

**Appendix A.**  
**(of “C. Atlantic Surfclam” SARC-37 Report):**

**The KLAMZ Assessment Model**

The KLAMZ assessment model (NEFSC 2000; 2001) is based on the Deriso-Schnute delay-difference equation (Deriso 1980; Schnute 1985; Quinn and Deriso 1999). The delay-difference equation is a relatively simple and implicitly age structured model that counts fish in either numerical or biomass units. It gives the same results as explicitly age-structured models (e.g. Leslie matrix model) if fishery selectivity is “knife-edged”, somatic growth follows the von Bertalanffy equation, and natural mortality is the same for all age groups in each year. Knife-edge selectivity means that all individuals alive in the model during the same year experience the same fishing mortality rate.<sup>1</sup> Natural and fishing mortality rates, growth parameters and recruitment may change from year to year, but delay-difference calculations assume that all individuals share the same mortality and growth parameters within each year.

As in many other simple models, the delay difference equation explicitly distinguishes between two age groups. In KLAMZ, the two age groups are called “new” recruits and “old” recruits. New recruits are individuals that recruited at the beginning or during the current year. Old recruits are all older individuals in the model. As described above, KLAMZ assumes that new and old recruits are fully vulnerable to the fishery. The most important differences between the delay-difference and other simple models (e.g. Prager 1994; Conser 1995; Jacobson et al. 1994) are that von Bertalanffy growth is used to calculate biomass dynamics and that the delay-difference model captures transient age structure effects due to variation in recruitment, growth and mortality exactly. Transient effects on population dynamics are captured exactly because, as described above, the delay-difference equation is algebraically equivalent to an explicitly age-structured model with von Bertalanffy growth. As described above, delay-difference calculations can be carried out in units of biomass or numerical abundance. The KLAMZ model includes simple numerical models as special cases (e.g. Conser 1995) because growth can be turned off so that all calculations are in numerical units (see below).

The KLAMZ model incorporates a few extensions to Schnute’s (1985) revision of Deriso’s (1980) original delay difference model. Most of the extensions facilitate tuning to a wider variety of data that anticipated in Schnute (1985). The KLAMZ model was programmed in both Excel and in C++ using AD Model Builder libraries<sup>2</sup>. The AD

---

<sup>1</sup> In applications, assumptions about knife-edge selectivity can be relaxed by assuming the model tracks “fishable”, rather than total, biomass (NEFSC 2000a; 2000b). An analogous approach assigns pseudo-ages based on recruitment to the fishery so that new recruits in the model are all pseudo-age  $k$ . The synthetic cohort of fish pseudo-age  $k$  may consist of more than one biological cohort. The first pseudo-age ( $k$ ) can be the predicted age at first, 50% or full recruitment based on a von Bertalanffy curve and size composition data (Butler et al. 2002). The “incomplete recruitment” approach (Deriso 1980) calculates recruitment to the model in each year  $R_t$  as the weighted sum of contributions from two or more cohorts due to spawning in

successive years (i.e.  $R_t = \sum_{a=1}^k r_a \Pi_{t-a}$  where  $k$  is the age at full recruitment to the fishery,  $r_a$  is the

contribution of fish age  $k-a$  to the fishable stock, and  $\Pi_{t-a}$  is the number or biomass of fish age  $k-a$  during year  $t$ ).

<sup>2</sup> Otter Research Ltd., Box 2040, Sydney, BC, V8L 3S3 (otter@otter-rsch.com).

Model Builder version is faster, more reliable and probably better for producing “official” stock assessment results. The Excel version is slower but useful in developing prototype assessment models, teaching and for checking calculations.

## Population dynamics

The assumed birth date and first day of the year are assumed the same in derivation of the delay-difference equation. It is therefore natural (but not strictly necessary) to tabulate catch and other data using annual accounting periods that start on the assumed biological birthday of cohorts.

Schnute’s (1985) delay-difference equation in the KLAMZ model is:

$$B_{t+1} = (1 + \rho) \tau_t B_t - \rho \tau_t \tau_{t-1} B_{t-1} + R_{t+1} - \rho \tau_t J_t R_t$$

where  $B_t$  is total biomass of individuals at the beginning of year  $t$ ;  $\rho$  is Ford’s growth coefficient (see below);  $\tau_t = \exp(-Z_t) = \exp[-(F_t + M)]$  is the fraction of the stock that survived in year  $t$ ,  $Z_t$ ,  $F_t$ , and  $M$  are instantaneous rates for total, fishing and natural mortality; and  $R_t$  is the biomass of new recruits (at age  $k$ ) at the beginning of the year. The natural mortality rate  $M_t$  may vary or be constant over time. Instantaneous mortality rates in KLAMZ model calculations are biomass-weighted averages if von Bertalanffy growth is turned on in the model. However, biomass-weighted mortality estimates in KLAMZ are the same as rates for numerical calculations because all individuals are fully recruited. The growth parameter  $J_t = w_{t-1,k-1} / w_{t,k}$  is the ratio of mean weight one year before recruitment (age  $k-1$  in year  $t-1$ ) and mean weight at recruitment (age  $k$  in year  $t$ ).

It is not necessary to specify body weights at and prior to recruitment in the KLAMZ model (parameters  $v_{t-1}$  and  $V_t$  in Schnute 1985) because the ratio  $J_t$  and recruitment biomass contain the same information. Schnute’s (1985) original delay difference equation is:

$$B_{t+1} = (1 + \rho) \tau_t B_t - \rho \tau_t \tau_{t-1} B_{t-1} + w_{t+1,k} N_{t+1} - \rho \tau_t w_{t-1,k-1} N_t$$

To derive the equation used in KLAMZ, substitute recruitment biomass  $R_{t+1}$  for the product  $w_{t+1,k} N_{t+1,k}$  and adjusted recruitment biomass  $J_t R_t = (w_{t-1,k-1} / w_{t,k}) w_{t,k} N_{t,k} = w_{t-1,k-1} N_t$  in the last term on the right hand side. The advantage in using the alternate parameterization for biomass dynamic calculations in KLAMZ is that recruitment is estimated directly in units of biomass and the number of growth parameters is reduced.

