Evaluating Two Assessment Methods for *Pandalus borealis* Stocks in the Northwest Atlantic

by

Steven X. Cadrin

National Marine Fisheries Service, Northeast Fisheries Science Center
166 Water Street, Woods Hole, MA, USA
Steven.Cadrin@NOAA.gov

Abstract

Analyses of simulated “shrimp-like” data with measurement errors, process errors, and specification errors demonstrate that a stage-based model (Catch-Survey) provides reliable estimates of stock abundance, and a biomass dynamics model (ASPIC) contributes accurate trends in stock biomass when catch and relative stock size indices are measured with moderate to high precision. However, accuracy of Catch-Survey estimates decreased when precision of input data was low (i.e., greater than 30% coefficient of variation), and estimates of stock size were frequently misleading. Despite bias in estimates of stock biomass from ASPIC, trends in stock size were relatively accurate.

Introduction

Lacking precise information on age and growth, some stock assessments of northern shrimp, *Pandalus borealis*, in the Northwest Atlantic have been based on models that integrate aggregate catch and stock size indices. Modal analysis of length distributions have had limited success in assessing the status of northern shrimp stocks (Hvingel and Savard, 1997). The most common alternatives to age-based or length-based assessment techniques are biomass dynamics models (e.g., Graham, 1935; Schaefer, 1954; Pella and Tomlinson 1969) and stage-based methods (e.g., Collie and Sissenwine, 1983; Schnute, 1987). Biomass dynamics models, commonly referred to as surplus production analyses, have been applied to northern shrimp fisheries off Iceland (Skúladóttir, 1979), in the Denmark Strait (Skúladóttir, 1985; Cadrin and Skúladóttir, 1998), in the Gulf of Maine (Cadrin *et al.* 1998, 1999) and off West Greenland (Hvingel and Kingsley 1998). Stage-based models incorporate information on recruitment, usually defined by a size range of shrimp that will recruit to the fishery, and have been used to assess northern shrimp stocks off Iceland (Stefánsson *et al.*, 1994) and in the Gulf of Maine (Cadrin *et al.* 1998, 1999).

When stock size is modeled as a forward projection through an observed time series, biomass dynamics methods (e.g., ASPIC; Prager, 1994) and stage-based methods (e.g., catch-survey, or C-S, model; Collie and Sissenwine 1983) essentially scale stock size indices according to their response to fishery removals by minimizing observation errors from an assumed population process. However, both approaches can provide misleading information on the scale of stock sizes if catchability of survey indices is poorly estimated. For example, ASPIC estimates of stock biomass for longfin inshore squid (*Loligo pealeii*) were less than area-swept biomass estimates from research surveys, which are considered to be minimum estimates (Cadrin and Hatfield, 1999). Similarly, C-S estimates of abundance for surfclams (*Spisula solidissima*; NEFSC, 1996) and sea scallops (*Placopecten magellanicus*; NEFSC, 1997) were also less than survey area-swept estimates, indicating that survey catchabilities were overestimated.
Simulation studies have been used to assess precision and bias of model estimates by emulating applications of C-S and ASPIC. For example, Prager et al. (1996) simulated age structured data for “swordfish-like” (*Xiphias gladius*) stocks and found that ASPIC generally provided reliable information on relative stock status despite specification error and measurement error. However, ASPIC analyses of other simulated data sets demonstrated that production models require a time series of catch and biomass indices with a relatively wide dynamic range of observations (NRC, 1998; Prager 1998). Collie and Kruse (1998) simulated “king crab-like” (*Paralithodes camtschaticus*) data to evaluate the performance of the C-S model and found that the model performed well on data with moderate measurement errors but was sensitive to some assumed parameters in the process equation. However, results of simulation studies are conditional on the specifics of data simulations and are difficult to generalize for other applications with different data characteristics.

The objective of this study was to evaluate the accuracy of ASPIC and C-S for “shrimp-like” simulated populations and fisheries under various assumptions about measurement error, process error, and specification error. The performance of analyses under simulated situations should help to understand the limitations of the approaches and their relative merits for assessing northern shrimp stocks.

