

# Three-Stage Catch-Survey Analysis Applied to Blue King Crabs

**Jeremy S. Collie and Allison K. DeLong**

*Graduate School of Oceanography, University of Rhode Island,  
Narragansett, Rhode Island*

**Gordon H. Kruse**

*School of Fisheries and Ocean Sciences, Juneau Center, University of  
Alaska Fairbanks, Juneau, Alaska*

## Abstract

Catch-survey analysis (CSA) is a stock assessment method that is applicable in data-limited situations. It requires time series of catch and relative abundance divided into two categories: recruits and post-recruits. This method is particularly useful for invertebrate populations (e.g., crabs) that cannot be aged but can be classified by their size distributions. In this study we extended the existing two-stage CSA to include a third stage, pre-recruits. An additional equation describes the transition of pre-recruits to the recruit stage as a function of survival and molt probability. We applied the three-stage CSA to two blue king crab (*Paralithodes platypus*) stocks: St. Matthew Island and the Pribilof Islands in the eastern Bering Sea. Sensitivity analyses and Monte Carlo tests were performed on data from the St. Matthew Island stock. The final version of the model includes bootstrapped confidence intervals on the estimated population abundances. Abundance of both stocks declined following high harvest rates in the early 1980s and the fisheries have been closed since 1999. Three-stage catch-survey analysis has several advantages over the two-stage CSA. It uses more data and therefore has a higher ratio of observations to parameters. Three-stage CSA also provides estimates of mature and legal male crabs, both of which are used for management. By accounting for observation errors for an additional stage, it provides more accurate estimation of abundance in the most recent year, which is critical for setting harvest guidelines.

## Introduction

Many fisheries agencies conduct standardized research surveys of relative abundance as a cornerstone of stock assessments of exploited fish and shellfish populations. To convert relative to absolute estimates of population size, survey gear must be calibrated with a survey catchability coefficient,  $q$ , defined as the proportion of the population sampled with one unit of survey effort. A limitation of research surveys is the requirement to assess vast areas with a limited number of observations. Random sampling can ensure unbiased abundance estimates, but can't eliminate large observation errors that arise from patchy spatial distributions and other chance elements of sampling.

Collie and Sissenwine (1983) developed catch-survey analysis (CSA) as one of the first stock-assessment methods to calibrate fishery-independent surveys while accounting for observation errors. This method was designed for populations with limited age information, either because of the difficulty or high cost of age determination. It requires only that the population be divided into two categories, recruits and post-recruits, which can often be distinguished by their size distributions. The size or age distribution of the commercial catch is not required, though it can be used to calculate the selectivity of the fishing gear. This simple formulation produces abundance estimates in close agreement with those of age-structured methods (Collie and Sissenwine 1983), length-based models (Zheng et al. 1996), and production models (Cadrin et al. 1999).

Catch-survey analysis has been applied to a variety of demersal fish (Collie and Sissenwine 1983), mollusk (Conser 1991), and crustacean species (Conser and Idoine 1992, Cadrin et al. 1999). It is particularly well suited to invertebrate species for which routine age determinations do not exist. If gear size selectivity is known, it can be used to partition survey catches into recruit and post-recruit categories (Cadrin et al. 1999). King crabs are particularly good candidates for CSA because, as adults, they molt, at most, once per year. Although absolute age determination is impossible, legal-sized crabs can be classified into recruits (legal size for less than one year) and post-recruits (legal size for more than one year) based on the size and relative age (i.e., new-shell/old-shell) of their carapace. On this basis, Collie and Kruse (1998) applied CSA to two red king crab (*Paralithodes camtschaticus*) stocks in Alaska.

The original catch-survey analysis estimated process errors (also known as equation errors) and observation errors (also known as measurement errors) (Collie and Sissenwine 1983). Process errors arise from mis-specifying the population dynamics equation; for example, variations in a natural mortality rate,  $M$  and  $q$  that are assumed constant. However, the mixed-error formulation requires the estimation of unobserved, latent variables, which increases the risk of overparameterization and non-convergence to a unique solution. Several studies since then included

simulations to test alternate model formulations and the sensitivity of the method to the magnitude of errors commonly encountered in survey data. A general result is that  $q$  can be estimated with low bias (<8%) with realistic levels of observation error in the survey data (Collie and Kruse 1998). Estimates of  $q$  are highly dependent on  $M$ , which must therefore be specified as an input parameter.

A version of CSA that assumed only observation errors, ignoring process errors, performed well when tested with data containing various levels of correlated and uncorrelated observation errors, and low levels of process error (Collie and Kruse 1998). With fewer parameters to estimate (no latent variables), this version reduces the risk of overparameterization. With high levels of process error (higher variance than in most real data sets) the all-observation error model resulted in more highly biased parameter estimates and some Monte Carlo trials failed to converge. Derived quantities of interest, such as the abundance of legal male crabs, were estimated with low bias (<4%). Monte Carlo simulations provided bootstrapped confidence intervals for the quantities used in decision making (e.g., absolute abundance, harvest rate).

Further sensitivity analyses have been conducted with data generated from a known fully age-structured model (Cadrin 2000, Mesnil 2003). Absolute abundance estimates are sensitive to relative catchability of one stage to the other, which must be specified externally. If this relative catchability is size dependent, its value can vary with time due to changes in the age structure of the population caused by fishing and/or recruitment variability (Mesnil 2003). Erroneously treating this relative catchability as constant can therefore introduce process errors in the estimation. Notwithstanding these potential sources of error, Mesnil (2003) found CSA to be "a very valuable method to support management advice in data limited contexts, in which age data are lacking or uncertain."

Populations of the blue king crab (*Paralithodes platypus*) occur in the eastern Bering Sea near St. Matthew Island and the Pribilof Islands. As there is no evidence of migration between the two populations, they are assessed and managed as two separate stocks (Otto and Cummiskey 1990). Whereas juveniles strongly prefer shallow-water habitats with shell hash that provides refuge from predators (Armstrong et al. 1985), adults are generally thought to occupy soft bottoms (mud-sand) adjacent to coarse bottoms (gravel-cobble-boulder-rock) at depths of 48-75 m.

Although subadult and adult males from both areas have the same mean growth increment (~14 mm carapace length, CL), molting probability declines sharply after attaining size at maturity, which is 105 mm CL for St. Matthew Island and 120 mm CL for the Pribilof Islands. Likewise, the mean size of landed blue king crabs is much larger for the fishery off the Pribilof Islands (3.5-3.6 kg) than off St. Matthew Island (1.8-2.2 kg). Harvest guidelines are based on a harvest rate applied to abundance estimates of mature male crabs, but only legal males ( $\geq 120$

mm CL for St. Matthew Island and  $\geq 135$  mm CL for the Pribilof Islands) may be retained.