## Growth

As described in Schnute (1985), biomass calculations in the KLAMZ model are based on Schnute and Fournier’s (1980) re-parameterization of the von Bertalanffy growth model:

$$w_a = w_{k-1} + (w_k - w_{k-1}) (1 + \rho^{1+a-k}) / (1 - \rho)$$

where  $w_k = V$  and  $w_{k-1} = v$ . Schnute and Fournier’s (1980) growth model is the same as the traditional von Bertalanffy growth model  $\{W_a = W_{max} [1 - \exp(-K(a - t_{zero}))]$  where  $W_{max}$ ,  $K$  and  $t_{zero}$  are parameters}. The two growth models are the same because  $W_{max} = (w_k - \rho w_{k-1}) / (1 - \rho)$ ,  $K = -\ln(\rho)$  and  $t_{zero} = \ln[(w_k - w_{k-1}) / (w_k - \rho w_{k-1})] / \ln(\rho)$ .

In the KLAMZ model, the growth parameters  $J_t$  can vary with time but  $\rho$  is constant. Use of time-variable  $J_t$  values with  $\rho$  is constant is the same as assuming that the von Bertalanffy parameters  $W_{max}$  and  $t_{zero}$  change over time. It is possible to accommodate a wide range growth patterns by changing only  $W_{max}$  and  $t_{zero}$ . Growth parameters are usually estimated externally, rather than directly in the KLAMZ model. The KLAMZ model uses catch-at-age information indirectly, if catch-at-age is used to estimate growth parameters.

#### *Numerical population dynamics (growth turned off)*

Growth can be turned on off so that abundance, rather than biomass, is tracked in the KLAMZ model. Set  $J_t=1$  and  $\rho=0$  in the delay difference equation, and use  $N_t$  (for numbers) in place of  $B_t$  to get:

$$N_{t+1} = \tau_t N_t + R_{t+1}$$

All of the calculations in KLAMZ for biomass dynamics are also valid for numerical dynamics.

#### *Instantaneous growth rates*

Instantaneous growth rate (IGR) calculations in the KLAMZ model are an extension to the original Deriso-Schnute delay difference model. IGRs are used extensively in KLAMZ for calculating catch biomass and projecting stock biomass forward to the time at which surveys occur. The IGR for new recruits depends only on growth parameters:

$$G_t^{New} = \ln\left(\frac{w_{k+1,t+1}}{w_{k,t}}\right) = \ln(1 + \rho - \rho J_t)$$

IGR for old recruits is a biomass-weighted average that depends on the current age structure and growth parameters. It can be calculated easily by projecting biomass of old recruits  $S_t=B_t-R_t$  (escapement) forward one year with no mortality:

$$S_t^* = (1 + \rho)S_t - \rho\tau_{t-1}B_{t-1}$$

where the asterisk (\*) means just prior to the start of the subsequent year  $t+1$ . By definition, the IGR for old recruits in year  $t$  is  $G_t^{Old} = \ln(S_t^*/S_t)$ . Dividing by  $S_t$  gives:

$$G_t^{Old} = \ln\left[(1 + \rho) - \rho\tau_{t-1}\frac{B_{t-1}}{S_t}\right]$$

IGR for the entire stock is the biomass weighted average of the IGR values for new and old recruits:

$$G_t = \frac{R_t G_t^{New} + S_t G_t^{Old}}{B_t}$$

All IGR values are zero if growth is turned off.

#### Recruitment

In the Excel version of the KLAMZ model, annual recruitments are calculated  $R_t = e^{\Omega_t}$  where  $\Omega_t$  is a log transformed annual recruitment parameter usually estimated in the model. In the C++ version, recruitments are calculated based on log geometric mean

recruitment ( $\mu$ ) and a set of annual log scale deviation parameters ( $\omega_t$ ):

$$\Omega_t = \mu + \omega_t$$

The deviations  $\omega_t$  are constrained to average zero.<sup>3</sup> With the constraint, estimation of  $\mu$  and the set of  $\omega_t$  values (1+  $n$  years parameters) is equivalent to estimation of the smaller set ( $n$  years) of  $\Omega_t$  values.

### Natural mortality

Natural mortality rates ( $M$ ) are assumed constant in the Excel version of the KLAMZ model but can change from year to year in the C++ version based on covariates (e.g. predator density) or natural mortality rate process errors. Natural mortality rate process errors represent variation in predation, disease, parasitism and other factors that affect natural mortality rates in fish populations. Annual process error parameters are estimated to improve model fit to survey and other data. Calculations are basically the same as for survey covariates and survey process errors described below.

### Fishing mortality and catch

Fishing mortality rates ( $F_t$ ) are calculated so that predicted and observed catch data (landings plus estimated discards in units of weight) “agree”. It is not necessary, however, to assume that catches are measured accurately (see “Observed and predicted catch”).

Fishing mortality rate calculations in Schnute (1985) are applicable when catches are in units of numbers but catch data are usually in units of weight. Calculation of predicted catches in units of weight is more complicated because somatic growth occurs throughout the year as fishing occurs.

The KLAMZ model uses a generalized catch equation that incorporates continuous growth through the fishing season. By the definition of instantaneous rates, the catch equation expresses catch as the product:

$$\hat{C}_t = F_t \bar{B}_t$$

where  $\hat{C}_t$  was predicted catch weight (landings plus discard) and  $\bar{B}_t$  is average biomass.

Following Ricker (1970) and Zhang and Sullivan (1988), let  $X_t = G_t - F_t - M_t$  be the net instantaneous rate of change for biomass.<sup>4</sup> If the rates for growth and mortality are equal, then  $X_t = 0$ ,  $\bar{B}_t = B_t$  and  $C_t = F_t B_t$ . If the growth rate  $G_t$  exceeds the combined rates of natural and fishing mortality ( $F_t + M_t$ ), then  $X_t > 0$ . If mortality exceeds growth, then  $X_t < 0$ . In either case, with  $X_t \neq 0$ , average biomass is computed:

$$\bar{B}_t \approx -\frac{(1 - e^{X_t})B_t}{X_t}$$

When  $X_t = 0$ , the expression for  $\bar{B}_t$  is an approximation because  $G_t$  approximates the rate of change in mean body weight due to von Bertalanffy growth. However, the approximation is reasonably accurate and preferable to calculating catch biomass with the

<sup>3</sup> The constraint is implemented by adding  $L = \lambda \bar{\omega}^2$  to the objective function, generally with  $\lambda = 1000$ .

<sup>4</sup> By convention, the instantaneous rates  $G_t$ ,  $F_t$  and  $M_t$  are always expressed as numbers  $\neq 0$ .

traditional catch equation that ignores growth during the fishing season.<sup>5</sup> Average biomass can be calculated for new recruits, old recruits or for the whole stock by using either  $G_t^{New}$ ,  $G_t^{Old}$  or  $G_t$ .

