D. PRODUCTION MODELING

PURPOSE

The purpose of this report is to summarize the working papers and discussions of the Methods Working Group concerning the terms of reference related to surplus production models. Each major topic is introduced and accompanied by a brief description of the methodology and one or more example applications. The various approaches are intended to be illustrative, rather than definitive, and should not be construed as revisions to existing assessments or biological reference points. The comments of the SARC on the various approaches are summarized at the end of this section.

INTRODUCTION

Surplus production models play a central role in the management of fisheries under the Sustainable Fisheries Act (SFA) of 1996 (USDC 1997). The SFA provides explicit guidance on the definition of maximum sustainable yield, fishing mortality and biomass targets and thresholds, and time frames for rebuilding of overfished stocks. This guidance has challenged stock assessment scientists to develop estimates of such parameters for a broad range of species. In this report we focus on the challenges relevant to assessments of fishery resources in the northeast region of the United States. Fishery resources in the northeast benefit from a long time series of synoptic survey data. For over 30 years the Northeast Fisheries Science Center has conducted two annual surveys of finfish populations. Since 1992 surveys have been conducted in the winter, and specialized surveys for shellfish are also conducted.

In spite of this wealth of data, estimation of parameters in surplus production models is difficult for many stocks. The difficulties stem from several sources. First, many of the stocks have been heavily fished for almost 100 years. Thus any simplifying assumptions about the state of resource when the surveys began are tenuous. Second, the fisheries are prosecuted by a wide range of fleets and gear types with varying levels of selectivity. Many of these fisheries generate substantial quantities of discards, which in many instances are poorly estimated. Third, many stocks have been subject to intense fishing mortality, first by foreign fleets before 200 mile coastal limits were imposed, and then by overcapitalized domestic fleets. Rebuilding of fish stocks, to the extent it has occurred, has been limited to the past 5-10 years. In terms of surplus production models, these conditions imply that one cannot assume that the initial population size in near the carrying capacity. Heterogeneous fisheries imply that not all removals are known. Many stocks have experienced consistent declines in abundance (the “one-way trip”) such that information about density dependent processes is difficult to discern. In aggregate, these conditions create an ironic circumstance in which a wealth of age-specific survey and catch data contains little information about density-dependent processes.

Nonetheless, models of surplus production have been developed for several species in the Northeast and the resulting parameters have been codified into control rules for fisheries management. Most of the major fish stocks are assessed using age-structured models, especially Virtual Population Analyses “tuned” to survey data. The tuned-VPAs create a further difficulty, in that the vector of derived fishing mortality rate (F) must be translated into an average F comparable to estimates derived from surplus production models. This melding of two model constructs is likely to be a transient condition as more synthetic models are developed. However in the short run it is important that current assessment results be compared to existing definitions of overfishing, however they have been defined.

The Methods Subcommittee of the SARC was asked to provide guidance on the use of surplus production models. The specific terms of reference are listed below:

TERMS OF REFERENCE

(A) Evaluate the use of production models in providing estimates of biomass and yield targets and thresholds consistent with provisions of SFA

(B) Provide guidance on the use and limitations of production model results for establishing management goals

(C) Evaluate various types of production models (age/stage structured, non-equilibrium, etc.) and provide guidance on the use of model types in differing circumstances of data availability, exploitation history and length of time series.

(D) Compare estimates of MSY, Fmsy and Bmsy from production models with those based on catch-at-age model results as a basis for understanding biases, stability and precision of such estimated parameters.

The Methods Subcommittee attempted to address these topics by examining a number of case studies. The approaches taken by the group can be classified into the following:

1) Diagnostic measures–is there evidence to support the underlying processes of density dependence?

2) Commensurate quantities–can we develop internally consistent measures of biomass and fishing mortality from age-structured and production models?

3) Advanced estimation procedures–are there more advanced estimation procedures which improve the accuracy and precision of biological reference points?

4) Identification of promising research areas–especially those related to integrated assessment approaches.

GRAPHICAL AND DIAGNOSTIC METHODS

Design Sufficiency

One of the major difficulties for Northeast fisheries is the determination of biomass targets under SFA. As noted above, high exploitation for many species occurred prior to the primary time series of fishery independent data. This can mean that the biomass that supports MSY has not been observed in the extant data. Restricting inferences of Bmsy to what has been observed may be myopic; extending estimates beyond the range of the data can be tenuous. The realized time series of catches and survey values can be considered as an outcome of an unplanned experiment. The same properties that constrain inferences in experimental design are also important for model based estimation. In particular, the number of observations that occur at each level of treatment factors is a critical factor in experimental design. In fisheries the relative proportion of observations at high and low levels of population level are important. Let population density be proportional to relative survey biomass and exploitation be proportional to catch divided by survey biomass. The relative frequency of observation in each cell provides an indication of the ability to estimate population parameters. A hypothetical example:

	Population Size	
Relative Harvest Rate	Low	High
Low	0.1	0.2
High	0.5	0.2

Note that most of the observations are at high harvest rates and low population size. An even more disturbing pattern will be evident for a heavily exploited stock.

	Population Size		
Relative Harvest Rate	Low	Medium	High
Low	0.05 (5)	0	0
Medium	0.25 (4)	0.1 (1)	0
High	0.4 (3)	0.2 (2)	0

The first number indicates the fraction of observations at each population state. The second number in parentheses indicates the order of observation. Hence the population may never have been monitored at high levels of abundance and the sequence of "treatments" has clearly not been randomized. In an experimental design, such conditions would merit at least a split plot design. In fisheries the analytical solution is not clear, but the warning message is the same. Analytical sophistication may not be sufficient to overcome fundamental problems of inference.

The Envelope Plot

To gain some insight into possible biomass targets it is useful to compute historical measures of abundance under a range of assumptions. For example, a catch series C_t can be used to create a range of possible population biomasses B_t by noting that $B_t = C_t / U$ where U is exploitation rate. If we assume that the catches are the realization of a consistently low exploitation rate then $B_{t(high)} = C_t / U_{low}$. Conversely, if exploitation rates have been consistently high then an ultimate lower bound on exploitable biomass is simply the catch series, i.e., $B_{t(low)} = C_t$. If a time series of survey data are available, swept area biomass estimates can be computed for varying levels of catchability or gear efficiency. Other model-based estimates of abundance, say from a VPA, can be superimposed on the same graph. Finally, estimates of biomass per recruit from standard YPR analyses can be multiplied by average recruitment to generate an estimate of expected population biomass. The resulting series of population estimates can be considered an envelope of feasible population sizes in which surplus production-based estimates of Bmsy should at least have some consistency. An example plot for summer flounder is depicted in Fig. 1 in which the high estimates of population size are based on an exploitation rate ~0.15.

Model Behavior Plots

The standard surplus production model can be written as a difference equation

$$B_{t+1} = B_t + rB_t - \frac{r}{K} B_t^2 - C_t \qquad (1)$$

If survey index I_t is proportional to B_t such that $I_t = q\,B_t$, then Eq. 1 can be written as

$$I_{t+1} = I_t + rI_t - \frac{r}{qK} I_t^2 \ - qC_t \quad (2)$$

Rearranging terms and dividing though by I_t

$$\frac{I_{t+1} - I_t}{I_t} \ = +rI_t - \frac{r}{qK} I_t \ \ - q\frac{C_t}{I_t} \quad (3)$$

If density dependence is evident in a population, then the rate of increase in relative population size should increase at low population levels and decrease at high population densities. A simple test of this concept is depicted in Fig. 2 in which relative population abundance at time t+1 is plotted against relative harvest rate for summer flounder. The size of the circle is proportional to the magnitude of the quantity $(I_{t+1} - I_t) / I_t$; open circles indicate negative values, filled circles represent positive values. Each point is labeled with the year of observation and time trend is denoted by the line. Figure 2 demonstrates the ongoing recovery of summer flounder as relative harvest rates appear to be decreasing. More importantly, however, the plot suggests that the rate of increase (i.e., circle diameter) at high density and low harvest rate is comparable to those values observed at low densities and high harvest rates. From this plot at least, there does not appear to be any evidence of density dependent reduction in biomass increase. The observed trajectory of summer flounder may in fact reflect the transient effects of several yearclasses surviving for more than a few years in the fishery. Such evidence suggests a priori that the K parameter (and hence B_{msy}) may be difficult to estimate for summer flounder.

<u>When Are Multiple Indices Useful?</u>

It is often assumed that multiple indices will improve model fits by using more information. This is true however, only if the indices are measuring the same attribute of the population in a given spatial domain. If so, information on population abundance will be improved by having multiple measures. In most assessment models, conflicting data trends are accommodated by “splitting the difference”. “Splitting the difference” may not be useful if the conflicting observations signal changes in the underlying process (e.g. shifts among spatial units), temporal changes in availability, or changes in the underlying harvest process. As an example of the latter process, consider a management measure that changes the seasonality of harvesting. The relationship between catch and index values will then change over time but the model will accommodate this change as an error term to be minimized.

A simple plot of the spring and fall indices for summer flounder (Fig. 3) suggests that the fall index has been increasing more rapidly than the spring survey since about 1990. A plot of the spring survey in the subsequent spring (t+1) against the fall index in year t suggests that the slope is decreasing. This figure is consistent with the hypothesis that an increasing fraction of the landings are occurring in the winter between the fall and spring surveys. The simple surplus production model with multiple indices cannot accommodate this change since each index is assumed to be representative of the population biomass. Moreover, catch is incorporated as total annual catch rather than temporally disaggregated values.

Response Surface Plots: Graphical Measures of Uncertainty

The sampling covariance between r and K (and hence F_{msy} and B_{msy}) in surplus production has been well studied in the literature. The nonlinear negative association can be particularly severe if the model does not fit particularly well. In these circumstances, both of the primary estimates of interest to management may be useless. Regardless of the degree of fit, it is clear that traditional measures of precision, based on asymptotic properties, are likely to underestimate the true variation. Bootstrap procedures address this issue in part and should be a component of any serious attempt to estimate population parameters. Additional insights can be gained by examining the loss function in the vicinity of the solution and by applying confidence intervals procedures more appropriate for nonlinear models.

To begin this examination, it is necessary to reparameterize the surplus production model in terms of F_{msy} and B_{msy}. This is accomplished by substituting the functional relationships r=2 F_{msy} and K=2 B_{msy} into Equation 2.

$$I_{t+1} = (1+2F_{MSY})I_t - \frac{F_{MSY}}{qB_{MSY}} I_t - qC_t \quad (4)$$

Eq. 4 permits one to immediately compute the primary parameters of interest, an advantage in some statistical packages. Comparisons of parameter estimates obtained by Eq. 2 and 4 were identical (thereby providing empirical evidence of the invariance principle of maximum likelihood estimators!).

In contrast to standard Wald-type estimators of confidence intervals , Cook-Weisberg (C-W) method (Cook and Weisberg 1990) is specifically designed for nonlinear models. The C-W method is conceptually similar to profile likelihood methods since the model is re-estimated for each alternative fixed value of the variable of interest. For example, in Eq. 4, the confidence interval for Fmsy is estimated by recomputing the best estimates of Bmsy and q for each fixed value of Fmsy in the vicinity of the solution. The residual sum of squares is asymptotically distributed as a t-statistic; in a profile likelihood approach the likelihood function would have a χ^2 distribution.

Approximate confidence regions for each parameter can be simply examined by evaluating the RSS in the vicinity of the solution. The C-W method was not applied to the confidence region; instead, the "significance level" was approximated with an F statistic, following the methods in Draper and Smith (1966) (See Fig 4a).