In the past, stock assessments were not based on population models; instead crab abundance was estimated by the area-swept method in which survey catches were expanded from the area swept by trawl tows to the entire region encompassed by the stratified systematic survey design (Stevens et al. 2002). Catch-survey analysis is an attractive assessment tool for blue king crabs because there is limited length-frequency data to support more data-intensive methods. A two-stage CSA has been used to assess the St. Matthew Island and Pribilof Islands blue king crab stocks (Zheng et al. 1997). The estimates of survey catchability were uncertain, partly because crab habitats around these islands include rugged, untrawlable bottoms with uncertain crab densities. To address this uncertainty, Zheng et al. (1998) incorporated an index of commercial catch per unit effort into the CSA, because commercial crab pots (traps) are deployed over a wider range of bottom types than survey trawls. As a result, the spatial distributions of the survey and the fishery do not fully overlap.

The objective of this study was to improve the performance of CSA by refining the estimate of legal blue king crab abundance and by including the estimation of mature crab abundance in the analysis. Though previous applications were promising, the output of CSA was not ideally suited to crab management, because mature male crabs were not included in the analysis. Mature crabs include pre-recruit males that have not yet molted into the legal size range. We therefore added a pre-recruit stage to the standard two-stage CSA. A three-stage CSA would make more complete use of the survey data but would also require the specification of additional parameters associated with the pre-recruit stage. In this study we develop a three-stage version of CSA and apply it to the estimation of blue king crab abundance. Monte Carlo simulations are used to evaluate the performance of the three- versus two-stage CSA.

## Methods

### *Model formulation*

The three-stage catch-survey analysis uses a simple population model to calibrate relative abundance indices from research surveys with the commercial catch in numbers  $C_t$ . The procedure estimates  $q$ , which can then be used to estimate a time series of absolute population abundance. The stages are defined as pre-recruit  $N$ , recruit  $R$ , and post-recruit  $P$ . As in the two-stage model, the recruit category is defined such that all individuals become post-recruits the following year and the post-recruit category is a “plus group” (Collie and Kruse 1998). The population model for true (absolute) abundance is

$$P_{t+1} = (R_t + P_t)e^{-M_t} - C_t e^{-M_t(1-T_t)},$$

where  $P_t$  and  $R_t$  are the respective post-recruit and recruit numerical abundances in year  $t$ ,  $M_t$  is the instantaneous natural mortality rate, and  $T_t$  is the time lag from the survey to the midpoint of the fishery. Most of the variability in the time lag is due to historical changes in fishery start dates and duration. Recruit abundance is a function of pre-recruit abundance  $N_p$  in the previous year:

$$R_{t+1} = mN_t e^{-M_t},$$

where  $m$  is the transition probability between the pre-recruit and recruit stages. Pre-recruits that don't molt remain as pre-recruits the following year, which adds the constraint

$$N_{t+1} \geq (1-m)N_t e^{-M_t}.$$

Only the male component of the population is modeled, because females cannot be legally retained by the fishery. Size limits were set to allow males one opportunity to mate before achieving legal size. So, the pre-recruit class includes mature males that are one molt increment shy of recruitment into the commercial fishery. The transition probability  $m$  is the probability that a pre-recruit crab will molt in a given year. Recruits are defined as crabs that entered the legal size class in that year. This determination is based on two criteria: they are "new-shell" meaning they molted less than one year before being caught in the survey and they are less than one molt increment greater than minimum legal size. The post-recruit size class includes all crabs greater than legal size that are not new recruits (Collie and Kruse 1998). Owing to differences in growth and maturity schedules between St. Matthew Island and the Pribilof Islands, size classes differ for each stock as defined in the next section.

As in the two-stage CSA, true (unknown) relative abundance can be written as the product of a catchability coefficient and absolute abundance:

$$n_t = q \theta N_p$$

$$r_t = q \phi R_p \text{ and}$$

$$P_t = q p_p$$

where  $n_t$ ,  $r_t$  and  $p_t$  are the respective pre-recruit, recruit, and post-recruit relative abundances in year  $t$ ,  $\phi$  is the catchability of recruits relative to post-recruits,  $\theta$  is the catchability of pre-recruits relative to post-recruits,

and  $q$  is the survey catchability coefficient for post-recruit crabs. Substituting these relationships into the population equations and solving for post-recruit and recruit relative abundance results in:

$$p_{t+1} = (r_t / \phi + p_t) e^{-M_t} - q C_t e^{-M_t(1-T_t)}$$

$$r_{t+1} = m n_t \frac{\phi}{\theta} e^{-M_t}.$$

Observed (survey) relative abundance indices, denoted below with the tilde ( $\sim$ ), are assumed to have lognormally distributed observation errors such that their logarithms can be written as:

$$\ln(\tilde{n}_t) = \ln(n_t) + \gamma_t,$$

$$\ln(\tilde{r}_t) = \ln(r_t) + \kappa_t, \text{ and}$$

$$\ln(\tilde{p}_t) = \ln(p_t) + \varepsilon_t,$$

where  $\gamma_t$ ,  $\kappa_t$ , and  $\varepsilon_t$  are either correlated or uncorrelated normal deviates (Collie and Sissenwine 1983). The model parameters are estimated by minimizing the sum of squared differences between the observed and predicted relative abundance indices, with the objective function:

$$\begin{aligned} SSQ = & \sum_{t=1}^{y-1} w_{pre}^2 [\ln(\tilde{n}_t + \delta) - \ln(n_t + \delta)]^2 + \sum_{t=1}^y w_{rec}^2 [\ln(\tilde{r}_t + \delta) - \ln(r_t + \delta)]^2 \\ & + \sum_{t=1}^y [\ln(\tilde{p}_t + \delta) - \ln(p_t + \delta)]^2, \end{aligned}$$

where  $\delta$  is a small constant added to ensure the equation remains defined,  $w_{pre}$  and  $w_{rec}$  are the weights of pre-recruits and recruit errors relative to post-recruit errors, and  $y$  is the number of years of data. The parameters  $\phi$ ,  $M_t$ , and  $T_t$  are assumed to be known independently; the model estimates  $\theta$ ,  $q$ ,  $r_1$ ,  $p_1$ , and pre-recruit relative abundance  $n_t$ , for each year but the last. When fit to  $y$  years of data, the model has  $y + 3$  parameters to estimate,  $3y - 1$  residual errors, and  $2y - 4$  degrees of freedom.

Estimated relative abundances are converted to absolute population estimates using

$$N_t = \frac{n_t}{q\theta}, \quad R_t = \frac{r_t}{q\phi}, \quad \text{and} \quad P_t = \frac{p_t}{q}.$$

The number of mature crabs each year is estimated as  $N_t + R_t + P_t$ , and the number of legal crabs each year is  $R_t + P_t$ . The three-stage formulation allows legal crab abundance to be estimated in all years and mature crab abundance to be estimated in all years but the last, for which the

observed pre-recruit index must be used. Historical harvest rates  $h$  are calculated relative to both legal and mature crab abundance:

$$h(\text{legal})_t = \frac{C_t}{(R_t + P_t)e^{-M_t T_t}}$$

$$h(\text{mature})_t = \frac{C_t}{(N_t + R_t + P_t)e^{-M_t T_t}}.$$

### **Data sources**

Abundance indices for blue king crabs were derived from the National Marine Fisheries Service (NMFS) eastern Bering Sea trawl survey (Stevens et al. 2002). This fishery-independent bottom-trawl survey samples a fixed 20 nm by 20 nm grid of stations to monitor the abundance of crabs and demersal fish. Survey strata adjacent to St. Matthew Island and the Pribilof Islands were used by NMFS to derive area-weighted estimates of blue king crab abundance. Though these survey data have been calibrated for the effective width of the trawl using measurements from electronic sensors, here we treat them as relative abundance estimates. Commercial catch data and timing of the commercial catch came from Alaska Department of Fish and Game (ADFG) annual management reports. We set  $M$  equal to 0.3, typical of assessments for these stocks (Zheng et al. 1997).