In the Excel version of KLAMZ, the modified catch equation is solved analytically for  $F_t$  given  $C_t$ ,  $B_t$ ,  $G_t$  and  $M$ . In the C++ version, fishing mortality rates are calculated using a log geometric mean parameter ( $\Phi$ ) and a set of annual log scale deviation parameters ( $\psi_t$ ):

$$F_t = e^{\Phi + \psi_t}$$

where the deviations  $\psi_t$  are constrained to average zero.

### Surplus production

Annual surplus production was calculated exactly by projecting biomass at the beginning of each year forward with no fishing mortality:

$$B_t^* = (1 + \rho) e^{-M} B_t - \rho e^{-M} L_{t-1} B_{t-1} - \rho e^{-M} J_t R_t$$

By definition, surplus production  $P_t = B_t^* - B_t$ .

### Per recruit modeling

Per recruit model calculations in the Excel version of the KLAMZ simulate the life of a hypothetical cohort of arbitrary size (e.g.  $R=1000$ ) with constant  $M$ ,  $F$  (survival) and growth ( $\tau$  and  $J$ ) in a population initially at zero biomass. In the first year:

$$B_1 = R$$

In the second year:

$$B_2 = (1 + \rho) \tau B_1 - \rho \tau J R_1$$

In the third and subsequent years:

$$B_{t+1} = (1 + \rho) \tau B_t - \rho \tau^2 B_{t-1}$$

This iterative calculation is carried out until the sum of lifetime cohort biomass from one iteration to the next changes by less than a small amount (0.0001). Total lifetime biomass, spawning biomass and yield in weight are calculated by summing biomass, spawning biomass and yield over the lifetime of the cohort (in each iteration). Lifetime biomass, spawning biomass and yield per recruit are calculated by dividing totals by initial recruitment ( $R$ ).

### Status determination variables

The user may specify a range of years (e.g. the last three years) to use in calculating recent average fishing mortality  $\bar{F}_{Recent}$  and biomass  $\bar{B}_{Recent}$  levels. These status determination variables are often useful in calculation of status ratios such as  $\bar{F}_{Recent} / F_{MSY}$  and  $\bar{B}_{Recent} / B_{MSY}$ .

<sup>5</sup> The traditional catch equation  $C_t = F_t(1 - e^{-Z_t})B_t / Z_t$  where  $Z_t = F_t + M_t$  underestimates catch biomass for a given level of fishing mortality  $F_t$  and overestimates  $F_t$  for a given level of catch biomass. The errors can be substantial for fast growing fish, particularly if recent recruitments were strong.

## Goodness of Fit and Parameter Estimation

Parameters estimated in the KLAMZ model are chosen to minimize an objective function based on a sum of weighted negative log likelihood (NLL) components:

$$\Xi = \sum_{v=1}^{N_{\Xi}} \lambda_v L_v$$

where  $N_{\Xi}$  is the number of NLL components ( $L_v$ ) and the  $\lambda_v$  are emphasis factors used as weights. The objective function  $\Xi$  may be viewed as a NLL or a negative log posterior (NLP) distribution, depending on the nature of the individual  $L_v$  components and modeling approach. Except during sensitivity analyses, weighting factors for objective function components ( $\lambda_v$ ) are usually set to one. An arbitrarily large weighting factor (e.g.  $\lambda_v=1000$ ) is used for “hard” constraints that must be satisfied in the model. Arbitrarily small weighting factors (e.g.  $\lambda_v=0.0001$ ) can be used for “soft” model-based constraints. For example, an internally estimated spawner-recruit curve or surplus production curve might be estimated with a small weighting factor to summarize stock-recruit or surplus production results with minimal influence on biomass, fishing mortality and other estimates from the model. Use of a small weighting factor for an internally estimated surplus production or stock-recruit curve is equivalent to fitting a curve to model estimates of biomass and recruitment or surplus production in the output file, after the model is fit (Jacobson et al. 2002).

### NLL kernels

NLL components in KLAMZ are generally programmed as “concentrated likelihoods” to avoid calculation of values that do not affect derivatives of the objective function. For  $x \sim N(\mu, \sigma^2)$ , the complete NLL for one observation is:

$$L = \ln(\sigma) + \ln(\sqrt{2\pi}) + 0.5 \left( \frac{x - \mu}{\sigma} \right)^2$$

The constant  $\ln(\sqrt{2\pi})$  can always be omitted because does not affect derivatives. If the standard deviation is known or assumed known, then  $\ln(\sigma)$  can be omitted as well because it is a constant that does not affect derivatives. In such cases, the concentrated negative log likelihood is:

$$L = 0.5 \left( \frac{x - \mu}{\sigma} \right)^2$$

If there are  $N$  observations with possible different variances (known or assumed known) and possibly different expected values:

$$L = 0.5 \sum_{i=1}^N \left( \frac{x_i - \mu_i}{\sigma_i} \right)^2$$

If the standard deviation for a normally distributed quantity is not known and is (in effect) estimated by the model, then one of two equivalent calculations is used. Both approaches assume that all observations have the same variance and standard deviation. The first approach is used when all observations have the same weight in the likelihood:

$$L = 0.5N \ln \left[ \sum_{i=1}^N (x_i - u)^2 \right]$$

where  $N$  is the number of observations. The second approach is equivalent but used when the weights for each observation ( $\lambda_i$ ) may differ:

$$L = \sum_{i=1}^N \lambda_i \left[ \ln(\sigma) + 0.5 \left( \frac{x_i - u}{\sigma} \right)^2 \right]$$

In the latter case, the maximum likelihood estimator:

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^N (x_i - \hat{x})^2}{N}}$$

(where  $\hat{x}$  is the average or predicted value from the model) is used for  $\sigma$ . The maximum likelihood estimator is biased by  $N/(N-d_f)$  where  $d_f$  is degrees of freedom for the model. The bias may be significant for small sample sizes but  $d_f$  is usually unknown.

In practice, it is often useful to use a different emphasis factor ( $\lambda_{v,i}$ ) for each observation so that the emphasis for specific observations or specific instances of a constraint can be increased or decreased. KLAMZ allows the user to specify observation- an instance-specific weights for most types of data and constraints.

### Observed and predicted catch

In the AD Model Builder version, fishing mortality rates (based on the parameters  $\Phi$  and  $\psi_f$ ) are estimated to satisfy a NLL for observed and predicted catches:

$$L = \sum_{t=0}^N w_t \left( \frac{\hat{C}_t - C_t}{\kappa_t} \right)^2$$

where the standard error  $\kappa_t = CV_{catch} \hat{C}_t$  with  $CV_{catch}$  and weights are  $w_t$  supplied by the user. The weights can be used, for example, if catch data in some years are less precise than in others. The AD Model Builder version of KLAMZ can potentially estimate any or every catch in the time series.