Results of the modified surplus production model fit are summarized in Table 2. The spring and fall survey indices were simply averaged for this heuristic example. For this model configuration, the estimated value of F_{msy} = 0.4 and the B_{msy} level is 59,268 mt. It is worth repeating –these values are used for illustration only.

Contour plots of the loss function for all possible pairings of F_{msy}, B_{msy} and q (not shown) demonstrated a wide range of values for even the nominal significance levels. One example (Fig. 4b) of the B_{msy} vs F_{msy} contour plot may be of general utility for development of uncertainty in control rules.

Funnel Plots—Evaluating the Value of Additional Data

As many authors have noted, long time series of catch data are not necessarily informative about underlying population dynamics in surplus production models. The surplus production model does not exhibit the convergence properties of VPAs and additional data may not improve the precision of estimates. On the contrary, additional data, especially if it is informative, may markedly alter one's perception of the population's dynamics. In principle, a data set derived from a population following a logistic growth model and subject to variations in harvest rates at different stock levels, should be sufficient to recover the underlying parameters. As the length of the time series increases, the estimates should converge to stable estimates of these parameters. Moreover, these parameters should be recoverable from series of any length and any starting point.

These concepts were merged to estimate a set of parameter estimates corresponding the enumeration of all possible series of length s from an initial series of length n. In more mathematical terms, let Θ_{sj} represent a vector of parameters corresponding to the j-th series on length s. For example the series can be enumerated as {j=1; t=1,2, ...s}, {j=2; t=2, 3, ...s+1}, ... {j=k; t=k, k+1, ...n}. This can be done for all series of length s up to n. The corresponding estimates can be displayed as a function of the number of contiguous points in the data set. These can be called funnel plots based on an expected shape. Series with fewer elements might be expected to exhibit greater variation, with a narrower range of estimates at the number of elements approaches the original number of observations. An example set of funnel plots for summer flounder is depicted in Fig. 5. The left column shows the set of estimates circumscribed by a convex hull. The right column shows a box plot of the estimates with a Lowess smooth through the data points. The plots suggest that the surplus production model parameters are not stable since removal of a small number of data points induces wide variations in estimates. The apparent trend in increasing values of F_{msy} and q is also undesirable. Similar concerns were noted by Terceiro (2001) who conducted a retrospective analysis. The funnel plot simply enumerates all possible retrospective patterns and reinforces Terceiro's concerns. As the number of data points in the series decreased, the number of estimation failures (i.e., no convergence) increased. For the shortest length series (m=13), over 35% of the runs failed to converge. Failure rates did not fall below 30% until at least 19 points were included in the time series.

Collectively, the graphical methods proposed herein should be viewed as complementary to existing approaches to derivation of suitable surplus production models. Traditional residual analyses are useful, but many features may not be discernible if the model fitting process masks changes in the underlying process. While it may not be possible to develop a formal proof, it seems logical to assert that the problems of model mis-specification are likely to be more pronounced in simple models. Therefore, considerable caution should be applied when attempting to derive biological reference points from surplus production models.

Use of Smoothed Indices in Surplus Production Models

Modern smoothing methods are an important tool for stock assessment but in the context of modeling methods that include catch, considerable caution is necessary. A simple example will suffice to illustrate the difficulties of interpretation. As before, let B_t represent the population biomass at time t and B'_t represent a simple moving average of B_t centered on time t. For a simple 3 point moving average $B'_t = (B_{t+1} + B_t + B_{t-1})/3$. If P_t denotes the surplus production at time t then

$$P_t = B_{t+1} - B_t + \delta C_t . \quad (5)$$

If B_t is replaced by its moving average then

$$P'_t = (B_{t+2} + B_{t+1} + B_t)/3 - (B_{t+1} + B_t + B_{t-1})/3 + \delta C_t$$

or

$$P'_t = (B_{t+2} - B_{t-1})/3 + \delta C_t \quad (6)$$

Thus the production in year t is written as function of catch in the current year, biomass in the previous year and biomass two years in the future. As the duration of the moving average period increases, the e discounting of the terminal points would become even smaller such that $_{Prt}$–>C_t.

Without additional smoothing of the catch series, the mechanisms that might make the above equation meaningful are unclear. If a more complicated n-point smoothing algorithm was applied, then the smoothed estimate of production in year t would be represented as a linear combination of n+1 biomass levels. Once again, it may be difficult to interpret such equations.

RELATION BETWEEN MSY AND AVERAGE CATCH

The subcommittee also addressed the issue of the expected relationship between estimates of MSY and average catch. Many have noted that MSY is often close to estimates of average catch. It can be shown that a lower bound on MSY can be written as

$$MSY \geq (K\overline{P} + \overline{C})\left(\frac{r}{1+r}\right)^2$$

where $\overline{P}$ is the average fraction of the population present and $\overline{C}$ is the average catch. Unfortunately, it is not possible to develop an upper bound on MSY from the catch series. Thus, the potential for huge MSY values persists as long as there is no direct evidence of density dependence in the time series.

BIOMASS WEIGHTED F-EFFECTS OF TRANSIENT CONDITIONS

Theory

Surplus production models (SPM) treat biomass as an undifferentiated pool in which each unit of biomass has an equal capacity for reproduction, growth and mortality. In contrast, age-structured models (ASM) admit differences in the properties with respect to age. Stochastic variations in recruitment and their subsequent effects on biomass production are subsumed into estimates of r in SPM. The transient effects of recruitment complicate the translation of vector-based Fs in ASM to scalar-based Fs in SPM.

One simplification that identifies the nature of the problem is to note that prediction of yield from an undifferentiated biomass pool is equivalent to that in the age-structured model.

Under the surplus production model $Y = F_{SPM}\overline{B}_{TOT}$. Under the age-structured model $Y = \sum_{i=1}^{A} F_i \overline{B}_i$. Combining these equations for yield and noting that $B_{TOT} = \Sigma \overline{B}_i$ leads to

$$F_{SPM} = \frac{\sum_{i=1}^{A} F_i \overline{B}_i}{\sum_{i=1}^{A} \overline{B}_i}$$

This implies that the biomass weighted F from an ASM is equivalent to the pooled F from a surplus production model (SPM).

It is important to note however, that the variations in age-specific biomass are induced by variations in the numbers of recruits associated with each cohort and their fishing history. Both factors will cause deviations from the weighting factors associated with constant recruitment and fishing history. A hypothetical age structure, based on the contemporary set of age specific Fs and a constant recruitment can be used to compare the magnitude of deviations in the current age structure.

Let the vector $\underline{F}_{ASM}(t)$ represent the estimated age-specific Fs in year t from an ASM. The expected number at age that would obtain under $\underline{F}_{ASM}(t)$ and constant recruitment R can be estimated as

$$N_{EQ,i} = \mathrm{Re}^{-\sum_{j=0}^{i-1} F_{ASM,j} + M}$$

The expected equilibrium biomass at age can be estimated as $B_{EQi} = N_i \overline{W}_i$ where $\overline{W}_i$ is the average weight at age i. The corresponding biomass-weighted F associated with equilibrium recruitment and F_{ASM} is

$$F_{EQ} = \frac{\sum_{i=1}^{A} F_{ASM,i} \overline{B}_{EQ,i}}{\sum_{i=1}^{A} \overline{B}_{EQ,i}}$$

If we denote the observed age-specific biomass estimates a as the difference between the biomass-weighted F and the equilibrium F i.e., F_{BW} - F_{EQ} can now be examined in terms of its departure from equilibrium conditions. Not that the differences in F are independent of the absolute magnitude of recruitment R and depend only on the vector F and average weights. The differences between F_{BW} and F_{EQ} can be decomposed into deviations associated with non- equilibrium conditions. For the vector difference B_{OBS} - B_{EQ} , positive values are indicative of either lower historical F or higher recruitment; negative values reflect the opposite.

<u>Application</u>

To illustrate the technique, the above equations were applied to an earlier ADAPT version of the Gulf of Maine cod. Ages 1 to 7+ were used in the VPA. The equilibrium estimate of population biomass in the plus group was estimated by extending the population age vector out to 25 years, and retaining the same age weight as employed in the VPA. The observed biomass weighted F from the VPA (i.e., F_{ASM}) = 0.2113 whereas the biomass weighted F under equilibrium conditions was 0.2296. Comparison of the observed and expected biomasses at age suggest that the largest disparity for age 2 (1998 year class) accounts for about 80% of the total deviation.

Discussion
The vector-based approach may be useful for characterizing the transient effects of non-equilibrium age structure on the derived biomass-weighted F. The prediction of an equilibrium biomass structure that would obtain under the observed F vector in the terminal year permits an analysis of how far the current age structure is from equilibrium. The total difference in average F can be computed and the age-specific contributions to the difference can be estimated.

As a discussion point, it could be argued that F_{EQ} is a better "point of entry " for fishery control rules based on surplus production models. By extension, it may also be argued that a total biomass estimate, derived as B^{T}_{EQ} .1, might be appropriate for the biomass axis of the control rule. In either case, the need to translate F's derived from age structured models into their surplus production equivalents (e.g., see Applegate et al. 1998), is a short-term problem that should be resolved as better estimates of biological reference points become available.

EXTERNAL SURPLUS PRODUCTION MODELS

Methodology
Annual surplus production in an unfished stock is defined as P_t=B_{t+1}-B_t (Ricker 1975). When fishing mortality is considered, surplus production is defined as

$$P_t = B_{t+1} - B_t + \delta C_t \quad (7)$$

where δ is a correction factor that adjusts biomass at the beginning of year t+1 for catch during year t. The factor δ accounts for surplus-production by fish taken in the fishery during year t so that the sum $B_{t+1} + \delta C_t$ is the hypothetical biomass that would have existed in year t+1 if there had been no fishing during year t (MacCall 1978). We assumed δ = 1 for all stocks in this analysis. This assumption is valid when the instantaneous rate of natural mortality (M) and average instantaneous somatic growth rate ($\overline{G}$) balance (i.e. where M-$\overline{G}$ is approximately zero for ages taken in the fishery).

Surplus production as defined in Eq. 7 can be estimated for any model that generates a time series of biomass estimates. Such estimates of production are useful in characterizing the response of populations to exploitation and investigating temporal trends. It is also possible to examine the degree to which productivity estimates agree with predictions of surplus productions models. This is accomplished by fitting a quadratic function (Schaefer 1957) to the estimated production estimates such that substituting biomass estimates ($\hat{B}_t$) from the "best available" stock assessment model for biomass (B_t) into Eq. (7), gives:

$$\hat{P}_t = \hat{B}_{t+1} - \hat{B}_t + \delta C_t = a\hat{B}_t + b\hat{B}_t^2 \quad (8)$$

where $\hat{P}_t$ is an "observed" estimate of P_t used as the best available "data" in externally estimated surplus production models. The fitted model of estimate surplus production can be written as

$$\tilde{P}_t = \tilde{a}\hat{B}_t + \tilde{b}\hat{B}_t^2 \quad (9)$$

Thus $\tilde{P}_t$ is the estimate of surplus production based on the zero intercept quadratic model that relies on biomass estimates derived from another model. Estimates of the $\tilde{a}$ and $\tilde{b}$ parameters can be used to derive estimates of the intrinsic rate of increase ($\hat{r} = \tilde{a}$) and carrying capacity ($\hat{K} = -\tilde{a}/\tilde{b}$). Other standard alegebraic deductions of Schaefer's model also follow such at $B_{MSY} = K/2$ (where B_{MSY} is the equilibrium biomass for MSY), MSY = $aK/4$ and $F_{MSY} = a/2$.