Input data files include time series of pre-recruit, recruit, and post-recruit relative abundance (survey) indices, commercial catch in numbers, timing of the commercial catch for each year the fishery was open, estimates of  $M$  for pre-recruits and recruits/post-recruits, and pre-recruit molt probability (Table 1). The pre-recruit and recruit size classes were 105 to 119 mm and 120 to 133 mm for the St. Matthew stock and 120 to 134 and 135 to 148 mm for the Pribilof stock, respectively. The recruit size class includes only new-shell crabs, such that all recruits become post-recruits the following year. For simplicity we used only the trawl-survey data and did not include auxiliary catch-per-unit-effort data.

Initial development of the three-stage model was based on data from 1980 to 1997 and initial baseline runs and simulation tests were conducted with this 18-year time series. For the final runs reported in this paper, we updated the data sets to 2003 and included several years of pre-1980 data that recently became available. For this analysis, we fixed  $\phi = 1.0$ , because the survey trawl catches recruit and larger crabs equally well. For the baseline runs the offset parameter in the sum of squared residuals term  $\delta$  was set to 0.001. All error weights ( $w$ ) were fixed at 1, though in principle they could be estimated from the within-year CV of the survey data.

**Table 1. Input files for the three-stage catch-survey analysis.****(a) St. Matthew Island blue king crabs**

Year	Abundance indices			Commercial catch (millions)	Timing of commercial catch	Mature harvest rates
	Pre-recruit	Recruit	Post-recruit			
1978	2.175	1.187	0.587	0.436	0.068	0.115
1979	1.793	1.455	0.333	0.053	0.060	0.011
1980	2.588	1.699	1.197	0.033	0.068	0.005
1981	1.480	1.195	1.648	1.046	0.052	0.133
1982	2.615	3.617	3.263	1.936	0.068	0.274
1983	1.639	1.399	1.956	1.932	0.121	0.405
1984	0.500	0.788	0.762	0.841	0.096	0.361
1985	0.431	0.541	0.708	0.485	0.137	0.350
1986	0.425	0.164	0.185	0.220	0.137	0.148
1987	0.757	0.492	0.292	0.235	0.137	0.117
1988	0.703	0.417	0.411	0.302	0.137	0.120
1989	1.235	0.940	0.954	0.248	0.134	0.081
1990	0.957	0.954	1.164	0.391	0.140	0.104
1991	1.636	1.353	0.889	0.727	0.178	0.161
1992	1.582	1.338	1.247	0.545	0.142	0.125
1993	1.994	1.605	2.000	0.630	0.178	0.138
1994	1.350	1.246	1.120	0.827	0.178	0.197
1995	1.321	0.993	0.902	0.667	0.178	0.153
1996	1.970	1.950	1.331	0.661	0.178	0.150
1997	2.319	2.213	1.853	0.940	0.178	0.254
1998	1.843	1.397	1.766	0.612	0.178	0.265
1999	0.215	0.179	0.436	0.000	NA	0.000
2000	0.310	0.323	0.488	0.000	NA	0.000
2001	0.527	0.345	0.681	0.000	NA	0.000
2002	0.266	0.186	0.576	0.000	NA	0.000
2003	0.330	0.188	0.379	0.000	NA	NA



**Table 1. (Continued.) Input files for the three-stage catch-survey analysis.****(b) Pribilof Islands blue king crabs**

Year	Abundance indices			Commercial catch (millions)	Timing of commercial catch	Mature harvest rates
	Pre-recruit	Recruit	Post-recruit			
1975	6.201	4.144	4.873	0.315	0.200	0.026
1976	0.647	1.412	2.147	0.856	0.200	0.084
1977	1.275	3.379	8.472	0.807	0.200	0.093
1978	3.525	1.684	3.918	0.797	0.250	0.087
1979	1.075	1.285	2.282	0.816	0.250	0.111
1980	1.158	1.284	3.767	1.497	0.252	0.259
1981	0.716	0.686	3.170	1.202	0.249	0.321
1982	0.299	0.596	1.657	0.588	0.175	0.281
1983	0.597	0.359	1.037	0.276	0.173	0.181
1984	0.215	0.218	0.335	0.040	0.142	0.037
1985	0.155	0.138	0.227	0.078	0.192	0.083
1986	0.018	0.098	0.394	0.037	0.233	0.057
1987	0.071	0.068	0.731	0.095	0.274	0.182
1988	0.000	0.034	0.164	0.000	0.000	0.000
1989	0.000	0.000	0.211	0.000	NA	0.000
1990	0.658	0.234	0.094	0.000	NA	0.000
1991	0.754	0.577	0.584	0.000	NA	0.000
1992	0.730	0.374	0.741	0.000	NA	0.000
1993	0.640	0.318	0.658	0.000	NA	0.000
1994	0.314	0.183	0.651	0.000	NA	0.000
1995	0.842	0.332	1.669	0.174	0.175	0.108
1996	0.938	0.145	1.071	0.124	0.175	0.076
1997	0.261	0.282	0.540	0.069	0.175	0.056
1998	0.192	0.099	0.739	0.069	0.175	0.067
1999	0.145	0.105	0.347	0.000	NA	0.000
2000	0.190	0.017	0.531	0.000	NA	0.000
2001	0.068	0.000	0.404	0.000	NA	0.000
2002	0.018	0.033	0.161	0.000	NA	0.000
2003	0.016	0.082	0.218	0.000	NA	NA

### **Estimation of molt probability**

Blue king crab pre-recruit molt probabilities were estimated from analysis of tag-recapture data collected by NMFS on the St. Matthew Island and Pribilof Islands blue king crab stocks and by ADFG on the St. Matthew Island stock (Blau 1996). A logistic equation that describes molt probability as a function of carapace length was fitted to these data:

$$prob_L = \frac{1}{1 + ae^{bL}} ,$$

where  $prob_L$  is the probability a crab of carapace length  $L$  will molt in any given year and  $a$  and  $b$  are parameters. Carapace length is the distance from the posterior margin of the right eye orbit to the posterior carapace margin. Logistic regression is the appropriate statistical model for these recapture data because each data point is weighted equally and the molt probability is a binomially distributed random variable (the crab did or did not molt), with an expected molt probability associated with the crab's release length (Neter et al. 1996). The logit transformation of the logistic molt probability function is

$$\ln\left(\frac{prob_L}{1 - prob_L}\right) = \ln\left(\frac{1}{a}\right) - bL .$$

We estimated the pre-recruit molt probability  $m$  by averaging the fitted equation over the length of the pre-recruit size-class for each stock.