### *Solving the generalized catch equation*

A few years of catches can be estimated in the Excel version of KLAMZ (see below) but catches are generally assumed measured without error. The Excel version does not compute a NLL for catch. Instead,  $F_t$  values are calculated iteratively using the Newton-Raphson method (Kennedy and Gentle 1980).

Subtracting predicted catch (from the generalized catch equation, see above) from the observed catch data gives:

$$g(F_t) = C_t + \frac{F_t(1 - e^{-X_t})}{X_t} B_t = 0$$

where  $X_t = G_t - M_t - F_t$ . If  $X_t = 0$ , then  $\bar{B}_t = B_t$  and  $F_t = C_t/B_t$ .

If  $X_t \neq 0$ , then the Newton-Raphson algorithm is used to solve for  $F_t$ . At each iteration of the algorithm, the current estimate  $F_t^i$  is updated using:

$$F_t^{i+1} = F_t^i - \frac{g(F_t^i)}{g'(F_t^i)}$$

where  $g'(F_t^i)$  is the derivative  $F_t^i$ . Omitting subscripts, the derivative is:

$$g'(F) = -\frac{Be^{-F}[(e^F - e^\gamma)\gamma + e^\gamma F\gamma - e^\gamma F^2]}{X^2}$$

where  $\gamma = G - M$ . Iterations continue until  $g(F_t^i)$  and  $abs[g(F_t^{i+1}) - g(F_t^i)]$  are both  $\leq 0.00001$ .

Initial values are important in algorithms that solve the catch equation numerically (Sims 1982). If  $M_t + F_t > G_t$  so that  $X_t < 0$ , then the initial value  $F_t^0$  is calculated according to Sims (1982). If  $M_t + F_t < G_t$  so that  $X_t > 0$ , then initial values are calculated based on a generalized version of Pope's cohort analysis (Zhang and Sullivan 1988):

$$F_t^0 = \gamma_t - \ln \left[ \frac{(B_t e^{0.5\gamma_t} - C_t) e^{0.5\gamma_t}}{B_t} \right]$$

### Initial population age structure

In the KLAMZ model, old and new recruit biomass during the first year ( $R_1$  and  $S_1 = B_1 - R_1$ ) and biomass prior to the first year ( $B_0$ ) are estimated as log scale parameters. Survival in the year prior to the first year ("year 0") is  $\tau_0 = e^{-F_0 - M_1}$  with  $F_0$  chosen to obtain catch  $C_0$  (specified as data) from the estimated biomass  $B_0$ . IGRs during year 0 and year 1 are assumed equal ( $G_0 = G_1$ ) in catch calculations.

Biomass in the second year of a series of delay-difference calculations depends on biomass ( $B_0$ ) and survival ( $\tau_0$ ) in year 0:

$$B_2 = (1 + \rho) \tau_1 B_1 - \rho \tau_1 \tau_0 B_0 + R_2 - \rho \tau_1 J_1 R_1$$

There is, however, there is no direct linkage between  $B_0$  and escapement biomass ( $S_1 = B_1 - R_1$ ) at the beginning of the first year.

The missing link between  $B_0$ ,  $S_1$  and  $B_1$  means that the parameter for  $B_0$  tends to be relatively free and unconstrained by the underlying population dynamics model. In some cases,  $B_0$  can be estimated to give good fit to survey and other data, while implying unreasonable initial age composition and surplus production levels. In other cases,  $B_0$  estimates can be unrealistically high or low implying, for example, unreasonably high or low recruitment in the first year of the model ( $R_1$ ). Problems arise because many different combinations of values for  $R_1$ ,  $S_1$  and  $B_0$  give similar results in terms of goodness of fit. This issue is common in stock assessment models that use forward simulation calculations because initial age composition is difficult to estimate. It may be exacerbated in delay-difference models because age composition data are not used.



The KLAMZ model uses two constraints to help estimate initial population biomass and initial age structure.<sup>6</sup> The first constraint links IGRs for escapement ( $G^{Old}$ ) in the first years to an adjacent value. The purpose of the constraint is to ensure consistency in average growth rates (and implicit age structure) during the first few years. For example, if IGRs for the first  $n_G$  years are constrained<sup>7</sup>, then the NLL for the penalty is:

$$L_G = 0.5 \sum_{t=1}^{n_G} \left[ \frac{\ln(G_t^{Old} / G_{n_G+1}^{Old})}{\sigma_G} \right]^2$$

where the standard deviation  $\sigma_G$  is supplied by the user. It is usually possible to use the standard deviation of  $Q_t^{Old}$  for later years from a preliminary run to estimate  $\sigma_G$  for the first few years. The constraint on initial IGRs should probably be “soft” and non-binding ( $\lambda \approx 1$ ) because there is substantial natural variation in somatic growth rates due to variation in age composition.

The second constraint links  $B_0$  to  $S_1$  and ensures conservation of mass in population dynamics between years 0 and 1. In other words, the parameter for escapement biomass in year 1 is constrained to match an approximate projection of the biomass in year 0, accounting for growth, and natural and fishing mortality. The constraint is intended to be binding and satisfied exactly (e.g.  $\lambda = 1000$ ) because incompatible values of  $S_1$  and  $B_0$  are biologically impossible. In calculations:

$$S_1^p = B_0 e^{G_1 - F_0 - M_1}$$

where  $S_1^p$  is the projected escapement in year 1 and  $B_0$  is the model’s estimate of total biomass in year 0. The instantaneous rates for growth and natural mortality from year 1 ( $G_1$  and  $M_1$ ) are used in place of  $G_0$  and  $M_0$  because the latter are unavailable. The NLL for the constraint:

$$L = \left[ \ln \left( \frac{S_1^p}{S_1} \right) \right]^2 + (S_1^p - S_1)^2$$

uses a log scale sum of squares and an arithmetic sum of squares. The former is effective when  $S_1$  is small while the latter is effective when  $S_1$  is large.

### Goodness of fit for survey trends

The NLL used to measure goodness-of-fit for observed and predicted abundance index data with lognormal errors is:

$$L = 0.5 \sum_{j=1}^{N_v} \left[ \frac{\ln \left( I_{v,j} / \hat{I}_{v,j} \right)}{\sigma_{v,j}} \right]^2$$

<sup>6</sup> Quinn and Deriso (1999) describe another approach attributed to a manuscript by C. Walters.