The use of the expression "external" reflects the fact that estimates of a and b are not incorporated into the original estimates of B_t. The external approach is a special case in a general family of internally estimated "composite" non-equilibrium surplus production models (Fournier and Warburton 1989), that also includes conventional all measurement error such as ASPIC (Prager 1994) as another special case. Additional details on the estimation and application of external estimates of surplus production parameters may be found in Jacobson et al. (in press). The following examples rely heavily on the methodology presented in Jacobson et al.

As noted earlier, F_{MSY} and B_{MSY} are often correlated and the b parameter in Eq. 9 may be difficult to estimate for heavily fished stocks with few data at high biomass levels. This circumstance implies that F_{MSY} is estimable but B_{MSY} is not. To test for this circumstance we fit the model with and without the quadratic term, and used t-tests (p=0.05) to determine which model was "better" for the available data. One-sided t-tests were used because the expected values of b is less than zero in Schaefer's model (Eq. 4).

The statistical model we used to fit external production models with independent errors was:

$$\hat{P}_t = \tilde{P}_t + \varepsilon_t \qquad (5)$$

where ε_t is an independent statistical error term that includes both measurement and process errors. When statistical errors were assumed to be autocorrelated, we used:

$$\hat{P}_t = \tilde{P}_t + \gamma_t = \tilde{P}_t + \varepsilon_t + \sum_{j=1}^{MaxLag} \lambda_{t-j}\varepsilon_{t-j} \qquad (6)$$

where the moving average parameters l_{t-j} Î (-1, 1) were for lags of 1-3 years, the simple residual $\gamma = \hat{P}_t - \tilde{P}_t$ was autocorrelated, and the time series residuals ε_t were independent. The moving average approach is easy to use and effective but it requires estimation of additional moving average parameters l_{t-i} (Schnute 1985). In theory, the independent errors prior to the first year (ε_{1-t}) should be estimated as well. For simplicity, we assumed that all independent errors prior to the first year were zero, and we restricted our models to lags £ 3 years.

We used t-tests for parameter estimates to help us determine how many moving average parameters were required in the best model for a particular data set (Schnute 1985). Residual patterns were another factor that considered in making these decisions.

Objective functions used in external surplus production model parameter estimation were weighted sum of squares proportional to one-half the negative log likelihood $L(\hat{P}|\hat{\theta})$ of the observed annual surplus production data ($\hat{P}_t$), given the surplus-production model parameter estimates $\hat{\theta}=\left(\hat{a},\hat{b},\hat{\lambda}_{t-1},\ldots\right)$. We assumed that independent errors (ε_t) in modeling surplus production were normally distributed to accommodate the potential for years with zero and negative surplus production. Models with independent statistical errors can be fit by quadratic linear regression (forced through the origin) with $\hat{P}_t$ as the dependent variable, $\hat{B}_t$ as the independent variable and weights, if required. However, we used non-linear regression (AD Model Builder software, Otter Research Ltd.) to fit external surplus production models with both uncorrelated and correlated errors.

The objective function with uncorrelated errors was:

$$L(\hat{P}|\hat{\theta})=0.5\sum_{t=1}^{N}\left(\frac{\hat{P}_t-\tilde{P}_t}{\sigma_t}\right)^2 \qquad (7)$$

where the standard errors s_t were from inverse variance weighting factors ($w_t=1/s_t^2$) supplied as input data. In the simple case of constant variance, the objective function is the same as an unweighted sum of squares. Following Schnute (1985), the objective function with correlated errors was:

$$L(\hat{P}|\hat{\theta})=0.5\sum_{t=1}^{N}\left(\frac{\varepsilon_t}{\sigma_t}\right)^2 \qquad (8)$$

We used a wide range of methods to characterize uncertainty and correlation in estimates of F_{MSY} and B_{MSY} and other model estimates. In particular, we used the delta method based on asymptotic variances for parameters (i.e. from the Hessian matrix in non-linear regression), empirical bootstrap (i.e. original weighted residuals e_y/s_y sampled with replacement, with appropriate calculations for autocorrelated errors), likelihood profiles, and numerical Markov Chain Monte Carlo (MCMC) techniques. Preliminary runs for some stocks indicated that the product $R_{MSY}=F_{MSY}B_{MSY}$ might be estimated robustly because of negative correlation between the individual terms (higher estimates of F_{MSY} tend to be offset in the product by lower estimates of B_{MSY}) so we estimated the variance of the product using all methods.

In interpreting bootstrap results, it is important to remember that the simulation analyses assume a true underlying model with all of the parameters at their estimated value. Bootstrap calculations give confidence intervals and variance estimates that can be compared to results using other techniques. Bivariate distributions for F_{MSY} and B_{MSY} estimates from bootstrap runs were plotted in three dimensions to illustrate the correlation between estimates of F_{MSY} and B_{MSY}.

Summary and Discussion

Externally estimated surplus production models are useful because they summarize assessment model results in terms of surplus production, use all of the information in the original stock assessment model, are simple enough to be carried out in a spreadsheet, depict surplus production-biomass relationships in a way that is easy to understand, and often provide useable estimates of MSY parameters. Moreover, they help assessment scientists avoid problems relating fishing mortality estimates from one model (e.g. VPA) to MSY reference point calculations from a second model (e.g. ASPIC). Hilborn (2001) and Jacobson et al. (in press) recommend carrying out external surplus production calculations routinely, even if results are not used to estimate MSY reference points.

In this paper, we fit a family of nested surplus production models with a linear term only, linear and quadratic terms, and with uncorrelated and correlated statistical errors to accommodate serial correlation in residuals, a common problem in surplus production modeling. The linear model is appropriate and useful when the dynamic range of the data is limited to low biomass levels. The nested model approach could be easily extended to asymmetric surplus production models with an additional parameter (e.g. Pella and Tomlinson 1969). However, data were not sufficient to estimate asymmetric surplus production curves for the stocks in this analysis.

Long time series are most useful in fitting surplus production models so biomass estimates from stock assessment models were supplemented in several cases by rescaling and smoothing bottom trawl survey data for early years. Sensitivity analyses were used to assess affects of combing data from different sources.

Surplus production was negative in 0% to 12% of years, depending on the stock. The best external surplus production models and MSY parameter estimates are summarized below. Models with moving average error terms and weighting were required for most stocks. All stocks showed P/B ratios declining with biomass suggesting density dependent production relationships. However, linear surplus production (rather than Schaefer surplus production) models were used for sea scallop in the Mid-Atlantic Bight and white hake due to lack of dynamic range in biomass levels.

Stock	F_{MSY}	B_{MSY}	Data/Model	Comment
Striped Bass	0.18	69,437	1982-2000 from VPA; 1965-1981 from scaled survey, catch and other data; Schaefer model; No weights; Independent errors	
Summer flounder	0.40	42,398	1982-2000 from VPA; 1974-1981 from scaled survey data; Schaefer model; Downweight data for 1978-1981; MA-1 errors	Implausible, F_{MSY} probably too high, B_{MSY} probably too low due to apparent low production in mid-1970's, possibly stemming from lack of recreational catch data
Redfish	0.087	135,241	1934-1999 from preliminary stock assessment; Schaefer model; Downweight data for 1934-1962; MA-3 errors	
White hake	0.23	Not estimate d	1989-1999 from a preliminary (and problematic!) VPA; Linear production model; No weights; Independent errors	B_{MSY} not estimable due to limited data and linear production model; F_{MSY} possibly biased low; Data from problematic VPA biomass estimates
Gulf of Maine Cod	0.44	22,988	1982-1999 from VPA (1963-1981 excluded due to lack of fit); Scaefer model; No weights; MA-2 errors	Implausible, F_{MSY} probably too high, B_{MSY} probably too low, possibly due to limited dynamic range and imprecise estimates

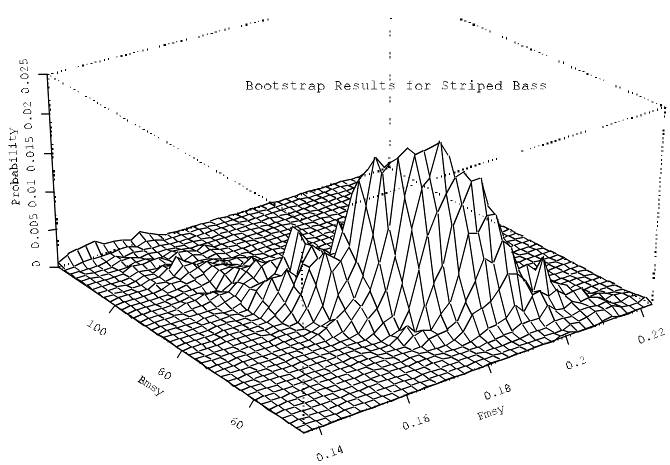
The F_{MSY} estimate (0.23 y^{-1}) from the external linear surplus production model for white hake was based on a problematic VPA. However, the estimate seems plausible (i.e. approximately the same as the assumed natural mortality M=0.2 y^{-1} and an estimate F_{MSY} =0.25 y^{-1} from Applegate et al. 1998). The external estimate may be robust if average surplus production and average biomass were measured accurately by the VPA.

External MSY parameter estimates were implausible for summer flounder with estimates of B_{MSY} probably too low and estimates of F_{MSY} probably. Problems with summer flounder stem from apparently low surplus production during the mid-1970's. Low surplus production, in the context of a surplus production model, indicates that the stock is at carrying capacity. However, summer flounder supported substantial catches for many years prior to the beginning of the mid-1970's. For example, survey age composition and survey abundance data (not included in the stock assessment or external surplus production model) indicate that summer flounder fishing mortality was high prior to the outset of the time series used in production modeling. We hypothesize that apparent low surplus production in early years for summer flounder may have been due to missing recreational catch data.

The best external estimate of F_{MSY} for striped bass was lower and the estimate of B_{MSY} was higher than proxy values used in managing the striped bass fishery. Results for summer flounder were implausible but the best external estimate of F_{MSY} was higher and the estimate of B_{MSY} was lower than proxy values used in managing the fishery. The best external estimate of F_{MSY} for redfish was similar to a preliminary estimate from ASPIC, but estimates of B_{MSY} were different. The best external estimate of F_{MSY} for white hake was similar to an estimate in Applegate et al. (1998) from ASPIC, but there was no external estimate of B_{MSY}. The external estimate of F_{MSY} and B_{MSY} for Georges Bank sea scallop, and the estimate of F_{MSY} for the Mid-Atlantic stock, were similar to proxies used in managing the fishery. External results for Gulf of Maine cod were similar to estimates from a preliminary ASPIC run.

B_{MSY} is more difficult to estimate than F_{MSY} for most of the stocks in this analysis. In practical terms, B_{MSY} may be inestimable currently for some stocks (e.g. white hake) regardless of the modeling approach due to lack of contrast in the available data.

We estimated uncertainty in estimates by bootstrapping, likelihood profiles, the delta method and Markov Chain Monte Carlo techniques. Bootstrapping usually gave the widest confidence intervals with confidence intervals from other methods that were similar. For most stocks, F_{MSY} was estimated more precisely than B_{MSY} because of the relatively large number of data points for low biomass levels and relatively small or lack of data points at high biomass levels. Statistical distributions measuring uncertainty were more symmetrical for F_{MSY} than for B_{MSY}. There was usually a strong negative correlation between F_{MSY} and B_{MSY} estimates. The distribution of bootstrap F_{MSY} and B_{MSY} estimates for striped bass (see below) is typical and shows the relative uncertainty in both parameters as well as their correlation.