For the Pribilof Islands stock, we followed the procedure of Otto and Cummiskey (1990). There were 188 recaptures of crabs that were at large between 1 and 1.5 years, which ensures the crabs had the opportunity to molt once and only once. A crab was presumed to have molted if the carapace length increased more than 7 mm and to have not molted if the recapture carapace length was within 3 mm of the release length. Four crabs with growth increments between 4 and 7 mm inclusive were excluded from the analysis as it could not be determined whether these crabs molted or their shells were measured with error.

For the St. Matthew Island stock, we pooled the NMFS and ADFG data and used recapture information from crabs that were at large between 1 and 1.5 years and between 2 and 2.5 years. As there were only 81 recaptures between 1 and 1.5 years, we included data from 22 crabs at large between 2 and 2.5 years to increase the sample size. We determined whether the crabs at large between 2 and 2.5 years molted in the first and/or second year by considering their growth increment and shell condition at recapture by the following logic. None of the crabs in this sample had carapace length increases of 7 mm or less. The aver-

age growth per molt for the St. Matthew Island stock follows a normal distribution with an approximate mean of 14.1 mm and variance of 9.6 mm (Otto and Cummiskey 1990). The symmetric interval from 8 to 20 mm about the mean approximates the 95% confidence interval for mean growth per molt. So, we designated crabs that molted once as those whose carapace length increased by 20 mm or less. Of these crabs, those recaptured with old shells were assumed to have molted in the first year and those with new shells in year two. Those recaptured crabs whose carapace length increased by more than 20 mm were assumed to have molted twice, regardless of the recapture shell condition. Only the first-year molt frequencies were used to calculate molt probabilities, so that each crab would be counted only once.

### ***Simulation tests***

We tested the sensitivity of the three-stage model to errors in input parameters and compared the bias and precision of the three-stage and two-stage model formulations with data from the years 1980 to 1997 for the St. Matthew Island stock only. Two types of analyses were performed. First, we looked at the sensitivity of the model's output parameters  $\theta$  and  $q$  to mis-specification of the input parameters, including pre-recruit and recruit catchability, pre-recruit mortality, and pre-recruit molt probability. Collie and Kruse (1998) performed an analogous exercise with the two-stage model.

Second, we performed Monte Carlo simulations of the three-stage model. The Fortran program (a) starts with a data set that fits the model perfectly (i.e., model estimates of relative abundance output from the baseline run), (b) accepts as input, the standard deviation and correlation of observation errors, and the standard deviation of process errors, (c) simulates the uncorrelated process errors and assumes these new relative abundance indices represent the true population, (d) simulates correlated or uncorrelated observation errors on the true abundance indices from step (c), (e) fits the simulated observed data with the three-stage model, and (f) repeats the simulation as many times as desired. Identical Monte Carlo simulations with the two-stage model were performed for comparison. The simulated data had the same state dynamics expected by the estimation method, except for the addition of process errors and correlated observation errors.

Process errors exist when the true population dynamics are not adequately described by the population model. We simulated the process errors by applying lognormally distributed errors to the input (fitted) survey indices with the following calculations, assuming that the new abundance indices represent the true population:

$$r_{t+1} = \left[ n_t m \frac{\phi}{\theta} e^{-M_t} \right] e^{z\sigma - \sigma^2/2},$$

$$p_{t+1} = \left[ (r_t / \phi + p_t) e^{-M_t} - q C_t e^{-M_t(1-T_t)} \right] e^{z\sigma - \sigma^2/2},$$

where  $z$  is a random standard normal deviate,  $\sigma$  is the standard deviation, and  $\sigma^2$  the variance of the process errors. In the simulations examining process error only, these simulated “true” relative abundance indices are equivalent to the observed indices.

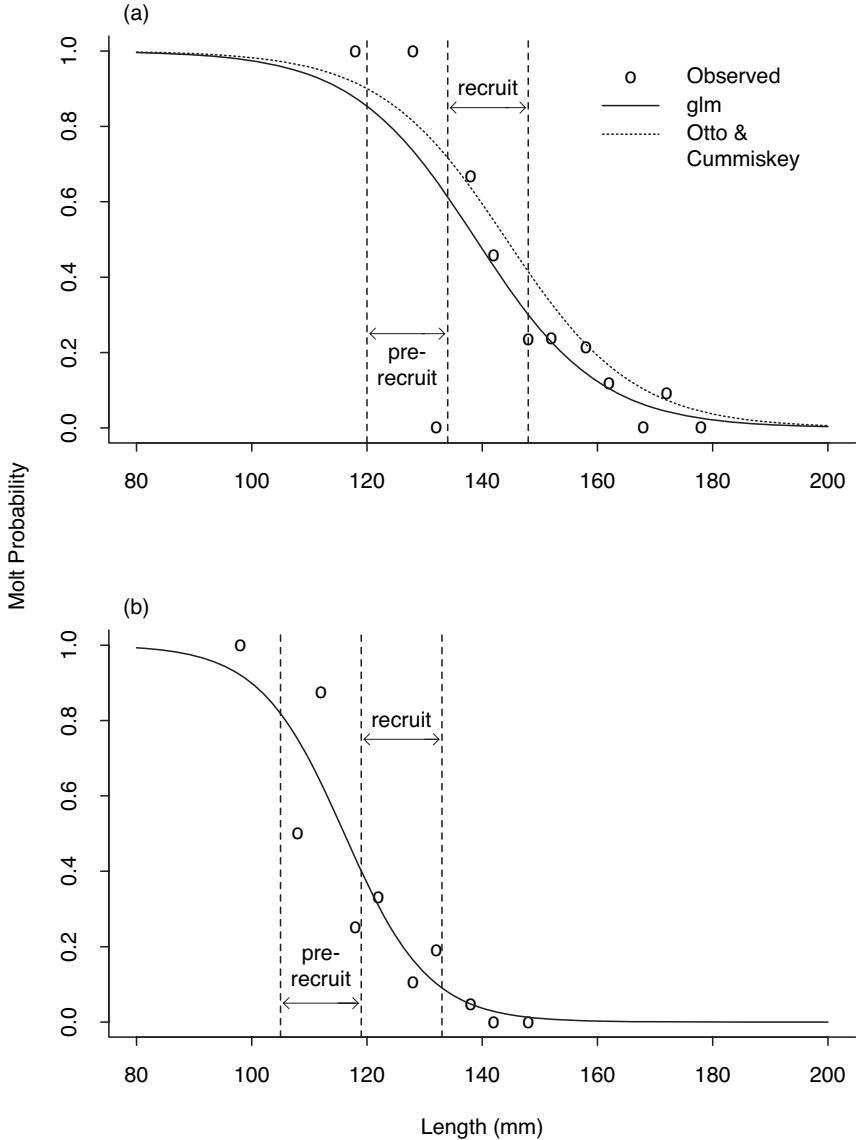
Correlated observation errors were placed on these indices with the following methodology. The  $3 \times 3$  matrix  $V$  contains the input variances and covariances, which in this case were assumed to be the same for all stages. If  $z$  is a vector of independent standard normal deviates, and  $L$  is the lower triangular matrix from the Cholesky decomposition of  $v$ , then  $l = Lz$  is a vector of cross-correlated random deviates (Kennedy and Gentle 1980). In each scenario, we simulated 1,000 replicates with the following error structures: (1) uncorrelated observation errors with standard deviation of 0.3, 0.5, or 0.7, (2) correlated observation errors with standard deviations as in (1) and a correlation of 0.5, and (3) a combination of correlated and uncorrelated observation errors and process error standard deviations of 0.1 and 0.3.