<sup>7</sup> Normally,  $n_G = 2$ .

where  $I_{v,t}$  is an abundance index datum from survey  $v$ , hats “ $\hat{\phantom{x}}$ ” denote model estimates,  $\sigma_{v,j}$  was a log scale standard error (see below), and  $N_v$  was the number of observations. There are two approaches to calculating standard errors for log normal abundance index data in KLAMZ and it is possible to use different approaches for different types of abundance index data in the same model (see below).

Abundance indices with statistical distributions other than log normal may be used as well, but are not currently programmed in the KLAMZ model. For example, Butler et al. (in press) used abundance indices with binomial distributions in a delay-difference model for cowcod rockfish.

#### *Standard errors for goodness of fit*

In the first approach, all observations for one type of abundance index share the same standard error, which is calculated based on overall goodness of fit. The first approach implicitly estimates the standard error based on goodness of fit, along with the rest of the parameters in the model (see “NLL kernels” above).

In the second approach, each observation has a potentially unique standard error that is calculated based on its CV. The second approach calculates log scale standard errors from arithmetic CVs supplied as data by the user (Jacobson et al. 1994):

$$\sigma_{v,t} = \sqrt{\ln(1 + CV_{v,t}^2)}$$

Arithmetic CV’s are usually available for abundance data. It is sometimes convenient to use  $CV_{v,t}=1.31$  to get  $\sigma_{v,t}=1$ .

There are advantages and disadvantages to both approaches. CV’s carry information about the relative precision of abundance index observations. However, CV’s usually overstate the precision of data as a measure of fish abundance.<sup>8</sup> Implicitly estimated standard errors are often larger and more realistic, but imply that all observations in the same survey are equally reliable.

#### *Predicted values for abundance indices*

Predicted values for abundance indices are calculated:

$$\hat{I}_{v,t} = Q_v A_{v,t}$$

where  $Q_v$  is a survey scaling parameter (constant here but see below) that converts units of biomass to units of the abundance index.  $A_{v,t}$  is available biomass at the time of the survey.

In the simplest case, available biomass is:

$$A_{v,t} = s_{v,New} R_t e^{-X_t^{New} \Delta_{v,t}} + s_{v,Old} S_t e^{-X_t^{Old} \Delta_{v,t}}$$

where  $s_{v,New}$  and  $s_{v,Old}$  are survey selectivity parameters for new recruits ( $R_t$ ) and old recruits ( $S_t$ );  $X_t^{New} = G_t^{New} - F_t - M_t$  and  $X_t^{Old} = G_t^{Old} - F_t - M_t$ ;  $j_{v,t}$  was the Julian date at the time of the survey, and  $\Delta_{v,t}=j_{v,t}/365$  was the fraction of the year elapsed at the time of the survey.

Survey selectivity parameter values ( $s_{v,New}$  and  $s_{v,Old}$ ) are specified by the user and must be set between zero and one. For example, a survey for new recruits would have

<sup>8</sup> The relationship between data and fish populations is affected by a host of factors (process errors) that are not accounted for in CV calculations.

$s_{v,New}=1$  and  $s_{v,Old}=0$ . A survey that measured abundance of the entire stock would have  $s_{v,New}=1$  and  $s_{v,Old}=1$ .

Terms involving  $\Delta_{v,t}$  are used to project beginning of year biomass forward to the time of the survey, making adjustments for mortality and somatic growth.<sup>9</sup> As described below, available biomass  $A_{v,t}$  is adjusted further for nonlinear surveys, surveys with covariates and surveys with time variable  $Q_{v,t}$ .

#### Scaling parameters ( $Q$ ) for log normal abundance data

Scaling parameters for surveys with lognormal statistical errors were computed using the maximum likelihood estimator:

$$Q_v = e^{\frac{\sum_{i=1}^{N_v} \left[ \ln \left( \frac{I_{v,i}}{A_{v,i}} \right) \right]^2 / \sigma_{v,i}^2}{\sum_{j=1}^{N_v} \left( 1 / \sigma_{v,j}^2 \right)}}$$

where  $N_v$  was the number of observations with individual weights greater than zero. The closed form maximum likelihood estimator gives the same answer as if scaling parameters are estimated as free parameters in the assessment model assuming lognormal survey measurement errors.

#### Survey covariates

Survey scaling parameters may vary over time based on covariates in the KLAMZ model. The survey scaling parameter that measures the relationship between available biomass and survey data becomes time dependent:

$$\hat{I}_{v,t} = Q_{v,t} A_{v,t}$$

and

$$Q_{v,t} = Q_v e^{\sum_{r=1}^{n_v} d_{r,t} \theta_r}$$

with  $n_v$  covariates for the survey and parameters  $\theta_r$  estimated in the model.

Covariates might include, for example, a dummy variable that represents changes in survey bottom trawl doors or a continuous variable like average temperature data if environmental factors affect distribution and catchability of fish schools. Dummy variables are either 0 or 1, depending on whether the effect was present in a particular year. With dummy variables,  $Q_v$  is the value of the survey scaling parameter with no intervention ( $d_{r,t}=0$ ). For ease in modeling, it is useful to center continuous covariates around their mean:

$$d_{r,t} = d'_{r,t} - \bar{d}'_r$$

where  $d'_{r,t}$  is the original covariate. With covariates that are continuous and mean-centered,  $Q_v$  is the value of the survey scaling parameter under average conditions ( $d_{r,t}=0$ )

<sup>9</sup> It may be important to project biomass forward if an absolute estimate of biomass is available (e.g. from a hydroacoustic or daily egg production survey), if fishing mortality rates are high or if the timing of the survey varies considerably from year to year.

and units for the covariate parameter are easy to interpret (for example, units for the parameter are  $1/^\circ\text{C}$  if the covariate is mean centered temperature in  $^\circ\text{C}$ ).

Covariate effects and available biomass are multiplied to compute an adjusted available biomass:

$$A'_{v,t} = A_{v,t} e^{\sum_{r=1}^{n_v} d_{r,t} \theta_r}$$

The adjusted available biomass  $A'_{v,t}$  is used instead of the original value  $A_{v,t}$  in the closed form maximum likelihood estimator described above.

It is possible to use a survey covariate to adjust for differences in relative stock size from year to year due to changes in the timing of a survey. However, this adjustment may be made more precisely by letting the model calculate  $A_{v,t}$  as described above, based on the actual timing data for the survey during each year.