Estimates of product MSY = F_{MSY} B_{MSY} were more usually more precise than estimates of B_{MSY}. Thus, the product may be useful in choosing a proxy value for B_{MSY} that corresponds to a particular proxy for F_{MSY}. For example, in striped bass, the product

$$F_{MSY} B_{MSY} \approx 17{,}000 \text{ mt y}^{-1},$$

and F30% (one of many potential proxies for F_{MSY})=0.3 y^{-1}. A corresponding B_{MSY} proxy value might be approximated as 17,000/0.3=57,000 mt.

Uncertainty about F_{MSY} and B_{MSY} is greater than estimated by statistical means in this analysis. Uncertainty is higher because estimation of B_{MSY} involves extrapolation beyond the available biomass data for most stocks and because there is increased uncertainty about model structure at high biomass levels. In particular, the symmetric Schaefer surplus production model may not fit at higher biomass levels. A plausible looking quadratic Schaefer surplus production curve could probably be fit to data from any stock, even if the underlying production curve was asymmetric. This would hold as long as the range of biomass levels did not extend beyond $B_{MSY,}$ because quadratic models are generally good approximations to any monotonic trend over a short interval. This apparent robustness of quadratic models does not imply that F_{MSY} and B_{MSY} parameters are robust, however, because the real F_{MSY} and B_{MSY} values depend on the surplus production relationship in nature, not on the curve fit to the data.

Uncertainty in the biomass estimates, natural mortality, catches and many other factors were not considered in estimation of confidence intervals for this paper. To evaluate the effects of these factors on uncertainty, it will be necessary to incorporate external or internal surplus production model calculations into the original stock assessment model (Jacobson et al. in press). If all calculations are carried out in the same computer program, bootstrap variance calculations for estimates of F_{MSY} would, for example, include uncertainty about biomass and production estimates.

If sufficient data are available, external fits provide useable estimates of MSY parameters and help avoid problems relating fishing mortality estimates from one model (e.g. VPA) to MSY reference point calculations from a second model. Biomass estimates and externally estimated MSY parameters are from the same data and imply the same levels and trends in fishing mortality, biomass and recruitment. However, potential problems due differences in units (e.g. reference points as biomass weighted F's and assessment model estimates of fully recruited F's) remain.

The biomass data B_t used in fitting external surplus production models may, in practice, be fishable biomass, total biomass, fishable abundance, total abundance, or calculated in the original assessment model according to any other convention that is reasonable under the circumstances. However, the interpretation of B_{MSY} and F_{MSY} may be affected. For example, if B_t measures fishable biomass and fishery selectivity is reasonably constant over time, then F_{MSY} estimates are equivalent to F_{MSY} for fully recruited individuals.

The calculations in this paper are based on most recent or preliminary assessment results and meant only to demonstrate the potential utility of using external surplus production models for a wide range of stocks off the northeastern US. Estimates of *MSY* reference points are not for use by managers unless reviewed, and possibly revised.

SENSITIVITY OF MSY REFERENCE POINTS TO RECRUITMENT MODEL

The Sissenwine-Shepherd (1987) age-based approach provides another alternative to surplus production models for estimating MSY. The Sissenwine-Shepherd approach incorporates more biological detail into the model but as noted by Mohn and Black (1998), such estimates are highly sensitive to the assumed relationship between spawning stock and recruits. To assess the implication of the S-R function on biological reference points, the working group considered the effects of five different recruitment models for Georges Bank yellowtail flounder stocks. At present the overfishing definition for this stock is based on a surplus production model but the estimates of B_{MSY} have been unstable. Thus it seemed appropriate to determine if model with more biological realism could improve the estimation of biological reference points.

Application to Georges Bank Yellowtail Flounder

Dynamic pool estimates of yield, mean biomass, and SSB per recruit were estimated for Georges Bank yellowtail flounder using 1994-2000 data (Stone et al. 2001). Five different stock-recruitment models for SSB (mt) and R (recruitment in millions) were assumed. Results of the model fits and comparisons with other alternatives are summarized below.

Model	Error	Years	MSY (mt)	SSB_{msy} (mt)	B_{msy}	F_{msy} (ages 4+)	Fmsy (wb)	Baseline
YPR		94-00						F_{max}=0.82 $F_{0.1}$=0.25 $F_{20\%}$=0.67
B-H: R=(50.3 S) /8.37+S)	lognormal	73-99	10,230	36,772	58,290	0.35	0.18	
B-H: R=(82.8 S)/(13.7+S)	Normal	73-99	16,860	60,620	96,080	0.35	0.18	
B-H: R=(105.8 S)/(24.69+S)	lognormal	60-99	19,620	80,620	123,530	0.30	0.16	
B-H: R=(83 S)/(10.39+S)	Normal	60-99	17,890	57,460	94,060	0.40	0.19	
Constant: R = 20.8	lognormal	73-99	5,530	10,080	20,480	0.82	0.27	
ASPIC-surplus production			14,140		43,470		0.33	

Estimates of MSY and Bmsy were sensitive to the assumed recruitment model. MSY estimates varied by a factor of three among all models and SSBmsy estimates varied by a factor of four. Although these results are deterministic, they suggest that stochasticity should include S-R specification error. Age-based estimates of Fmsy were consistently less than those from ASPIC but MSY and Bmsy were sensitive to the assumed model for the stock-recruitment relationship. While integrated estimates of biological reference points may be more appropriate when based on age-structured models, the choice of a stock-recruitment relationship is likely to become the predominant factor in the estimation. Hence the justification of an appropriate S-R model or set of S-R models should be rigorous. Formal methods of model selection (eg. AIC) may be useful in these instances.

BAYESIAN SURPLUS PRODUCTION MODELS FOR GULF OF MAINE-GEORGES BANK REDFISH

A Bayesian surplus production (BSP) model was applied to catch data and relative abundance indices for redfish to address: (1) whether initial population biomass (B_1) was an estimable parameter when it was not assumed to be equal to carrying capacity (K); (2) whether nonlinear models for survey catchability could be reliably estimated; (3) whether BSP models results were robust to the choice of the prior distribution for carrying capacity; and (4) how BSP model results compare with age-structured and ASPIC-based results (Mayo et al. 2001). The BSP model, the prior distributions, the input data, and selected alternative models are described in the following sections. We then address the four questions and discuss some implications of our findings.

Bayesian Surplus Production Model

We use a Bayesian state-space formulation of the Schaefer surplus production model (Meyer and Millar 1999, NEFSC 2000, Brodziak et al. 2001). This model uses a reparameterized form of the Schaefer surplus production model which relates the fraction of carrying capacity ($P_t = B_t/K$) to intrinsic growth rate, carrying capacity, and the catch time series as

$$P_t = P_{t-1} + rP_{t-1}(1 - P_{t-1}) - \frac{C_{t-1}}{K}$$

This relationship is the basis of the state equations for the state-space model. Stock biomass changes through time due to harvest and biomass production. Under the assumption that B_1=K, the state equations determine changes in relative stock biomass through time (t=1,...,N) via:

$$P_1 = \exp(u_1)$$

$$P_t = \left(P_{t-1} + rP_{t-1}(1 - P_{t-1}) - \frac{C_{t-1}}{K} \right) \exp(u_t) \quad \textit{for } t \geq 2$$

where the independent lognormal process errors for relative biomass are $\exp(u_t)$ with

$u_t \sim N(0,\sigma^2)$.

Relative abundance in year t is measured by either standardized fishery CPUE or the swept-area index (I_t) from the NEFSC autumn and spring bottom trawl surveys. In the simplest form, the CPUE or survey index is assumed to be proportional to stock biomass with constant catchability (Q) throughout the assessment time horizon. This is the *linear catchability model*

$$I_t = QB_t$$

Alternatively the CPUE or survey index is assumed to be proportional to stock biomass raised to a power (β) with constant catchability throughout the assessment time horizon. This is the *nonlinear catchability model*

$$I_t = QB_t^{\beta}$$

Either the linear or nonlinear relationship forms the basis of the observation equations for the state-space model. Stock biomass is measured by the time series of survey indices. For linear catchability, the observation equations relate the observed survey indices to parameters

$$I_t = QKP_t \cdot \exp(v_t) \ \textit{for} \ t = 1,\ldots,N$$

where the independent lognormal observation errors are $\exp(v_t)$ with $v_t \sim N(0,\tau^2)$. Similarly, the observations equations for nonlinear catchability are

$$I_t = Q(KP_t)^{\beta} \cdot \exp(v_t) \ \textit{for} \ t = 1,\ldots,N$$

Using fishery CPUE and two surveys as tuning indices, all with nonlinear catchability, the BSP model has fifteen parameters (r, K, σ^2, Q_{CPUE}, CPUE_σ^2, CPUE_τ^2, CPUE_β, Q_{FALL}, fall_σ^2, fall_τ^2, fall_β, Q_{SPR}, spr_σ^2, spr_τ^2, spr_β) and N unknown relative biomasses (P_t) for a total of N+15 unknowns. To describe the Bayesian estimation procedure, let the joint prior of the parameters and unobservables be p(Θ). Further, let the joint likelihood of the survey indices given the parameters and unobserved states be p(Data | Θ) and the joint posterior distribution of the unobservables be p(Θ | Data).

Bayes' theorem determines the posterior as a function of the prior and likelihood as

$$p(\Theta|Data) = \frac{p(Data|\Theta)\,p(\Theta)}{\int_{\Theta} p(Data|\Theta)\,p(\Theta)\,d\Theta}$$

Direct calculation of the posterior distribution is not possible for the BSP model because the integral in the denominator of the right hand side is not tractable. As a result, Markov chain Monte Carlo (MCMC) methods were used to obtain samples from the posterior distribution of a Bayesian model (Gilks et al. 1996, Brooks 1998). Gibbs sampling is one type of MCMC algorithm that can be readily applied using the BUGS software (Gilks et al. 1994; Meyer and Millar 1999). Computer code to fit the BSP model was implemented using the WINBUGS1.3 software.

Prior distributions

The prior distribution for carrying capacity was chosen to be either informative or uninformative. The *informative prior* distribution for K was a lognormal distribution with parameters chosen to set the 10^{th} and 90^{th} percentiles of the distribution. These percentiles were 100 kmt and 1,000 kmt, respectively. The *uninformative prior* for K was a broad uniform distribution where K~Uniform[1 kmt, 10000 kmt]. Similarly, the prior distribution for intrinsic growth rate was a broad uniform distribution with r~Uniform[0.01, 1.99].

The prior distribution for the inverse of CPUE or survey catchability was chosen to be a high-variance gamma distribution. In particular, the inverse of Q was assumed to be distributed as Gamma(0.001,0.001). This choice gives a relatively flat prior for Q, p(Q), that is approximately proportional to 1/Q, that is,

$p(Q) \propto 1/Q$. In addition, the range of possible values of Q was bounded to fall within the interval [0.01, 10000] to ensure that model predictions of survey biomass indices (QKP_t) were also bounded. The prior for process error variance parameter (σ^2) was also chosen to be an inverse gamma distribution. The inverse of σ^2 was distributed as Gamma(4.00, 0.01). This choice led to a 10% and 90% quantiles for σ of 0.04 and 0.08, respectively. Similarly, the prior for observation error variance (τ^2) was chosen to be an inverse gamma distribution for each tuning index. The inverse of τ^2 was distributed as Gamma(2.00, 0.01). This choice led to a 10% and 90% quantiles for τ of 0.05 and 0.14, respectively. This implied that observation error was somewhat larger than process error, although these parameters were freely estimated using the MCMC algorithm.