Confidence intervals on the estimates of population abundance for both stocks were estimated by parametric bootstrapping. The bootstrapping program first fit the input data to the model and calculated the standard deviation of the residuals. Unless otherwise specified (as in the Pribilof Islands simulation), the bootstrap used this standard deviation to simulate lognormally distributed errors on the relative abundances calculated in the initial run. Observation errors were simulated in the same manner as in the Monte Carlo analysis with uncorrelated errors. We performed 1,000 replicates and calculated the mean, standard deviation, and 95% empirical confidence intervals of the bootstrapped fits.

## Results

### *Molt probabilities*

Chi-square and deviance goodness of fit tests indicated that both of these data sets did not significantly differ from the logistic model (Neter et al. 1996). The logistic regression for the Pribilof Islands stock resulted in a pre-recruit molt probability of 0.75, which was used in all subsequent analyses (Fig. 1a). Using Otto and Cummiskey's (1990) parameter estimates resulted in a pre-recruit molt probability of 0.82. For the St. Matthew stock, the logistic regression resulted in a molt probability of 0.63, which was used for all the runs presented here (Fig. 1b). The reduced data set, consisting of 81 crabs at liberty between 1 and 1.5 years, resulted in



**Figure 1. Comparison of molt probability curves for the Pribilof Islands (a) and St. Matthew Island (b) blue king crab populations. Vertical lines indicate range of the pre-recruit and recruit size classes for each population. The solid curves are the generalized linear model fits (glm) to the binomial data with logit transformation.**

**Table 2. Residual analyses of the baseline runs.**

Statistic	St. Matthew	Pribilof Islands <sup>a</sup>
Number of years	18	18
Catchability coefficient ( $q$ )	1.105	1.405
Standard error of $q$	0.363	0.857
Catchability of pre-recruits ( $\theta$ )	0.516	0.625
Standard error of $\theta$	0.357	0.900
Sum of squared residuals	4.063	18.070
Standard deviation	0.356	0.751
Prob of normal distribution	0.925	0.055
Pre-recruit to recruit		
Correlation	0.702	-0.118
Prob correlation = 0	0.002	0.674
Prob of equal variances	0.955	0.871
Correlation (pre-recruit lag-1)	-0.585	-0.930
Prob correlation = 0	0.014	0.000
Pre-recruit to post-recruit		
Correlation	0.588	0.454
Prob correlation = 0	0.013	0.089
Prob of equal variances	0.834	0.114
Correlation (pre-recruit lag-1)	-0.407	-0.364
Prob correlation = 0	0.105	0.181
Recruit to post-recruit		
Correlation	0.613	-0.052
Prob correlation = 0	0.007	0.842
Prob of equal variances	0.876	0.151
Correlation (pre-recruit lag-1)	-0.380	0.626
Prob correlation = 0	0.132	0.009

<sup>a</sup>Zero survey indices were observed in 1988 and 1989 for the Pribilof stock. Zero indices are problematic because of the lognormal error structure. The sum of squares and standard deviation were calculated and the residual analyses performed with the zero survey indices omitted.

Prob = probability.

slightly lower molt probabilities and a less significant fit. These estimates are lower than the value 0.76, obtained by Otto and Cummiskey (1990), by lumping together all crab recaptures after 1 year at liberty.

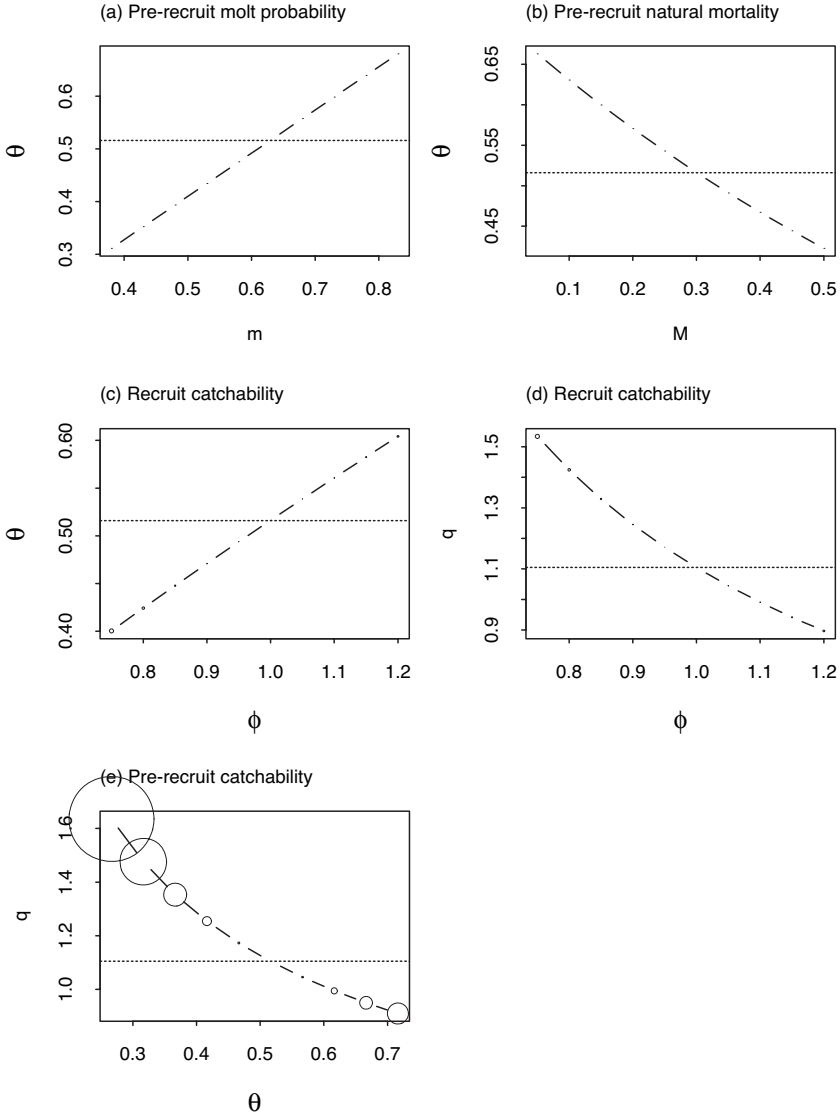
### **Baseline runs**

The distribution of residuals from the baseline runs (Table 2) was examined with histograms and quantile-quantile plots (not shown). Residual distributions for both stocks were skewed but were not significantly different from a normal distribution according to chi-square tests. The variances of the residuals from the different size classes of crabs were tested for homogeneity with an *F*-ratio test. There were no significant differences between size classes for either stock.