### *Nonlinear abundance indices*

With nonlinear abundance indices, and following Methot (1990), the survey scaling parameter is a function of available biomass:

$$Q_{v,t} = Q_v A_{v,t}^\Gamma$$

so that:

$$\hat{I}_{v,t} = (Q_v A_{v,t}^\Gamma) A_{v,t}$$

Substituting  $e^\gamma = \Gamma + 1$  gives the equivalent expression:

$$\hat{I}_{v,t} = Q_v A_{v,t}^{e^\gamma}$$

where  $\gamma$  is a parameter estimated by the model and the survey scaling parameter is no longer time dependent. In calculations with nonlinear abundance indices, the adjusted available biomass:

$$A'_{v,t} = A_{v,t}^{e^\gamma}$$

is computed first and used in the closed form maximum likelihood estimator described above to calculate the survey scaling parameter. In cases where survey covariates are also applied to a nonlinear index, the adjustment for nonlinearity is carried out first.

### *Survey Q process errors*

The AD Model Builder version of the KLAMZ model incorporates a very useful ability to let survey scaling parameters change, in a tightly controlled fashion, from year to year (NEFSC 2002):

$$Q_{v,t} = Q_v e^{\varepsilon_{v,t}}$$

where the deviations  $\varepsilon_{v,t}$  are constrained to average zero. Variation in survey Q process errors is controlled by the NLL penalty:

$$L = 0.5 \sum_{j=1}^N \left[ \frac{\varepsilon_{v,j}}{\sigma_v} \right]^2$$

where the log scale standard deviation  $\sigma_v$  is supplied by the user (e.g. see NEFSC 2002).

## Recruitment models

Recruitment parameters in KLAMZ may be freely estimated or estimated around an internal recruitment model, possibly based on spawning biomass. An internally estimated recruitment model may be used to reduce variability in recruitment estimates (often necessary if data are limited), to summarize stock-recruit relationships, or to make use of information about recruitment in similar stocks. There are four types of internally estimated recruitment models in KLAMZ: 1) random variation around a constant mean; 2) random walk around a constant mean (autocorrelated variation); 3) random variation around a Beverton-Holt recruitment model; and 4) random variation around a Ricker recruitment model.

The first step in recruit modeling is to calculate the expected log recruitment level  $E[\ln(R_t)]$  given the recruitment model. For random variation around a constant mean, the expected log recruitment level is the log geometric mean recruitment:

$$E[\ln(R_t)] = \sum_{j=1}^N \ln(R_j) / N$$

For a random walk around a constant mean recruitment, the expected log recruitment level is the logarithm of recruitment during the previous year:

$$E[\ln(R_t)] = \ln(R_{t-1})$$

with no constraint on recruitment during the first year  $R_1$ .

For the Beverton-Holt recruitment model, the expected log recruitment level is:

$$E[\ln(R_t)] = \ln[e^a T_{t-\ell} / (e^b + T_{t-\ell})]$$

where  $a=e^\alpha$  and  $b=e^\beta$ , the parameters  $a$  and  $b$  are estimated in the model,  $T_t$  is spawning biomass, and  $\ell$  is the lag between spawning and recruitment. Spawner-recruit parameters are estimated as log transformed values ( $e^\alpha$  and  $e^\beta$ ) to enhance model stability and ensure the correct sign of values used in calculations. Spawning biomass is:

$$T_t = m_{new} R_t + m_{old} S_t$$

where  $m_{new}$  and  $m_{old}$  are maturity parameters for new and old recruits specified by the user. For the Ricker recruitment model, the expected log recruitment level is:

$$E[\ln(R_t)] = \ln(S_{t-\ell} e^{a-bS_{t-\ell}})$$

where  $a=e^\alpha$  and  $b=e^\beta$ , and the parameters  $a$  and  $b$  are estimated in the model.

Given the expected log recruitment level, log scale residuals for the recruitment model are calculated:

$$r_t = \ln(R_t) - E[\ln(R_t)]$$

Assuming that residuals are log normal, the NLL for recruitment residuals is:

$$L = \sum_{t=\ell}^N \lambda_t \left[ \ln(\sigma_r) + 0.5 \left( \frac{r_t}{\sigma_r} \right)^2 \right]$$

where  $\lambda_t$  is an instance-specific weight usually set equal one. The additional term in the NLL  $[\ln(\sigma_r)]$  is necessary because the variance  $\sigma_r^2$  is estimated internally, rather than specified by the user.

The log scale variance for residuals is calculated using the maximum likelihood estimator:

$$\sigma_r^2 = \frac{\sum_{j=t_{first}}^N r_j}{N}$$

where  $N$  is the number of residuals. For the recruitment model with constant variation around a mean value,  $t_{first}=1$ . For the random walk recruitment model,  $t_{first}=2$ . For the Beverton-Holt and Ricker models,  $t_{first}= +1$  and the recruit model imposes no constraint on variability of recruitment during years 1 to (see below). The biased maximum likelihood estimate for  $\sigma^2$  (with  $N$  in the divisor instead of the degrees of freedom) is used because actual degrees of freedom are unknown. The variance term is calculated explicitly because it is used in other calculations.

#### *Constraining the first few recruitments*

It may be useful to constrain the first years of recruitments when using either the Beverton-Holt or Ricker models if the unconstrained estimates for early years are erratic. In the KLAMZ model, this constraint is calculated:

$$NLL = \sum_{t=1}^{t_{first}-1} \lambda_t \left\{ \ln \left( \sigma_r + 0.5 \left[ \frac{\ln(R_t / E(R_{t_{first}}))}{\sigma_r} \right]^2 \right) \right\}$$

where  $t_{first}$  is the first year for which expected recruitment  $E(R_t)$  can be calculated with the spawner-recruit model. In effect, recruitments that not included in spawner-recruit calculations are constrained towards the first spawner-recruit prediction. The standard deviation and weights used are the same as used in calculating the NLL for the recruitment model.

#### Prior information about abundance index scaling parameters ( $Q$ )

A constraint on one or more survey scaling parameters ( $Q_v$ ) may be useful if prior information about potential values is available (e.g. NEFSC 2000; NEFSC 2001; NEFSC 2002). In the Excel version, it is easy to program these (and other) constraints in an *ad-hoc* fashion as they are needed. In the AD Model Builder version, log normal and beta distributions may be used as prior information in estimating  $Q_v$  for any abundance index

The user must specify which surveys have prior distributions, minimum and maximum legal bounds ( $q_{min}$  and  $q_{max}$ ), the arithmetic mean ( $\bar{q}$ ) and the arithmetic CV for the prior the distribution. Goodness of fit for  $Q_v$  values outside the bounds ( $q_{min}$ ,  $q_{max}$ ) are calculated:

$$L = \begin{cases} 10000 (Q_v - q_{max})^2 & \text{if } Q_v \geq q_{max} \\ 10000 (q_{min} - Q_v)^2 & \text{if } Q_v \leq q_{min} \end{cases}$$

Goodness of fit for  $Q_v$  values inside the legal bounds depend on whether the distribution of potential values is log normal or follows a beta distribution.