The prior distributions for the relative biomasses (P_t) were lognormal distributions for each year, with the possible exception of the initial year. The prior distribution for relative biomass in the initial year of the assessment time horizon was either lognormal with a mean set to B_1=K or an uniform prior. The assumption that B_1=K was relaxed by choosing a broad uninformative prior for P_1 to examine the consequences of not assuming that initial population abundance was at carrying capacity. This prior was

$$P_1 = Uniform[0.01, 1000]$$

For subsequent years, the conditional prior distribution of P_t (conditioned on values of P_{t-1}, K, r, and σ^2) was

$$P_t \sim Lognormal\left(P_{t-1} + rP_{t-1}(1 - P_{t-1}) - \frac{C_t}{K}, \sigma^2 \right)$$

Thus, the prior distribution for relative biomass in year t was dependent upon the previous year's relative biomass, intrinsic growth rate, carrying capacity, and the process error parameter.

Input Data and Alternative BSP Models

Input data were taken from Mayo et al. (2001). These consisted of time series of catch biomass for 1934-2000, standardized fishery CPUE for 1952-1989, and NEFSC autumn and spring survey biomass indices. These input data were the same as the data used for the age-structured model and ASPIC model presented in Mayo et al. (2001).

We fit six alternative BSP models to address (1) whether initial population biomass (B_1) was an estimable parameter when it was not assumed to be equal to carrying capacity (K); (2) whether nonlinear models for survey catchability could be reliably estimated; (3) whether BSP models results were robust to the choice of the prior distribution for carrying capacity; and (4) how BSP model results compare with age-structured and ASPIC-based results. Each model used the same input data. The six models differed in the initial population biomass assumption (B_1=K or $B_1 \neq$K), the tuning index catchability assumption (linear or nonlinear), and the prior assumed for K (informative or uninformative). The six models were:

1. Informative prior on K, B_1=K, and nonlinear catchability
2. Uninformative prior on K, B_1=K, and nonlinear catchability
3. Uninformative prior on K, $B_1 \neq$K, and nonlinear catchability
4. Informative prior on K, $B_1 \neq$K, and nonlinear catchability

5. Informative prior on K, B_1=K, and linear catchability
6. Uninformative prior on K, B_1=K, and linear catchability

<u>Results</u>

Figures D7 through D7 show the results of fitting the six BSP models to redfish input data. In these figures, estimates of exploitable biomass and their 80% CI are depicted as well as median estimates of r, K, maximum surplus production (MSP), the ratio of exploitation rate in year 2000 to the MSP exploitation rate (HRATIO), initial biomass in 1934 (B_{1934} also denoted as B_1), terminal biomass in 2001 (B_{2001}), along with their estimated coefficients of variation (CVs) in parentheses. In addition, the range of CVs for the tuning index catchability coefficients (Qs) are listed in parentheses. Note that the coefficients of variation are provided to give an indication of the precision of the parameter estimates and are not intended for hypothesis testing. Each of the six models shows a long-term decline in redfish biomass from the 1930s to the 1950s, a moderate increase in biomass in the late 1960s, followed by a further decline in biomass through the late 1980s, and an increase in biomass during the 1990s. Overall, the primary difference between the model results is the scale of the biomass trajectory.

Is initial population biomass estimable if it is not assumed to be equal to carrying capacity?
The answer appears to be "No". The redfish BSP models where $B_1 \neq K$ (Figures D8 and D9) have extremely large CVs on initial biomass (92% and 136%) which indicates that this parameter is imprecisely determined. Although we did not have time to complete analyses with linear catchability and $B_1 \neq K$, it is likely that this imprecision is an inherent feature that would not be affected by the choice of catchability submodel, based on our experience with this BSP model. Overall, the two BSP models (3 and 4) with $B_1 \neq K$ are less credible than the others due to this imprecision in B_1.

Are parameters of the nonlinear catchability models reliably estimated?
The answer is probably not, unless they are interpreted as nuisance parameters that can be expected to have high correlation due to nonlinear model structure, as, for example, one might expect in the estimation of L_∞ and K in the von Bertalanffy growth model. The range of CVs for the catchability coefficients (Qs) of the BSP models where B_1 = K with nonlinear catchability is large (Figures D6 and D7), on the order of 100% and these parameters are imprecisely determined. In contrast, the power coefficients (βs) of the nonlinear catchability submodel had lower Cvs, on the order of 20-40%. This suggests that there was probably insufficient information to estimate two parameters for each catchability submodel. We note that this behavior was also apparent for the models where $B_1 \neq K$.

Are the BSP model results robust to the choice of prior distribution for carrying capacity?
The answer appears to be "Yes". For each of the pairs of BSP models using informative and uninformative priors for K, e.g., models 1 & 2 (Figures D6 and D7), models 3 & 4 (Figures D8 and D9), and models 5 & 6 (Figures D10 and D11), the results are generally consistent, with similar estimates of r, K, MSP, HRATIO, B_{1934} , and B_{2001} obtained using informative and uninformative priors for K.

How do the BSP model results compare with age-structured and ASPIC-based results?
The results for the two BSP models with $B_1 \neq K$ are consistent with an initial version of the age-structured dynamics model for redfish

(not shown) where the size of the plus-group in the initial year was estimated as a free parameter. In this case, the plus-group size was estimated to be very large in comparison to subsequent recruitment estimates, similar to the large initial population biomasses in Figures D8 and D9. This initial age-structured model was discounted by the NDWG because there was no information to discern recruitment strengths of year classes in the plus-group during the initial year and because it implied that the redfish population was far from an equilibrium state in 1934. Overall, this suggests that estimates of initial biomass different from carrying capacity are not likely to be well determined for redfish. However, since no directed fishery for redfish existed prior to 1934, we believe it is reasonable to assume that B_{1934} was near carrying capacity.

The results for the two BSP models with nonlinear catchability (Fig. D12) are very similar to the results of the age-structured dynamics model for redfish. With the exception of an increase in biomass in the late 1960s, the age-structured and BSP biomass trajectories are quite similar after 1952 when the earliest tuning index (CPUE) begins. This similarity in biomass trajectories over the range of years where there was tuning information (1952-2000) is probably the result of similar modeling assumptions. In particular, the age-structured dynamics model includes a nonlinear catchability submodel for fishery CPUE, to account for non-random behavior of fishing fleets in relation to redfish density, and also includes a nonlinear catchability submodel for the NEFSC spring survey, to account for differences in redfish schooling behavior and availability to survey trawl gear during this season. In contrast, the BSP model results with nonlinear catchability submodels are less consistent with the ASPIC model results, most likely because ASPIC assumes a linear catchability submodel.

Similarly, the results for the two BSP models with linear catchability (Figure D13) are similar to the results of the ASPIC model for redfish, with the exception of the late 1960s. Presumably this is a consequence of both BSP and ASPIC models using the same catchability submodel for the tuning indices.

Discussion

The result that it is not probably not possible to estimate an initial population biomass for redfish that differs from carrying capacity is not surprising given the available data. In particular, the tuning indices for redfish extend from 1952-2000 and so the only information on population dynamics at the beginning of the modeling time horizon (1934-2001) is the catch. Regardless of this indeterminacy, it seems satisfactory to assume that initial redfish biomass was probably near carrying capacity because it is a long-lived species with low natural mortality (e.g., analogous to a K-selected species) that was not subject to a directed fishery prior to 1934. Moreover, it is encouraging to observe that the BSP results for a particular model configuration were robust to the choice of either an informative or uninformative prior for carrying capacity.

The higher precision of the estimates of Qs for BSPs with linear versus nonlinear catchability submodels is consistent with our observation that the marginal posterior densities of the linear models were much smoother. This visual diagnostic shows that the mixing rate for the MCMC chains was much better with the linear versus the nonlinear catchability assumption. Note, however, this does not imply that catchability is in fact linear. Instead, it merely shows that there is insufficient information input to the BSP model to give precise estimates of the particular nonlinear model that we examined. Nonlinear catchability submodels may be more appropriate for modeling the catchability of redfish, but there is no way to discern this using only a biomass dynamics model.

The inclusion of process error in the BSP models allowed for deviations from the simple Schaeffer model dynamics to fit observed tuning indices. This increased flexibility is the reason that biomass estimates from the BSP models and the ASPIC model differed in the late 1960s. In general, allowing for process error is more realistic but comes at the expense of reduced precision, due to the need to estimate the unknown biomass in each year (N=67). We believe that this trade-off worth it because the Bayesian model provides a more realistic depiction of the error processes, and hence a better quantification of the underlying uncertainty in the management parameters of interest for decision analyses. Or as Tukey once put it, "*Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise.*"

RESEARCH RECOMMENDATIONS

The Methods Subcommittee considered a wide range of topics. The resulting spectrum of research recommendation is similarly broad. It was noted that many of the problems of estimation arise from the lack of integration of reference points in current assessment models. Models that allow direct estimation of biological reference points within the context of biomass and mortality estimation should be useful. However, it should be noted that lack of contrast in the data cannot be overcome by more sophisticated estimation procedures. In that regard, the apparent recovery of many stocks in the Northeast will afford considerable insights into population dynamics. It is important that managers, scientists and industry conduct specific studies and experiments during this period.

The existing software for age-structured projections (e.g., AGEPRO) might be modified to allow for direct search of MSY and biological reference points. There was insufficient time to explore this option for the present SARC. Similarly, age-structured production models and delay difference approaches may provide additional realism in the estimation of biological reference points.

It will be useful to apply the diagnostic measures proposed herein to other species. Improbable changes in reference points over short time periods are probably indicative of poor fits. Similarly, variations in estimates among models are unlikely to occur if the underlying data support the model. Simulation tests of this principle would be helpful.

The subcommittee reviewed a number of species with widely varying life histories, fisheries, and data quality. It is unlikely that a single model or approach will sufficient to capture the underlying dynamics and biological reference points. Bayesian and state-space approaches may also prove to be adequate to incorporate the necessary realism for such species.

SARC COMMENTS

The SARC reviewed working documents and presentations covering several aspects of production modeling. Production models are of special importance in resource assessment because of their role in the development of biological reference points such as those used by the Council.

Many of the biological reference points now in use, however, were developed from age-aggregated production models that pool information across age groups, whereas many of the stocks now managed are fully assessed using age-disaggregated models (e.g. age-structured, cohort-analysis or VPA models). As pointed out in the November 2000 Report of the Groundfish Overfishing Definition Committee, many of the biological reference points developed using production modeling now need to be updated using more comprehensive approaches, which may include production modeling as one component.

Inconsistencies between age-aggregated and age-disaggregated assessments often result from differences in the information content of the data, as well as how biomass and fishing mortality have been defined. These differences have led to attempts by scientists and others to convert the output from one approach so that it corresponds to the recommendations from another. For example, a biomass weighted F from an age-disaggregated model is needed to evaluate the attainment of certain targets such as Fmsy, which are derived from age-aggregated production modeling exercises.

The difficulty here is not production modeling as such, rather it is the development of a better understanding of the association between age-aggregated and age-disaggregated modeling results and how to incorporate new and better information into the management process when that information becomes available. In reviewing this issue, the SARC broadly interpreted the terms of reference and examined not only production-model-based biological reference points, but also considered alternative formats for presenting information important to management, and in particular discussed how information resulting from age structured modeling approaches could be utilized.

SARCs have not formally reviewed such fundamental methodological issues in recent years. And although the current SARC had sufficient expertise to evaluate the production modeling approaches presented at this session, time constraints and the priority given to current stock assessment evaluations limited the discussion on production models and on the specification of biological references points in general.

In addition, although the presentations and discussions provided by Center scientists on these topics were enlightening and useful, the standard Advisory Report format for reporting scientific stock status and stock specific management advice to the Council is not well suited to reviews of technical issues that are more methodological in nature. Therefore we provide only a brief summary of our findings and indicate areas requiring further consideration, perhaps in a workshop setting supported by considerable analysis.