Methods

**Catch-Survey stage-based model (C-S)**

The method described by Collie and Sissenwine (1983) was used to estimate abundance of recruits and fully-recruited shrimp. The model essentially fits a simple process equation to a time series of catch and relative abundance indices of two size classes of shrimp:

\[
N_{t+1} = [(N_t + R_t)e^{-0.5M} - C_t]e^{-0.5M}
\]

where \(t\) is an annual fishing season, and recruited northern shrimp \((N_t + R_t, \) fully recruited plus new recruits to the fishery\) experience a half-year of natural mortality \(e^{-0.5M}\), catch \((C, \) in numbers\) is removed, then the survivors from the fishery \([(N_t + R_t)e^{-0.5M} - C_t]\) experience another half-year of natural mortality. Similar to ASPIC, abundance is related to indices of relative abundance:

\[
n_t = q_n N_t e^{ht}
\]

and

\[
r_t = q_r R_t e^{dt}
\]

where \(r_t\) and \(n_t\) are observed survey indices of recruits and fully-recruited northern shrimp, \(q\) is catchability of the survey gear, and \(e^{ht}\) and \(e^{dt}\) are lognormally distributed measurement errors. The process equation is derived by substituting survey indices into equation 3 and including lognormally distributed process error \((e^{ct})\):

\[
n_{t+1} = [(n_t + r_t/s_t)e^{-0.5M} - q_n C_t]e^{-0.5M} e^{ct}
\]

where

\[
s_t = q_t / q_n
\]

Observation and process errors are minimized by iteratively solving for \(n_t, r_t,\) and \(q_n\). Similar to analyses of northern shrimp in the Gulf of Maine, \(M\) was assumed to be 0.25 and \(s_t\) was assumed to be 0.9 for analysis of simulated data.

**A Stock Production Model Incorporating Covariates (ASPIC)**

Software described by Prager (1994, 1995) was used to estimate biomass. In comparison to other production models, ASPIC has the desirable features of a being time series method (Hilborn and Walters 1992), an observation error estimator.
(Polachek et al., 1993), and is conditional on catch (Xiao 1997). The model assumes logistic population growth, in which the change in stock biomass over time \( \frac{dB_t}{dt} \) is a quadratic function of biomass \( B \) minus yield \( Y \), in weight:

\[
\frac{dB_t}{dt} = rB_t - \left( \frac{r}{K} \right)B_t^2 - Y_t
\]

where \( r \) is the instantaneous annual growth rate, and \( K \) is the carrying capacity. Relative biomass indices from either research surveys or catch-per-unit-effort (CPUE) of a fishing fleet are used to calibrate the predicted biomass trajectory. Stock biomass is related to indices of relative biomass:

\[
b_t = qB_t e^{\epsilon_t}
\]

where \( b_t \) is an observed index of biomass, \( q \) is the catchability of the index, and \( e^{\epsilon_t} \) is a lognormally distributed measurement error. The sum of squared measurement errors are minimized by iteratively solving for values of \( r, K \), and \( q \), and biomass in the first year.

**Simulated observation error**

To emulate the Gulf of Maine northern shrimp fishery, true abundance of recruits and fully recruited shrimp, total biomass, and catch were taken directly from the most recent stock assessment estimates (1984-1998; Clark et al., this volume). Observed abundance and biomass indices were generated by randomly sampling lognormal measurement errors assuming a 10% coefficient of variation (CV) which is appropriate for the Gulf of Maine shrimp survey (Cadrin et al., 1998, 1999).

\[
n_t = q_n N_t e^{\epsilon}
\]

\[
r_t = s_r q_n R_t e^{\epsilon}
\]

\[
b_t = q B_t e^{\epsilon}
\]

where \( q_n \) was 0.5 and \( s_r \) was 0.9. Error in observed catch (in numbers) and yield (in weight) was sampled from a normal distribution of error with a mean of 0 and a standard deviation of 50 mt (approximately 10% of observed catch):

\[
C_t = C_t + \epsilon
\]

\[
Y_t = Y_t + \epsilon
\]

**Simulated process error**

A fifteen-year time series of survey indices and catches for shrimp-like fisheries were generated using the Collie-Sissenwine process equation with random errors in \( M \) and \( s_r \). In the first two years, recruitment to the fishery was sampled from a distribution with a mean of 800 million and a CV of 50%. Fully-recruited abundance in the first year was sampled from a distribution with a mean of one billion and a CV of 0.5. A time series of \( F \) was sampled from a lognormal distribution with a mean of 0.5 and a coefficient of variation (CV) of 50%. A time series of observed catch was generated with a series of different precisions (10%, 20%, 30%, 40%, and 50% CV):

\[
C_t = F_t (N_t + 0.5R_t) \left[ 1 - e^{(F-M)} \right] (M+F)^{-1} + \epsilon
\]