#### *Lognormal case*

Goodness of fit for lognormal  $Q_v$  values within legal bounds is:

$$L = \left[ \frac{\ln(Q_v) - \tau}{\varphi} \right]^2$$

where the log scale standard deviation  $\varphi = \sqrt{\ln(1 + CV)}$  and  $\tau = \ln(\bar{q}) - \frac{\varphi^2}{2}$  is the mean of the corresponding log normal distribution.

#### *Beta distribution case*

The first step in calculation goodness of fit for  $Q_v$  values with beta distributions was to calculate the mean and variance of the corresponding “standardized” beta distribution:

$$\bar{q}' = \frac{\bar{q} - q_{\min}}{D}$$

and

$$Var(q') = \left( \frac{\bar{q} CV}{D} \right)^2$$

where the range of the standardized beta distribution is  $D = q_{\max} - q_{\min}$ . Equating the mean and variance to the estimators for the mean and variance for the standardized beta distribution (the “method of moments”) gives the simultaneous equations:

$$\bar{q}' = \frac{a}{a + b}$$

and

$$Var(q') = \frac{ab}{(a + b)^2 (a + b + 1)}$$

where  $a$  and  $b$  are parameters of the standardized beta distribution.<sup>10</sup> Solving the simultaneous equations gives:

$$b = \frac{(\bar{q}' - 1)[Var(q') + (\bar{q}' - 1)\bar{q}']}{Var(q')}$$

and:

$$a = \frac{b\bar{q}'}{1 - \bar{q}'}$$

Goodness of fit for beta  $Q_v$  values within legal bounds was calculated with the NLL:

$$L = (a - 1)\ln(Q'_v) + (b - 1)\ln(1 - Q'_v)$$

where  $Q'_v = Q_v / (Q_v - q_{\min})$  is the standardized value of the survey scaling parameter  $Q_v$ .

#### Surplus production modeling

Surplus production models can be fit internally to biomass and surplus production estimates in the model (Jacobson et al. 2002). Models fit internally can be used to

<sup>10</sup> If  $x$  has a standardized beta distribution with parameters  $a$  and  $b$ , then the probability of  $x$  is

$$P(x) = \frac{x^{a-1}(1-x)^{b-1}}{\Gamma(a,b)}.$$

constrain estimates of biomass and recruitment, to summarize model estimates in terms of surplus production parameters, or as a source of information in tuning the model. The NLL for goodness of fit assumes normally distributed process errors in the surplus production process:

$$L = 0.5 \sum_{j=1}^{N_P} \left( \frac{\tilde{P}_j - P_j}{\sigma} \right)^2$$

where  $N_p$  was the number of surplus production estimates (number of years less one),  $\tilde{P}_t$  was a predicted value from the surplus production curve,  $P_t$  was the assessment model estimate, and the standard deviation  $\sigma$  was supplied by the user based, for example, on preliminary variances for surplus production estimates.<sup>11</sup> Either the symmetrical Schaefer (1957) or asymmetric Fox (1970) surplus production curve may be used to calculate  $\tilde{P}_t$  (Quinn and Deriso 1999).

It may be important to use a surplus production curve that is compatible with assumptions about the underlying spawner-recruit relationship. More research is required, but the asymmetric shape of the Fox surplus production curve appears reasonably compatible with the assumption that recruitment follows a Beverton-Holt spawner-recruit curve (Mohn and Black 1998). In contrast, the symmetric Schaefer surplus production model appears reasonably compatible with the assumption that recruitment follows a Ricker spawner-recruit curve.

The Schaefer model has two log transformed parameters that are estimated in KLAMZ:

$$\tilde{P}_t = e^\alpha B_t - e^\beta B_t^2$$

The Fox model also has two log transformed parameters:

$$\tilde{P}_t = -e \left( e^{e^\alpha} \right) \frac{B_t}{e^\beta} \log \left( \frac{B_t}{e^\beta} \right)$$

See Quinn and Deriso (1999) for formulas used to calculate reference points ( $F_{MSY}$ ,  $B_{MSY}$ ,  $MSY$ , and  $K$ ) for both surplus production models.

### Catch/biomass

Forward simulation models like KLAMZ may estimate absurdly high fishing mortality rates. The likelihood constrain used to prevent this potential problem was calculated:

$$L = 0.5 \sum_{t=0}^N d_t^2$$

where:

<sup>11</sup> Variances in NLL for surplus production-biomass models are a subject of ongoing research. The advantage in assuming normal errors is that negative production values (which occur in many stocks, e.g. Jacobson et al. 2001) are accommodated. In addition, production models can be fit easily by linear regression of  $P_t$  on  $B_t$  and  $B_t^2$  with no intercept term. However, variance of production estimate residuals increases with predicted surplus production. Therefore, the current approach to fitting production curves in KLAMZ is not completely satisfactory.



$$d_t = \begin{cases} (C_t/B - \kappa) & \text{if } C_t/B > \kappa \\ 0 & \text{otherwise} \end{cases}$$

with the threshold value  $\kappa$  normally set by the user to about 0.95. Values for  $\kappa$  can be linked to maximum F values using the modified catch equation described above. For example, to use a maximum fishing mortality rate of about  $F = 4$  with  $M=0.2$  and  $G=0.1$  (maximum  $X=4+0.2-0.1=4.1$ ), set  $\kappa = F/X(1-e^{-X})=4 / 4.1 (1-e^{-4})=0.96$ .

## Uncertainty

The AD Model Builder version of the KLAMZ model automatically calculates variances for parameters and quantities of interest (e.g.  $R_t$ ,  $F_t$ ,  $B_t$ ,  $F_{MSY}$ ,  $B_{MSY}$ ,  $\bar{F}_{Recent}$ ,  $\bar{B}_{Recent}$ ,  $\bar{F}_{Recent} / F_{MSY}$ ,  $\bar{B}_{Recent} / B_{MSY}$ , etc.) by the delta method using exact derivatives. If the objective function is the log of a proper posterior distribution, then Markov Chain Monte Carlo (MCMC) techniques implemented in AD Model Builder libraries can be used estimate posterior distributions representing uncertainty in the same parameters and quantities.

