. 1978. Parameter estimation for the Pella–Tomlinson stock production model under nonequilibrium conditions. U.S. Fish. Bull. 76: 523–534.

SCHAEFER, M.B. 1954. Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. Bull. Inter-Am. Trop. Tuna Comm. 1: 25–56.

SCHNUTE, J. 1977. Improved estimates from the Schaefer production model: theoretical considerations. J. Fish. Res. Board. Can. 34: 583–603.

SPARRE, P., E. URSIN, AND S.C. VENEMA. 1989. Introduction to tropical fish stock assessment, Part 1 — Manual. FAO Fish. Tech. Pap. 306/1: 429 p.

UHLER, R.S. 1980. Least squares regression estimates of the Schaefer production model: some Monte Carlo simulation results. Can. J. Fish. Aquat. Sci. 37: 1284–1294.

VENZON, D.J., AND S.H. MOOLGAVKOR. 1988. A method for computing profile-likelihood-based confidence intervals. Appl. Stat. 37: 87–94.

WALTERS, C.J., AND R. HILBORN. 1976. Adaptive control of fishing systems. J. Fish. Res. Board Can. 33: 145–159.

YEH, S-Y., T-S. TSOU, AND H-C. LIU. 1991. Assessment of the south Atlantic albacore resource by using surplus production models, 1976–1988. Colln. Vol. Sci. Pap. ICCAT. 39: 166–170.