Recruitment in years 2-15 was simulated as a linear function of lagged fully-recruited abundance (from Cadrin et al. 1998) with 50% CV around the predicted value:

\[
R_t = N_{t-2} \cdot 0.7 \cdot e^\epsilon
\]

Fully-recruited abundance in years 3-15 was generated with the C-S process equation and a CV of 50%.

\[
N_{t+1} = \left[ (N_t + R_t/s_r) e^{0.5Mr} - q_n C_t \right] e^{0.5Mr} \cdot e^\epsilon
\]

Catchability of fully-recruited shrimp \( q_n \) was assumed to be 0.5, and \( s_r \) varied randomly with a mean of 0.9 and a normally distributed 10% CV. Stochastic natural mortality \( Mr \) varied randomly with a mean of 0.25 and a normally distributed 10%
CV. A time series of observed abundance indices were simulated with a series of different precisions (10%, 20%, 30%, 40%, and 50% CV) as in equations 8 and 9.

For surplus production analyses, 30-year time series of survey indices and catches for shrimp-like fisheries were generated using the logistic growth equation with random errors in the parameters $r$ and $K$. Stock biomass in the first year was sampled from a distribution with a mean of 10,000 mt and a CV of 50%. Yield was calculated from simulated biomass and $F$, which was sampled from a lognormal distribution with a mean of 0.5 and a coefficient of variation (CV) of 50%. Total biomass in years 2-30 was simulated with the biomass dynamics process equation:

\[ B_{t+1} = r' B_t - \left(\frac{r}{K'}\right) B_t^2 - Y_t \]

Catchability was assumed to be 0.5, $r'$ varied randomly with a mean of 1.0 and a normally distributed 10% CV, and $K'$ varied randomly with a mean of 20,000 mt and a normally distributed 10% CV. A time series of observed biomass indices were simulated as in equation 10 with a CV of 10%.

**Simulated specification error**

Unlike simulated data with observation or process error, which were produced by the same model used to analyze the data (i.e., C-S or logistic growth), specification error was simulated with a different model. Survey indices and catches for shrimp-like fisheries were generated using a density-dependent, age-structured model based on a dynamic pool model of Gulf of Maine northern shrimp (Cadrin *et al.*, 1999). Abundance of age-0 shrimp in the first year was sampled from a distribution with a mean of 2 billion and a CV of 10%. Abundance of older ages ($N_2$ to $N_7$) in the first year were calculated as

\[ N_{a+1} = N_a e^{-M - pF} \]

where $F$ in each year was sampled from a lognormal distribution with a mean of 0.5 and a coefficient of variation (CV) of 50%. Partial recruitment at age ($p$) was determined from predicted length ($L$) at age (from Schick and Brown, 1997):

\[ p = 1/[1 + e^{(6.75 - 0.3007 L)}] \]

Length at age in the first year was derived directly from a von Bertalanffy growth function (McInnes, 1986).

Recruitment in subsequent years was derived as a function of female biomass (SSB) with lognormal error (10% CV) in the previous year:

\[ N_{0,t+1} = \frac{SSB/(0.0003 SSB + 3)}{e^{34.95 - 1.384 L}} \]

Weight at age was derived from length at age (Haynes and Wigley, 1969). Proportion female ($P$) at age is based on observed proportion female at length (Cadrin *et al.*, 1999):

\[ P = 1/[1 + e^{34.95 - 1.384 L}] \]

Abundance at older ages in years 2-40 were calculated as

\[ N_{a+1,t+1} = N_{a,t} e^{M - pF} \]