<u>Purpose</u>
The Terms of Reference for this review address a number of issues that may be classified broadly as:

Technical Issues:

1) How can we make better use of, and develop further, production modeling approaches?

2) What are the alternatives to production modeling and how do results from alternative approaches compare to proxy biological reference points? In particular, what should we do in circumstances where additional information, such as age-disaggregated information, becomes available?

Non-technical Issues:

1) What are the implications for management of characterizing a fishery using a broader information base? In particular, how can we best use age-disaggregated information and model results in forming management objectives that have been based historically upon outputs from production models (including outputs, such as Fmsy and Bmsy)?

2) What are the directions for future research in these areas?

3) Is this the appropriate forum for developing, reviewing and presenting results from that methodological research?

Results and Discussion

Center scientists presented information and analyses covering:

D1) Graphical and diagnostic evaluation of production models

D2) External surplus production models

D3) Methods for estimating production and Fmsy in any stock model

D4) Bayesian surplus production models

The SARC listened to these presentations, discussed the implications and potential implementation of each subject presented, then moved into the broader questions of whether management objectives should be constrained to production model type outputs, how more comprehensive information, such as that available from age-disaggregated data and modeling, could be used, and what the appropriate format might be for such discussions in the future.

What follows is a brief summary discussion of these presentations, a discussion of production modeling in the context of other approaches to fisheries management, and a brief statement about the usefulness to scientists and the Council of having methods such as these discussed by the SARC.

NEFSC Presentations

Graphical approaches to the presentation of information and its use in facilitating diagnosis in model estimation have evolved rapidly in the last two decades. Part of this is the availability of easy to use graphical and statistical presentation techniques that make use of the powerful way humans can interpret information visually. Center scientists developed a number of these techniques to demonstrate how they might be used to analyze and interpret production modeling results. The methods presented included exploratory data analysis (EDA) applied to survey catch-rate comparisons to evaluate the adequacy of the data for answering certain questions, envelope plots for presenting the bounds on uncertainty in biomass estimates, model behavior plots to assess density dependence as exhibited by a stock, smoothed time dependent and regression plots to explore correlations between survey indices, response surface plots to describe the uncertainty of parameters of interest, such as biological reference points (which should also be useful in assessing risk associated with decision making), and funnel plots, a newly proposed concept to assess the information content of correlated data. The funnel plots that were discussed are a powerful extension to the widely used retrospective analysis approach and will be useful for other assessment models as well as production models.

It is encouraging to see such novel visual approaches being used and developed. In fact, the use of these approaches is not limited to examining surplus production modeling and can of course be used to examine data (such as survey data) directly as well as be applied to more comprehensive approaches, for example age-structured modeling. The routine use and review of these methods will enhance the use of production modeling and other modeling approaches, increase the likelihood that application of such models is appropriate to the available data, and will allow Council members and their support committees to readily visualize the strengths and weaknesses of the data available.

External production models were investigated as a means to link age-structured models to the more familiar age-aggregated surplus production model. External refers to where in the process a production model is applied. In contrast, internal approaches estimate model parameters simultaneously while other estimates, for example age-structured estimates, are derived. The results, while not encouraging for the data examined, point to important efforts at trying to compare and reconcile production modeling with other modeling and estimation techniques. The National Research Council reports (dates) on stock assessment and improving the collection and use of data encourage applying alternative approaches to data to better understand the information it contains. One consequence, pointed out by the SARC, of comparing age-structured analyses with, for example, production model outputs may be the recognition that earlier production modeling results may need to be reevaluated and updated. These comparisons also force us to clearly define the population biomass in the context of fishing selectivity, fishing effort (possibly from multiple sources), other sources of mortality or population change (such as through discarding, migration, or environmental change), and in the context of management objectives and constraints. Debate still exists on preferred approaches for estimating biological reference points, but efforts such as those discussed are encouraged and should continue to be evaluated through the SAW and SARC processes, as well as through peer review from the broader scientific community.

A presentation was made specifically on estimating biological reference points, such as Fmsy, from any stock assessment model. This, of course can be done, as was pointed out in this work, but it begs the larger question of whether results from more comprehensive models should necessarily be winnowed down to the classical production model outputs.

Fishery models, in general, tend to assume that the fundamental dynamics of a population are in some form of approximate equilibrium. The crude equilibrium assumptions of the past have been abandoned, but by definition models are simplifications of nature and may not capture all population responses to changing abundance (e.g., changes in fecundity patterns, growth, or age structure) or responses to habitat changes over time (e.g., from contaminants or development). This issue may be particularly important in areas, like New England, where most fish stocks are quite far from equilibrium, and is certainly important when planning recovery of depressed stocks. Thus, we consider it especially valuable to explore population dynamics with a variety of models of differing underlying assumptions.

This presumes that the information to conduct an age or size structured assessment is available, but even if it is not, alternatives exist. Several were raised during the SARC discussion on this issue including delay-difference analyses and analyses simply involving catch-rate or survey indices. And looking more towards the future, there exists the possibility of expanding these approaches to the problem of multispecies and trans-boundary stocks.

The final presentation, on a Bayesian surplus production model applied to redfish, showed that the initial population size (if not assumed to be at carrying capacity) was difficult to estimate and that non-linear catchability was also relatively difficult to capture. The Bayesian results, however, were consistent with those of other age structured models and ASPIC (an age-aggregated production model). Again, this is evidence that Center scientists are engaging in research approaches that will broaden the level of information available to stock assessment and decision making. Analogous to discussions made earlier on graphical methods and model comparisons, Bayesian approaches can also be applied more generally, such as to age-structured population analyses, or even to estimating more elementary statistics such as estimates of catch-rate or survey abundance. The Center should be encouraged to continue to explore methods for making the best use of the information they gather through surveys and fishing records.

For each of these topics, stock specific recommendations have been incorporated into the advice for the respective species.

Production Modeling in the Context of Other Approaches to Fisheries Management

Production modeling is a valuable tool in the stock assessment toolbox. It provides a reasonable method of synthesizing information, especially in those situations where very little information is available. (It relies simply upon recorded total catch and an index of abundance or fishing effort.) And, as demonstrated in the presentations discussed, there exist a variety of means of determining the quality of the information available from these methods and for presenting the information they contain.

And yet, there are many instances where there is more information available to the scientist and manager than that provided by total catch or a survey index alone, where the stock may be far from equilibrium, or where the more immediate consequences of the biological response to management actions may be as important as the longer-term consequences of these actions. In these instances, a more comprehensive approach may be required, and scientists and managers should not feel constrained to fitting the results from these more encompassing approaches into the statistics provided from a simple production model analysis.

Unfortunately, many of the biological reference points currently used as management proxies may fit into this category. Stock management will benefit from making use of this broader information base, but scientists will have to respond to that need by providing updated measures that characterize populations more broadly and indicate where additional information is needed and how the population, as defined, should be interpreted.

The utility of production and age-structured modeling is often improved if their interpretation is linked with simpler analyses, that may include simple exploratory data analyses, as well as to more complex analyses, such as age-structured or multispecies models.

All models benefit from longer data series that demonstrate a higher contrast in biomass levels in response to harvest rates. And it goes without saying that if the data in general is poor then no model will suffice. At the other extreme, good information may be poorly utilized, and consideration should be given to which summary statistics are the most informative and robust to uncertainties in the data.

One recommendation is the use of ratios with regard to biological reference points. For example, in representing current F to Fmsy it may be more reliable to consider the ratio of one to the other than considering either estimate in absolute terms. In many instances the absolute levels will change, while their relationship to one another remains stable. Possibilities for deriving more informative and robust measures should be explored.

The SARC notes there has been progress made on a number of fronts on production modeling and data analysis in general. Additional work is needed on utilizing information from production models, age-structured models and more generalized approaches in order to facilitate means for managers and stakeholders to interpret this information in the context of management decisions. This may indicate that we will need to step beyond a few simple biological reference points to viewing alternative means (alternative pathways) towards achieving our goal of sustainable fisheries. This also indicates that vehicles for development, review and implementation of these methods are needed and should be established.

The work reported was useful in generating advice for the stocks under consideration and for assessments in general.

In age-structured models the focus of the uncertainty in the estimation of biological reference points shifts to the specification of a stock-recruit relationship and depends more heavily upon the dynamics observed in recent recruitments. In some instances, age structured approaches may improve upon the estimation of biological reference points or even provide a broader base of reference upon which to make decisions.

<u>SARC Input into Methodological Review and Development</u>

The SARC recognized the benefit of having a discussion related to methodological development and implementation and thought that such discussions should continue. However, it was not clear if this was the most appropriate forum or structure for these discussions. (These discussions certainly will influence and would be influenced by other stock assessment discussions outside the ones currently being considered, namely cod, white hake, and redfish.) Nor was it clear to whom the discussions were to be directed. (Are the discussions aimed at directing fisheries scientists alone, or Council advisors, or the Council itself?) As a consequence, the SARC broke this report into components dealing with the specific presentations, as given above, and into a general discussion-debate. The notes to follow represent a list of ideas based on that general discussion-debate, and should form a good starting point for designing future explorations.

Evaluation of Production Models and Modeling Approaches in General

- Model exploration illustrates the consequences of model choice and provides guidance on uncertainty.

- Model exploration demonstrates the limitation of both models and data.

- Multiple models may provide a needed perspective on uncertainty.

- Exploration of alternative methods demonstrates the limitations and consequences of lack of information.

- Graphical methods, diagnostic approaches, and model comparisons provide a good way of understanding the behavior of models to different pieces of information.

- A single number or statistic may give false sense of security (certainty) about the question being addressed.

- Both real data and simulated data are useful in understanding and characterizing model performance.

RECOMMENDATIONS

! Complex systems may require alternative perspectives and approaches.

! The methods working group should consider a decision theoretic framework under certain management conditions.

! Evolving methods and expanding an information base available for fishery management implies that managers will perceive Amoving goal posts@. This does not mean the rules are changing, but rather that new information has been brought to bear on the problem. This suggests that input controls, such as effort control, may be more a more effective and stable management tool than output controls, such as catch limits. Consideration should be given to such approaches.

! Scientists and managers should be encouraged to use modeling exercises to explore the effectiveness of control rules in achieving production and standing stock objectives and to explore the consequences and risks of alternate management actions.

! Adaptive approaches are encouraged in order to find limits of productivity.

! As information changes it will continue to be important to chronological changes in the fishery, the stock, and the catch so that information from new scientific and management approaches can be linked to what has happened in the past. In other words, preserve history.

BIOLOGICAL REFERENCE POINTS

An evaluation of how biological reference points such as MSY, Fmsy and Bmsy compare between models is a useful exercise, but it is better done on a stock-by-stock basis where the units of comparison and models of choice are clearly defined.

In many instances, it may be difficult to compare Bmsy, for example, from one model to another in absolute terms as the definition of biomass implicit in the model may vary from one model to the next. In one instance biomass may be best defined as the exploitable stock biomass, in another instance it may be best defined as the reproductive stock biomass, or even stock numbers.

This is all very difficult, of course, when the biological and legal settings have been framed in terms of these numbers. This suggests some alternative methods for using this information. First, how stock biomass has been defined should be made explicit for analysis, goal setting, and deliberations. Second, comparisons should be viewed in a relative rather than an absolute sense. For example, one might ask instead what is the ratio of current biomass to Bmsy, what is the current fishing mortality rate is relative to Fmsy, or what the current yield is relative to MSY. These comparisons are less likely to change as models and model estimates are updated than are the absolute values themselves. Finally, recognize that if two or more analytical approaches exist, one can always ask for short-term and long-term predictions under, for example, a statis quo scenario or an Fmsy scenario for each approach to see what consequences, if any, exist under each perspective.