The standard deviation of residuals was 0.36 for the St. Matthew stock and 0.89 for the Pribilof stock. The Pribilof stock had four large positive residuals associated with zero abundance indices, in which case the offset added to prevent undefined logarithms substantially increased the residual standard deviation (Table 2). When the three residuals for the zero indices were removed, the standard deviation for the Pribilof stock was reduced to 0.75. The pre-recruit, recruit, and post-recruit residuals each year tended to be positively correlated. This might occur because all three indices come from the same survey and the total survey catch is subject to a common observation error. The correlations were stronger for the St. Matthew Island stock, and weaker for the Pribilof Islands stock, with negative, insignificant pre-recruit/recruit and recruit/post-recruit correlations.

First-order autocorrelation of the residuals was weak. However, one-year lagged correlations were sometimes significant (Table 2). For St. Matthew, the lag-1 correlation between pre-recruits and recruits was negative and significant. A correlation involving the pre-recruit residual lagging the recruit residual by one year would occur when there is a low survey year followed by a high survey year; the pre-recruit residual would be positive and the recruit residual negative to "split the difference." The lag-1 correlation between recruits and post-recruits and between pre-recruits and post-recruits were also negative but not significant. For the Pribilof stock, the lag-1 correlation between pre-recruits and recruits was negative and significant and between recruits and post-recruits was positive and significant.

In summary, the residuals were normally distributed and homoscedastic, as assumed in fitting the model and in creating the bootstrapped standard deviations of the abundance estimates. The main features of the residuals that were not accounted for in fitting the model are positive covariances and negative lagged covariances. This residual structure could be addressed in future extensions of catch-survey analysis.



**Figure 2.** Sensitivity of the estimated post-recruit catchability coefficient ( $q$ ) and pre-recruit catchability ( $\theta$ ) to errors in specifying the input values of pre-recruit molt probability ( $m$ ), natural mortality ( $M$ ), recruit catchability ( $\phi$ ), and  $\theta$ . The reference values were  $q = 1.105$ ,  $\theta = 0.516$ ,  $m = 0.63$ ,  $M = 0.3$ , and  $\phi = 1$ . In each subplot, the value of the abscissa was varied over the range of points shown, and the value of the ordinate re-estimated with three-stage CSA. The areas of the circles are proportional to the increase in the sum of squared residuals.



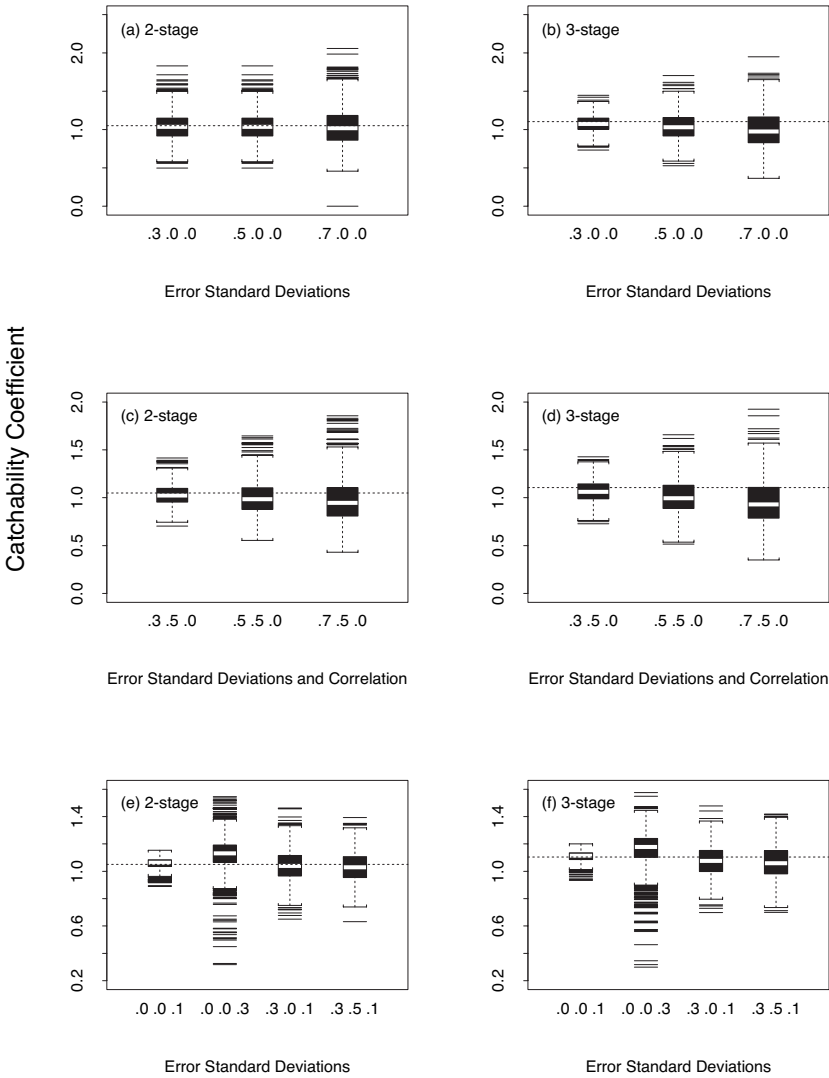
### **Sensitivity tests**

The St. Matthew stock was used as a representative example for sensitivity tests and simulation analyses. We examined the sensitivity of the estimated post-recruit  $q$ , and pre-recruit  $\theta$ , catchability coefficients to errors in specifying the input values of pre-recruit molt probability  $m$ , natural mortality  $M$ , recruit catchability  $\phi$ , and pre-recruit catchability  $\theta$  (Fig. 2). The absence of an increase in the sum of squared residuals indicates that pre-recruit catchability  $\theta$ , pre-recruit molt probability, and natural mortality are structurally confounded (Fig. 2a,b). It is therefore possible to estimate only one of these three parameters in the model fit, in this case pre-recruit catchability. Pre-recruit catchability and natural mortality are negatively correlated. The way the three-stage model is formulated, molt probability of pre-recruits is analogous to natural survival and thus inversely related to  $M$ . Thus, estimates of  $\theta$  should be positively related to molt probability (Fig. 2a). A small fraction of pre-recruits may molt directly into the post-recruit size range without being recruits. However, this is offset by a small fraction of recruits whose molt increment is insufficient to put them in the post-recruit stage.

Errors in molt probability or pre-recruit natural mortality directly influence pre-recruit abundance estimates. For example, if we assume that pre-recruit mortality equals recruit and post-recruit mortality (when in fact it is higher for smaller crabs),  $\theta$  would be overestimated and pre-recruit abundance would be underestimated. Thus, accurate estimates of the absolute abundance of pre-recruits depend on good estimates of molt probability and natural mortality. Estimates of both pre-recruit catchability and recruit catchability are inversely related to estimates of post-recruit catchability (Fig. 2d,e). However, as the three-stage model compensates for errors in recruit catchability by changing pre-recruit catchability (Fig. 2c), errors in pre-recruit catchability increase the sum of squared residuals much more than errors in recruit catchability (Fig. 2e).