In years 2-40, $p$ was based on density dependent length at age ($L$ in equation 18) in which the predicted length at age was reduced as a function of stock abundance in the previous year, based on observed reductions in length at age during periods of high abundance (Clark *et al.*, this volume):

\[ L_{a,t} = 35.2 (1 - a e^{-0.36} - 0.00033(N_{a,t-1})) \]
For C-S analysis, abundance of fully-recruited shrimp (N) was calculated from the sum product of abundance at age and partial recruitment at age, and abundance of recruits (R) was calculated from the sum product of remaining abundance at age and end-of-the-year selectivity at age (Cadrin et al., 1999). Annual catch in numbers was calculated as

\[ C_t = F_t \left( p_r (R_t + N_{0,t}) \left(1 - e^{M-F_t}\right) / (M+F_t) \right) \]

where \( p_r \) was assumed to be 0.5. Indices of recruit and fully-recruited abundance were simulated as described in equations 8 and 9 with CVs of 10% to emulate survey indices of Gulf of Maine shrimp.

For surplus production analyses, stock biomass was calculated as the sum product of abundance at age and weight at age, and yield in weight was calculated as the sum product of catch at age and weight at age. Indices of stock biomass were simulated as described in equation 10 with a CV of 10%. Relative performance of methods was evaluated using relative root mean square error between estimates and true values from simulations.

**Results**

Catch-Survey analysis of simulated data based on estimated observation error reported for the Gulf of Maine provided a range of stock sizes that are similar to reported confidence intervals (Cadrin et al. 1999; Figure 1). Surplus production analysis data with simulated observation error provided similar results and confirm that estimates from the first several years are the least reliable (Prager, 1995; Figure 2).

Catch-Survey analysis of data with simulated observation and process error shows that accuracy of stock size estimates and relative trends in stock size increase with precision of input data (Table 1; Figure 3). At 10% CV in catch and survey indices (approximately the level of imprecision in the Gulf of Maine shrimp assessment) all abundance estimates were relatively accurate (i.e., within 10% of the ‘true’ value). At 20% CV in catch and survey indices, estimates of abundance were frequently inaccurate (e.g., average relative errors were >10% for 50% of simulations, and >25% for 13% of simulations), however all relative trends were accurate. At 30% CV in input data, relative errors were similar, but some relative trends in estimates were misleading. At 40% CV in input data, relative errors in abundance estimates were frequently greater than 50%, and relative trends were misleading for many simulations. For results based on higher levels of imprecision in input data (i.e., CV>20%) it appears that the most inaccurate estimates were from simulations with narrow ranges of survey observations (Figure 4). For example, all simulations with an average relative error greater than 30% had a relatively restricted range of observed abundance indices (CV<35%), whereas all simulations with broader range had relative errors of 20% or less. Estimates of bias were not significantly different than zero and were inconsistent among different levels of variance (i.e., bias was slightly positive at 10% CV, slightly negative at 20% CV, slightly positive at 30% CV, and slightly negative at 40% CV).

Surplus production analysis of data with simulated observation and process error were only moderately accurate. The relative error of biomass estimates was 25%, and 20% of trials had greater than 50% relative error (Figure 5). Analysis of data with greater levels of observation error (e.g., 20%, 30%, and 40%) had frequent problems with convergence.

Catch-Survey analysis of data with simulated specification error shows that estimates are relatively accurate and unbiased (Figure 6; relative errors of abundance estimates averaged 14% when 10% observation error was simulated). However, ASPIC estimates of biomass from data with simulated specification error were significantly biased low (Figure 7). Despite unreliable estimates of catchability, relative biomass trends from ASPIC were accurate.

**Discussion**

Although results from analyses of data with simulated observation error may seem trivial because the same model and parameters were used to simulate the data, comparison with reported variance estimates (Cadrin et al. 1999) serve as a reminder that conventional estimates of variance from asymptotic theory, jackknifing (Hvingel and Kingsley, 1999), or bootstrapping (Cadrin et al. 1999) only encompass observation error and do not include other sources of uncertainty.

Inaccuracy in estimates from data with lower precision stresses the need to routinely evaluate the variance of catch and survey biomass estimates. Results from different levels of observation and process error are somewhat consistent with those from simulated king crab data in that relative error of estimates increases at greater levels of input error (Collie and
Kruse 1998). However, unlike the results from king crab data, the present results do not indicate an increase in bias as observation and process error increase. Similar to surplus production models (Hilborn and Walters, 1992; Prager et al., 1996), it appears that estimates from C-S are more reliable when there is more contrast in the data.