### Bootstrapping

A FORTRAN program called BootADM can be used to bootstrap survey data in the KLAMZ model. BootADM extracts the standardized residuals:

$$r_{v,j} = \frac{\ln \left( I_{v,j} / \hat{I}_{v,j} \right)}{\sigma_{v,j}}$$

log scale standard deviations ( $\sigma_{v,j}$ , originally from survey CV's or estimated from goodness of fit), and predicted values ( $\hat{I}_{v,j}$ ) for all active survey observations in a "base case" KLAMZ model run. The standardized residuals are resampled from a single pool with replacement to form new sets of bootstrapped survey "data":

$${}^x I_{v,j} = \hat{I}_{v,j} e^{r \sigma_{v,j}}$$

where  $r$  is a resampled residual. BootADM builds new KLAMZ data files and runs the KLAMZ model repetitively, collecting the bootstrapped parameter and other estimates at each iteration and writing them to a comma separated text file that can be processed in Excel to calculate bootstrap variances, confidence intervals, bias estimates, etc. for all parameters and quantities of interest (Efron 1982).

### Projections

Stochastic projections can be carried out using another FORTRAN program called SPROJDDF based on bootstrap output from BootADM. Basically, bootstrap estimates of biomass, recruitment, spawning biomass, natural and fishing mortality during the terminal years are used with recruit model parameters from each bootstrap run

to start and carry out projections.<sup>12</sup> Given a user-specified level of catch or fishing mortality, the delay-difference equation is used to project stock status for a user-specified number of years. Recruitment during each projected year is based on simulated spawning biomass, log normal random numbers, and spawner-recruit parameters (including the residual variance) estimated in the bootstrap run. This approach is similar to carrying out projections based on parameters and state variables sampled from a posterior distribution for the basecase model fit. It differs from most current approaches because the spawner-recruit parameters vary from projection to projection.

## References

- Butler, J.L., L.D. Jacobson, J.T. Barnes, and H.G. Moser. 2003. Biology and population dynamics of cowcod (*Sebastes levis*) in the southern California Bight. *Fish. Bull.* 101: 260-280.
- Chapman, D.W. 1971. Production. In: W.E. Ricker (ed.). *Methods for assessment of fish production in fresh waters*. IBP Handbook No. 3. Blackwell Scientific Publications, Oxford, UK.
- Conser, R.J. 1995. A modified DeLury modeling framework for data limited assessment: bridging the gap between surplus production models and age-structured models. A working paper for the ICES Working Group on Methods of Fish Stock Assessment. Copenhagen, Denmark. Feb. 1995.
- Deriso, R. B. 1980. Harvesting strategies and parameter estimation for an age-structured model. *Can. J. Fish. Aquat. Sci.*, 37: 268-282.
- Efron, B. 1982. *The jackknife, the bootstrap and other resampling plans*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 92 p.
- Fournier, D., and C.P. Archibald. 1982. General theory for analyzing catch at age data. *Can. J. Fish. Aquat. Sci.* 39: 1195-1207.
- Fox, W. W. R. Jr. 1970. An exponential yield model for optimizing exploited fish populations. *Trans. Am. Fish. Soc.* 99: 80-88.
- Jacobson, L.D., N.C.H. Lo, and J.T. Barnes. 1994. A biomass based assessment model for northern anchovy *Engraulis mordax*. *Fish. Bull. U.S.* 92: 711-724.
- Jacobson, L.D., J.A.A. De Oliveira, M. Barange, M.A. Cisneros-Mata, R. Felix-Uraga, J.R. Hunter, J.Y. Kim, Y. Matsuura, M. Niquen, C. Porteiro, B. Rothschild, R.P. Sanchez, R. Serra, A. Uriarte, and T. Wada. 2001. Surplus production, variability and climate change in the great sardine and anchovy fisheries. *Can. J. Fish. Aquat. Sci.* 58: 1891-1903.
- Jacobson, L.D., S.X. Cadrin, and J.R. Weinberg. 2002. Tools for estimating surplus production and  $F_{MSY}$  in any stock assessment model. *N. Am. J. Fish. Mgmt.* 22:326-338
- Kennedy, W.J. Jr., and J.E. Gentle. 1980. *Statistical computing*. Marcel Dekker, Inc., New York.
- Methot, R.D. 1990. Synthesis model: an adaptable framework for analysis of diverse stock assessment data. *Int. N. Pac. Fish. Comm. Bull.* 50: 259-277.

---

<sup>12</sup> At present, only Beverton-Holt recruitment calculations are available in SPROJDDF.

- Mohn, R., and G. Black. 1998. Illustrations of the precautionary approach using 4TVW haddock, 4VsW cod and 3LNO plaice. Northwest Atlantic Fisheries Organization NAFO SCR Doc. 98/10.
- NEFSC. 2000. Ocean quahog, p. 172-304. *In*: Report of the 31<sup>st</sup> Northeast Regional Stock Assessment Workshop (31<sup>st</sup> SAW): Stock Assessment Review Committee (SARC) consensus summary of assessments. Northeast Fish. Sci. Cent. Ref. Doc. 00-15.
- NEFSC. 2001. Sea scallops. *In*: Report of the 32<sup>st</sup> Northeast Regional Stock Assessment Workshop (32<sup>st</sup> SAW): Stock Assessment Review Committee (SARC) consensus summary of assessments. Northeast Fish. Sci. Cent. Ref. Doc. 01-18.
- NEFSC. 2002. Longfin squid. *In*: Report of the 34<sup>th</sup> Northeast Regional Stock Assessment Workshop (34<sup>th</sup> SAW): Stock Assessment Review Committee (SARC) consensus summary of assessments. Northeast Fish. Sci. Cent. Ref. Doc. 02-07.
- Prager, M.H. 1994. A suite of extensions to a nonequilibrium surplus-production model. *Fish. Bull.* 92: 374-389.
- Quinn II, T.J., and R.B. Deriso. 1999. Quantitative fish dynamics. Oxford University Press, NY.
- Ricker, W. E. 1975. Computation and interpretation of biological statistics of fish populations. *Bull. Fish. Res. Bd. Can.* No. 191. 382p.
- Schaefer, M.B. 1957. A study of the dynamics of the fishery for yellowfin tuna in the eastern tropical Pacific Ocean. *Inter-Am. Trop. Tuna Comm. Bull.* 2: 245-285.
- Schnute, J. 1985. A general theory for analysis of catch and effort data. *Can. J. Fish. Aquat. Sci.* 42: 414-429.
- Sims, S.E. 1982. Algorithms for solving the catch equation forward and backward in time. *Can. J. Fish. Aquat. Sci.* 39: 197-202.
- Zhang, C. I., and P. J. Sullivan. 1988. Biomass-based cohort analysis that incorporates growth. *Trans. Am. Fish. Soc.* 117: 180-189.