### **Monte Carlo simulation analyses**

Using the St. Matthew Island fit to data from 1980-1997 as the "true" population, we performed the same simulations with the two-stage and three-stage models. We performed Monte Carlo simulations assuming correlated and uncorrelated observation errors, process errors, and mixed process and observation errors. Generally, we expect a combination of both measurement and process errors. Table 3 contains details of the errors used in the simulations and a comparison of the bias and standard deviation of  $q$ , the average percent bias of the abundance estimates, average magnitude (absolute value) of the percent bias of the abundance estimates, the average coefficient of variation, of the abundance estimates and the number of converged replicates out of 1,000 for the two-stage and three-stage models. About 98% of the simulations converged in all



**Figure 3.** Distributions of the catchability coefficients ( $q$ ) estimated in the Monte Carlo simulations of the 2- and 3-stage CSA. Each box plot summarizes the distribution of 1,000  $q$  values. The white bar in the interior of the box is the median. The height of the box is equal to the interquartile distance (IQD). The whiskers extend 1.5 times the IQD, or to the extreme range of the data, whichever is less. Individual lines beyond the whiskers are outliers. The three numbers under each box indicate the standard deviation of the observation errors, the correlation of observation errors, and process error standard deviation (from left to right).

Table 3. Results of the Monte Carlo simulations of the 2- and 3-stage CSA.

Standard deviation of measurement error	Correlation of measurement errors	Standard deviation of process errors	Number of replicates that converged		Bias of $q$ estimates		Standard error of the $q$ estimates	
			2-stage	3-stage	2-stage	3-stage	2-stage	3-stage
0.3	0.0	0.0	999	999	-0.01	-0.03	0.11	0.11
0.5	0.0	0.0	998	1,000	-0.01	-0.06	0.18	0.18
0.7	0.0	0.0	995	998	-0.02	-0.10	0.25	0.24
0.9	0.0	0.0	995	983	-0.02	-0.14	0.31	0.30
0.3	0.5	0.0	1,000	1,000	-0.02	-0.04	0.11	0.11
0.5	0.5	0.0	999	1,000	-0.06	-0.09	0.17	0.18
0.7	0.5	0.0	997	995	-0.08	-0.15	0.23	0.23
0.9	0.5	0.0	979	985	-0.10	-0.21	0.29	0.29
0.0	0.0	0.1	988	983	0.01	0.00	0.04	0.04
0.0	0.0	0.3	989	931	0.07	0.05	0.14	0.14
0.3	0.0	0.1	998	993	-0.01	-0.03	0.12	0.12
0.3	0.5	0.1	1,000	995	-0.02	0.00	0.11	0.04

Table 3. (Continued). Results of the Monte Carlo simulations of the 2- and 3-stage CSA.

Standard deviation of measurement error	Correlation of measurement errors	Standard deviation of process errors	Legal abundance estimates					
			Average % bias		Average absolute % bias		Average CV	
			2-stage	3-stage	2-stage	3-stage	2-stage	3-stage
0.3	0.0	0.0	-0.6	0.3	0.7	0.2	15.29	13.22
0.5	0.0	0.0	-2.2	-0.2	2.5	0.8	26.72	24.06
0.7	0.0	0.0	-4.9	-2.5	5.3	2.8	38.36	37.35
0.9	0.0	0.0	-9.0	-6.4	9.7	6.3	51.02	54.94
0.3	0.5	0.0	0.7	1.4	0.6	1.2	14.71	13.42
0.5	0.5	0.0	0.8	2.4	1.4	2.4	25.66	24.66
0.7	0.5	0.0	-0.9	1.1	2.8	4.5	37.96	39.37
0.9	0.5	0.0	-4.3	-2.2	6.1	7.0	50.93	59.55
0.0	0.0	0.1	-0.3	-0.3	0.4	0.4	11.00	12.83
0.0	0.0	0.3	-3.2	-5.2	3.2	5.2	46.69	58.94
0.3	0.0	0.1	-0.1	0.5	0.4	0.6	19.56	19.23
0.3	0.5	0.1	0.6	1.3	0.5	1.3	19.28	19.41

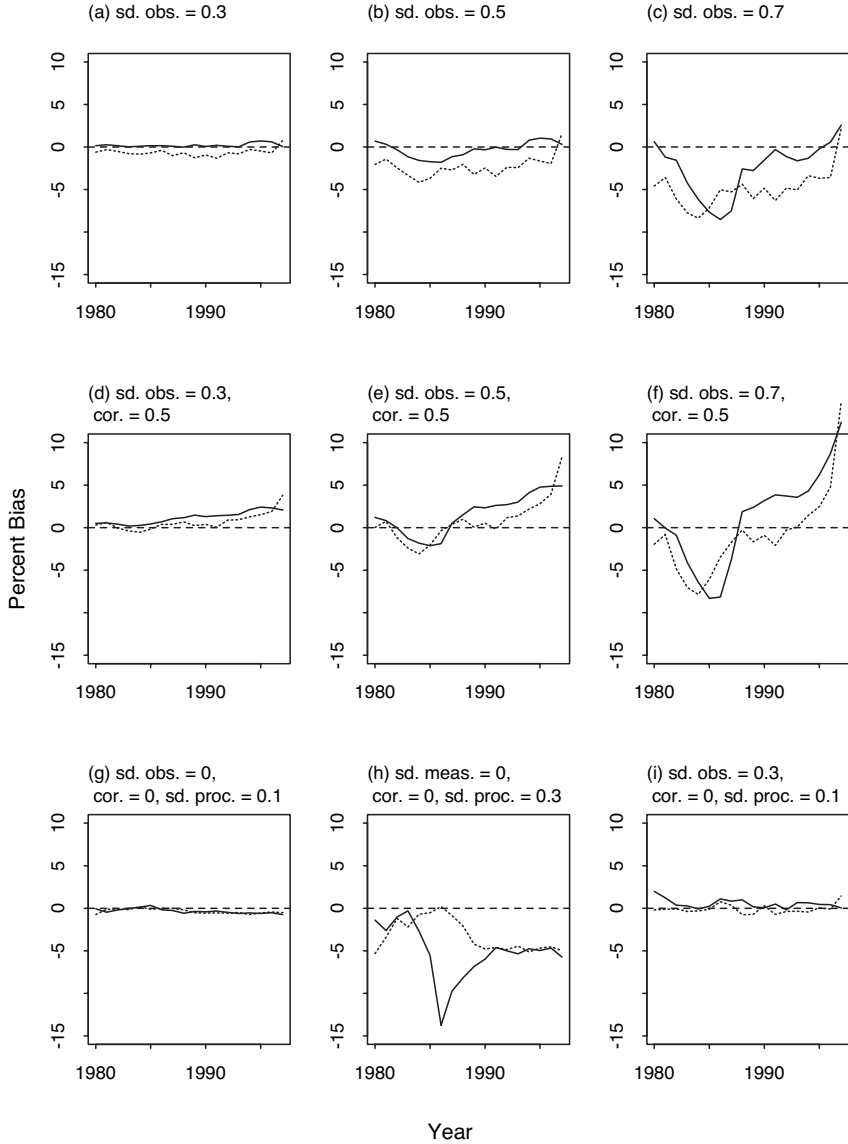
cases except for one; only 93% of the runs using the three-stage model and process errors with standard deviation of 0.3 converged.