Moderate inaccuracies in ASPIC estimates of stock biomass from data with simulated process error is consistent with simulation the results of Ludwig et al. (1988). The present results confirm that estimates of $q$ from ASPIC can be unreliable, but also illustrate that relative trends in biomass may be useful (Prager 1994). The bias in ASPIC estimates from data with specification error probably result from the age-based model not conforming to logistic growth. However, components of the age-based simulations (e.g., stock-recruit data) are based on very limited information and may not accurately reflect the dynamics of Gulf of Maine northern shrimp.

Specification error does not appear to be a problem with C-S, presumably because the underlying process equation is empirical, not theoretical (i.e., recruitment is observed rather than implicitly predicted from a population growth model). The present analysis suggests that the C-S application to shrimp-like data is robust to interannual variability in length at age, transition at age, and selectivity at age. Although the process equation is relatively robust to different population dynamics, C-S results are sensitive to incorrect model assumptions (e.g., natural mortality, commercial selectivity, relative survey catchability of recruits and recruited shrimp). The present study simulated imprecision in assumed C-S process parameters, but did not simulate incorrect assumptions (i.e., bias in process parameters). Simulation can be used to assess concerns about model assumptions, such as temporal changes in selectivity (Prager et al. 1996) and different survey catchabilities of R and N (Collie and Kruse, 1998). Because these results are conditional on the specifics of the simulation methods (e.g., number of years, level of F, assumed levels of imprecision) they may not apply to all applications of either C-S or ASPIC. Therefore, analysis of simulated data which emulates a specific application should be a routine method for evaluating stock assessment methods.

Analysis of data that simulated the approximate levels of process and observation errors that represent data for the Gulf of Maine northern shrimp fishery indicates that abundance estimates from the C-S model and relative trends in biomass from ASPIC are relatively accurate. However, unless ASPIC biomass estimates are confirmed by another method (e.g., C-S biomass estimates of Gulf of Maine shrimp; Cadrin et al., 1999), assessments should be limited to relative trends in biomass (e.g., Denmark Strait shrimp; Cadrin and Skúladóttir, 1998).

Acknowledgments

I am grateful to Mike Prager for providing and supporting ASPIC software and his continued guidance. This study also benefitted from discussions with Steve Clark, Han-Lin Lai, Jim Weinberg, and Larry Jacobson.

References


Table 1. Relative error (root mean square error / true value) in estimates of survey catchability (q), fully-recruited abundance (N), and recruit abundance (R) from Catch-Survey analysis of simulated shrimp-like data at various levels of observation and process error (10%, 20%, 30% and 40% coefficient of variation, CV).

<table>
<thead>
<tr>
<th>CV</th>
<th>q</th>
<th>N</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>10%</td>
<td>0.09</td>
<td>0.08</td>
<td>0.06</td>
</tr>
<tr>
<td>20%</td>
<td>0.12</td>
<td>0.15</td>
<td>0.12</td>
</tr>
<tr>
<td>30%</td>
<td>0.19</td>
<td>0.25</td>
<td>0.16</td>
</tr>
<tr>
<td>40%</td>
<td>0.20</td>
<td>0.35</td>
<td>0.25</td>
</tr>
</tbody>
</table>
Figure 1. True values (bold line) and Catch-Survey estimates (thin lines) of recruit and fully-recruited abundance based on simulated shrimp-like data with 10% observation error.
Figure 2. True values (bold lines) ASPIC estimates (thin lines) of stock biomass based on simulated shrimp-like data with 10% observation error.
Figure 3a. True values and Catch-Survey estimates of fully-recruited abundance based on simulated shrimp-like data with 10% observation and process error.
Figure 3b. True values and Catch-Survey estimates of fully-recruited abundance based on simulated shrimp-like data with 40% observation and process error.
Figure 4. Relative error in Catch-Survey estimates of catchability (root mean square error / true value) and interannual variability of survey indices of fully-recruited shrimp at two levels of observation and process error (30% and 40% coefficient of variation).
Figure 5. True values and ASPIC estimates of stock biomass based on simulated shrimp-like data with 10% observation and process error.
Figure 6. True values and Catch-Survey estimates of fully-recruited abundance based on simulated shrimp-like data with specification error.
Figure 7. True values and ASPIC estimates of stock biomass based on simulated shrimp-like data with specification error.