One major difference between production models and age-structured models is in how new biomass enters the standing stock. In production models the influx of new biomass comes in, usually instantaneously, as a proportion of the current biomass. In age-structured models new biomass comes in through recruitment, usually with a time lag and many times based upon a stock-recruitment relationship. Both representations are subject to assumptions and simplifications. It is the robustness of inferences to these assumptions that should form the basis of debate.

These suggestions will not solve all problems encountered in reference point comparisons, but consideration of these issues should move the process towards uses of these measures that have greater stability under uncertainty.

REFERENCES

Applegate, A., S. Cadrin, J. Hoenig, C. Moore, S. Murawski, and E. Pikitch. 1998. Evaluation of existing overfishing definitions and recommendations for new overfishing definitions to comply with the Sustainable Fisheries Act. New England Fishery management Council Report, Newburyport, MA, 179 p.

Brodziak, J.K.T., E.M. Holmes, K.A. Sosebee, and R.K. Mayo. 2001. Assessment of the silver hake resource in the northwest Atlantic in 2000. Northeast Fisheries Science Center Ref. Doc. 01-03.

Brooks, S. 1998. Markov chain Monte Carlo method and its application. Statistician, 47:69-100.

Cook, R. D. and S. Weisberg. 1990. Confidence curves in nonlinear regression. J. Amer. Stat. Assoc. 82:221-230.

Draper, N. R. and H. Smith. 1966. Applied regression analysis. Wiley. NY.

Gilks, W.R., A. Thomas, and D.J. Spiegelhalter. 1994. A language and program for complex Bayesian modeling. Statistician, 43:169-178.

Gilks, W.R., S. Richardson, and D.J. Spiegelhalter. 1996. Markov chain Monte Carlo in practice. Chapman and Hall, London.

Hilborn, R. 2001. Calculation of biomass trend, exploitation rate, and surplus production from survey and catch data. Canadian Journal of Fisheries and Aquatic Sciences 58: 579-584.

Jacobson, L.D., S.X. Cadrin, and J.R. Weinberg. In press. Tools for estimating surplus production and FMSY in any stock assessment model. N. Am. J. Fish. Mgmt.

Jacobson, L. D., and fourteen coauthors. In press. Surplus production, variability and climate change in the great sardine and anchovy fisheries. Canadian Journal of Fisheries and Aquatic Sciences.

MacCall, A. D. 1978. A note on production modeling of populations with discontinuous reproduction. California Fish and Game 64:225-227.

Mayo, R.K., J. Brodziak, M. Thompson, J. Burnett, and S. Cadrin. 2001. Biological characteristics, population dynamics, and current status of redfish, *Sebastes fasciatus* Storer, in the Gulf of Maine - Georges Bank region. SAW 33/SARC Working Paper C1.

Meyer, R. and R.B. Millar. 1999. BUGS in Bayesian stock assessments. Can. J. Fish. Aquat. Sci. 56:1078-1086.

Mohn, R. and G. Black. 1998. Illustrations of the precautionary approach using 4TVW haddock, 4VsW cod and 3LNO American plaice. NAFO SCR Doc. 1998/10.

Northeast Fisheries Science Center [NEFSC]. 2000. 31st Northeast regional stock assessment workshop: Stock assessment review committee consensus summary of assessments. NEFSC Ref. Doc. 00-15.

Prager, M. H. 1994. A suite of extensions to a nonequilibrium surplus-production model. Fishery Bulletin 92:374-389.

Ricker, W. E. 1975. Comparison and interpretation of biological statistics of fish populations. Fisheries Research Board of Canada, Bulletin Number 191.

Schaefer, M. B. 1957. A study of the dynamics of the fishery for yellowfin tuna in the eastern tropical Pacific Ocean. Inter-American Tropical Tuna Commission Bulletin 2:245-285.

Schnute, J. T. 1985. A general theory for analysis of catch and effort data. Canadian Journal of Fisheries and Aquatic Sciences 42:414-429.

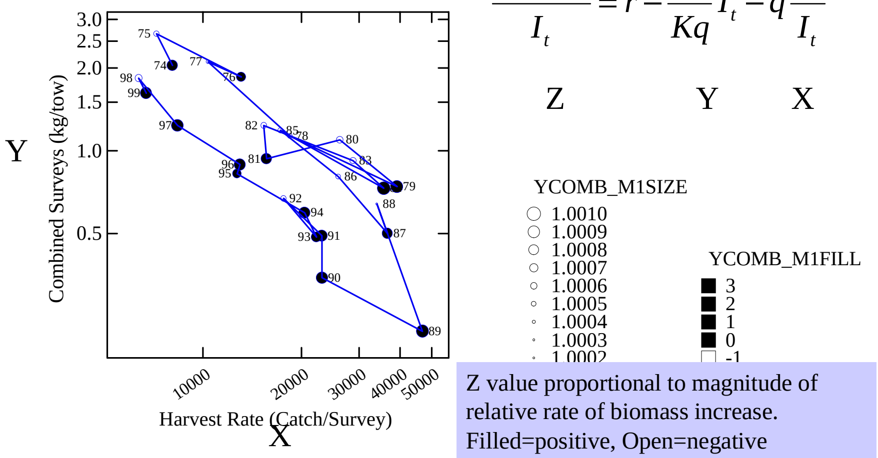
Sissenwine, M.P. and J. G. Shepherd. 1987. An alternative perspective on recruitment overfishing and biological reference points. Can. J. Fish. Aquat. Sci. 44:913-918.

Stone, H.H., C.M. Legault, S.X. Cadrin, S. Gavaris, and P. Perley. 2001. Stock assessment of Georges Bank (5Zjmnh) yellowtail flounder for 2001. DFO Res. Doc. 2001 (in press).

Thompson, W. F. and F. H. Bell. 1934. Effect of changes in intensity upon total yield and yield per unit of gear. Report of the International Fisheries Commission 8:7-49.

U.S. Department of Commerce. 1997. National Oceanic and Atmospheric Administration. 50 CFR Part 600. Docket No. 970708168-7168-01. Proposed rule Federal Register, August 4, 1997. Vol. 62, No. 149. Pages 41907-41920.

Design Sufficiency: Application to combined index of spring and fall research survey average weight per tow



$$\frac{I_{t+1}-I_t}{I_t} = r - \frac{r}{Kq} I_t - q\frac{C_t}{I_t}$$

Z Y X

Fig D1. Example plot to illustrate the notion of "design sufficiency for summer flounder stock.

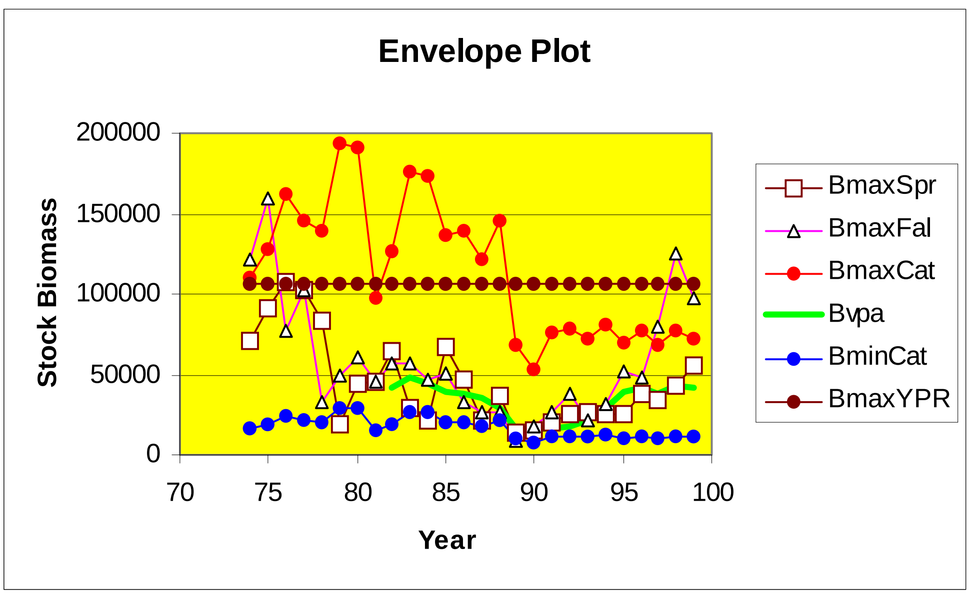
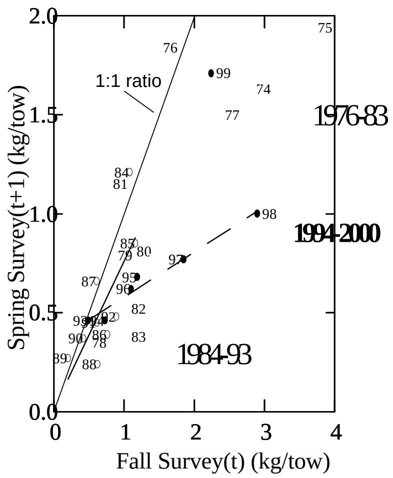


Fig. D2. Construction of an envelope plot for summer flounder. BminCat and BmaxCat are lower and upper bounds respectively, of biomass estimates based on assumed upper and lower bounds of fishing mortality rates. Bvpa is the derived estimate from the ADAPT VPA model. The BmaxSpr and BmaxFal represent biomass levels for swept area estimates from spring and fall survey estimates. BmaxYPR represents the proxy biomass target constructed as the product of B/R at Fmax and the average recruitment estimated from the VPA. See text for additional details.

Potential changes in the relationship between spring and fall surveys over time



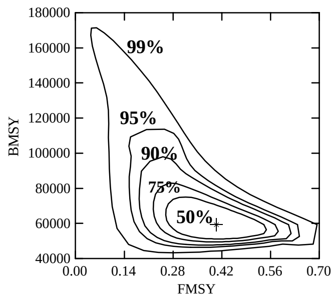
Linear regression lines plotted with bisquare function to downweight residuals.

Take home message: Multiple indices should consistently measure the same set of population attributes over time

Fig. D3. Evaluation of apparent changing relationship between spring and fall survey indices for summer flounder for three stanzas: 1976-1983, 1984-1993, and 1994-2000.

4A

Response surface for NLLS with approximate 50, 75, 90,95,99 % probability levels. Based on F stat.



Approximate Confidence Region:
Find the set of points satisfying the following equation

$$S(\Theta) = S(\hat{\Theta})\left(1 + \frac{p}{n-p} F_{p,n-p,1-\alpha}\right)$$

Where S(Θ) =residual sum of squares in model as a function of

$$\Theta = \begin{Bmatrix} F_{MSY} \\ B_{MSY} \\ q \end{Bmatrix}$$

4B

Apply a typical control rule to the confidence region for reference points

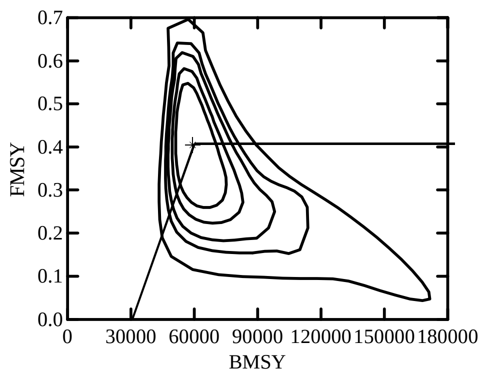


Fig. D4. Approximate confidence regions for F_{MSY} and B_{MSY} for example application to summer flounder. Panel A illustrates the computation methodology. Panel B llustrates the application of the uncertainty in point estimation to development of hypothetical control rule.

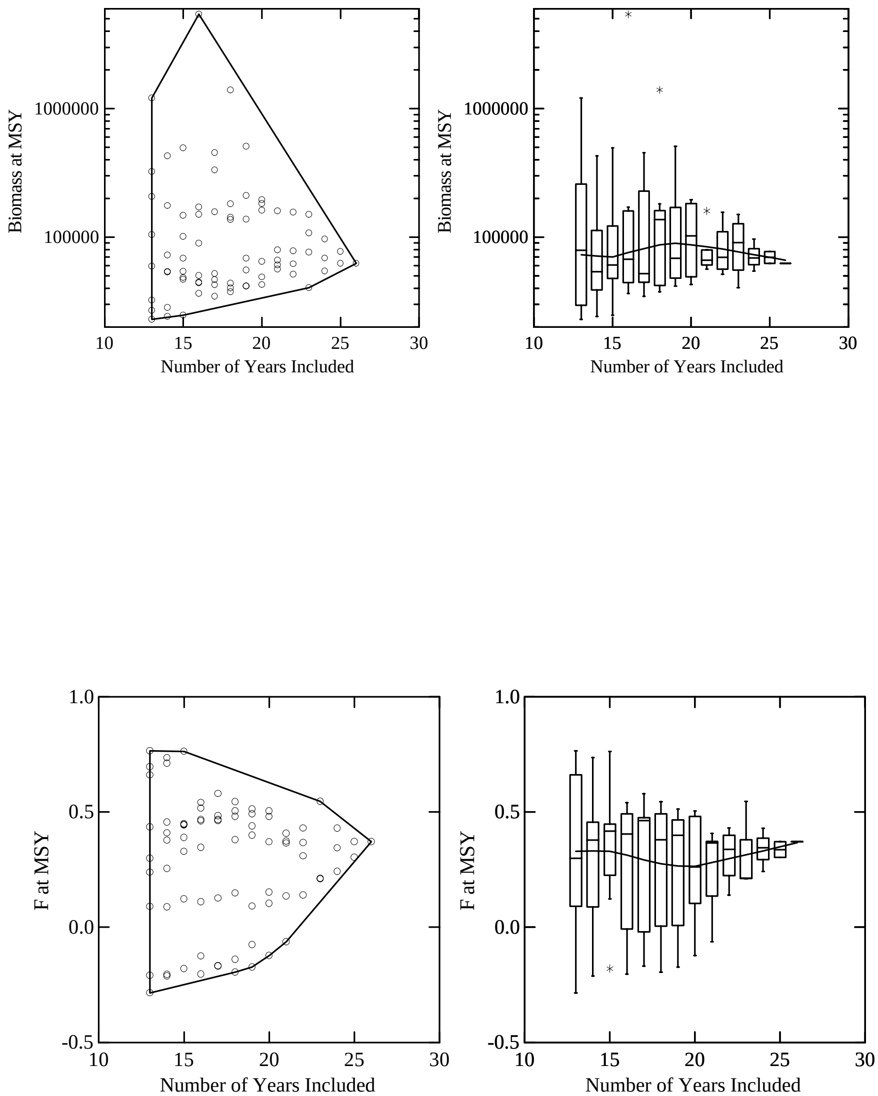


Fig D5. Example “funnel plots” for estimation of parameters of surplus production model for summer flounder. The top two panels show the convergence of estimates of F_{MSY} as additional years of data are incorporated. The lower two panels illustrate the convergence of estimates of B_{MSY}. See text for additional details.

Fig. D6 Results for redfish BSP model 1.

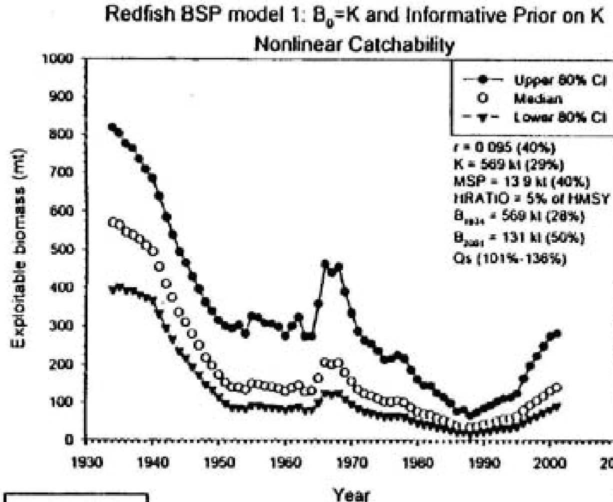


Fig. D7 Results for redfish BSP model 2.

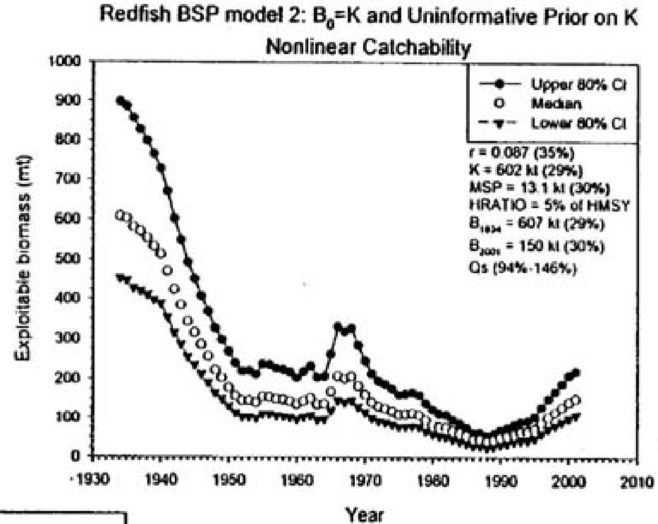


Fig. D8 Results for redfish BSP model 3.

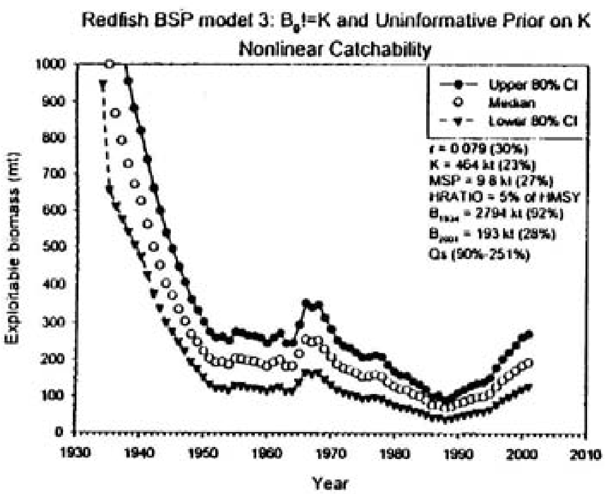


Fig. D9 Results for redfish BSP model 4.

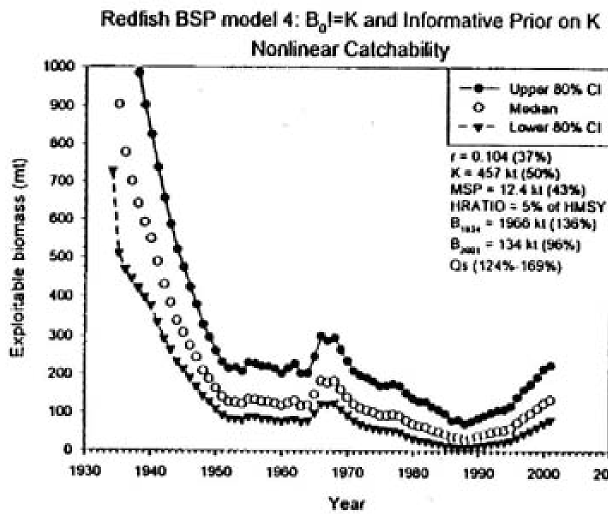


Fig. D10 Results for redfish BSP model 5.

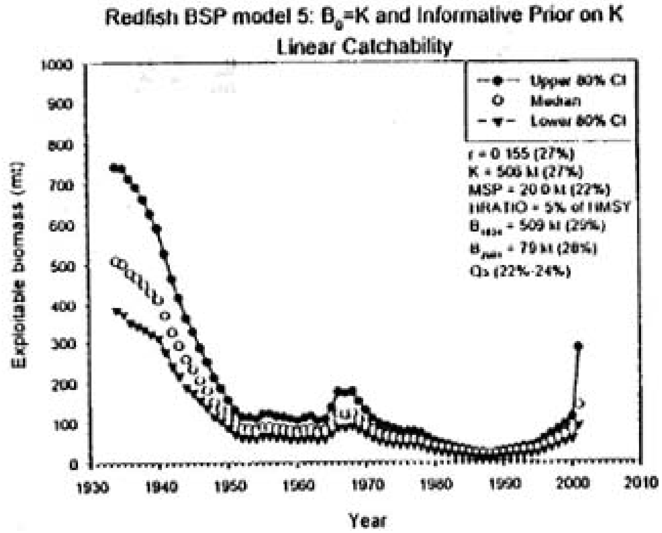


Fig. D11 Results for redfish BSP model 6.

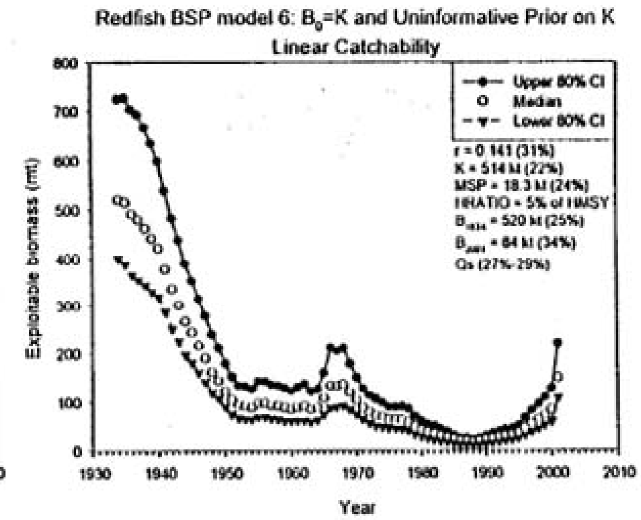


Fig. D12 Comparison of BSP results using nonlinear catchability with assessment results.

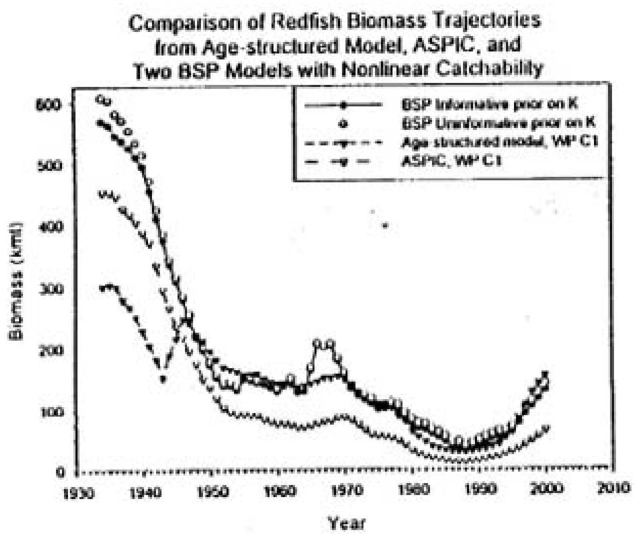


Fig. D13 Comparison of BSP results using linear catchability with assessment results.