The simulation results from both models were qualitatively similar (Fig. 3, Table 3). According to a chi-square goodness of fit test, the distribution of the logarithm of the simulated  $q$  and  $\theta$  values from the three-stage model did not differ from the normal distribution, indicating that these parameters are lognormally distributed. In all simulations, the bias of  $q$  tended to be smaller with the two-stage than with the three-stage model, but the standard errors of  $q$  were similar between the two models. The standard error of  $q$  increased as the standard deviation of the observation errors was increased. Adding correlated observation errors did not affect the standard error of  $q$ , but caused an increase in the bias of  $q$  in both models. In the simulations with process errors only, the bias of  $q$  was large and positive with both models. However, with a combination of correlated and uncorrelated measurement and small process errors, the bias of  $q$  became very small in both models (Fig. 3).

The main interest for stock assessment is the ability of the models to estimate population abundance accurately. Perhaps the best comparison of the two models is the average percent bias of legal abundance, average magnitude of the bias of legal abundance, and the CV of legal abundance from the two models (Table 3). In the runs with uncorrelated observation errors, the output from the models indicates that the average bias, average magnitude of the bias, and the CV of the abundance estimates are smaller with the three-stage model than with the two-stage model. With uncorrelated observation errors, the three-stage formulation appeared better able to smooth this variability than the two-stage model.

The average bias of the legal abundance estimate in each year from the two-stage model increased and became more negative as the standard deviation of observation errors increased (Fig. 4a,b,c). The largest bias occurred during 1982-1985 when the population abundance was decreasing to very low levels. The average yearly bias of the three-stage model also increased and became more negative as the standard deviation of the observation errors increased. However, in comparison with the two-stage model, the magnitude of the bias with the three-stage model was greater during the period 1982-1985 and less during the remainder of the time series.

In general, one may expect a slightly positive bias in the abundance estimates because the simulated lognormal process errors cause some large positive deviations from the fitted input data. However, we simulated the lognormal observation errors to have mean 1.0, which causes the median of the observation errors to be less than 1.0. As the standard deviation of the observation errors increases, the median shifts to a lower and lower value (0.94 with a standard deviation of 0.3, and 0.81 with a standard deviation of 0.7). Furthermore, with a standard deviation of 0.3 and 0.5, there was a significant correlation ( $P < 0.07$ ) that the correlation



**Figure 4. Percent bias by year of legal abundance from Monte Carlo simulations with the 3-stage (solid line) and 2-stage (dotted line) models with different combinations of observation and process errors. The abbreviations indicate the standard deviation of observation errors (sd. obs.), the correlation of observation errors (cor.), and the standard deviation of process errors (sd. proc.).**

is zero) between the median of the observation error for one run and the average percent bias of the population estimates of the run. As a result, as the standard deviation increased, the bias became more negative.

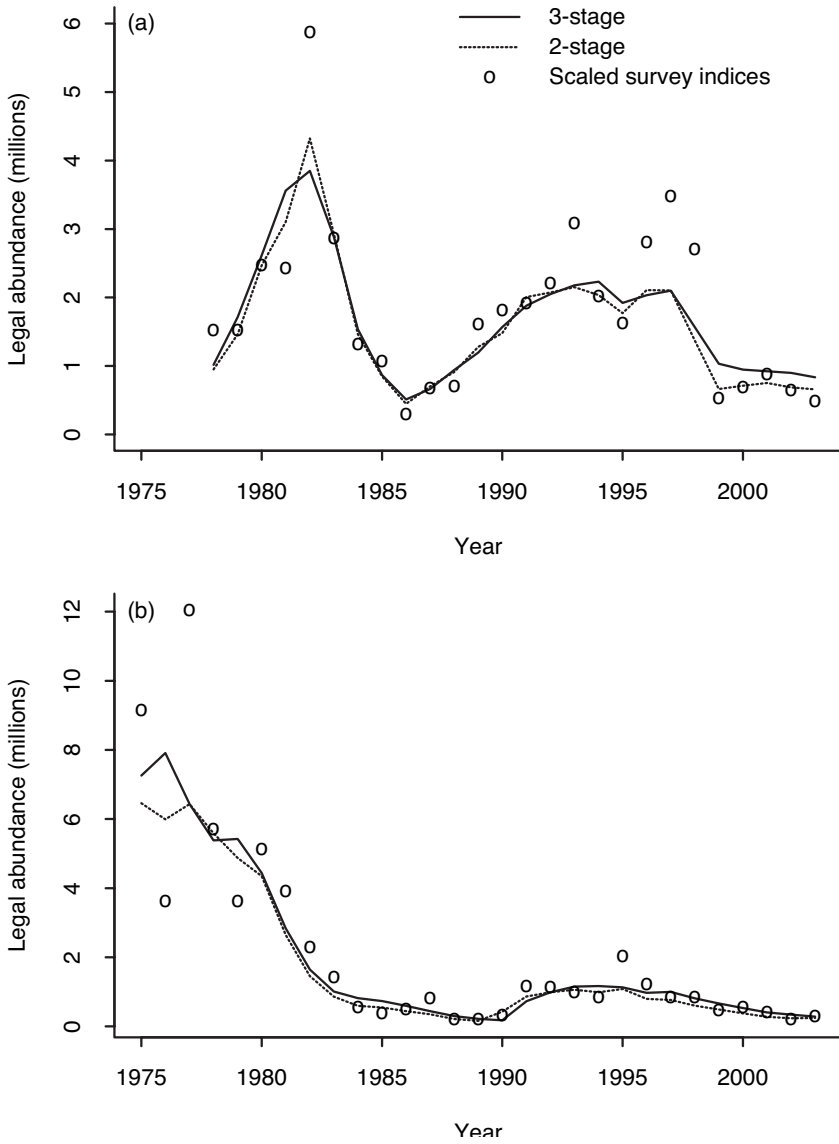
The percent bias of legal abundance was quite large during the period 1982-1985 (Fig. 4). During the intermediate years (1982-1987), the legal abundance decreased and became quite low in comparison to the earlier and later years. The percent bias was most likely amplified during the years of low abundance because of the logarithmic transform in the objective function. The simulated relative abundance estimates could become very small for these intermediate years, especially as the standard deviation of the observation error increased. The logarithm in the objective function places greater emphasis on the residuals during the years of low abundance and the model tries hard to fit these small values. This bias may be more extreme with the three-stage model than with the two-stage model because there are three residuals associated with the small cohorts rather than two.

In the simulations with correlated observation errors, the two-stage model appeared to out-perform the three-stage model. It also appeared to do better than the two-stage model with uncorrelated observation errors (Table 3). This counter-intuitive result is opposite from the findings of Collie and Kruse (1998) in their simulations with the two-stage model applied to the Bristol Bay red king crabs and from our simulations of correlated and uncorrelated errors with the three-stage model. However, over the range of expected observation errors (standard deviation of 0.3-0.5), the average magnitude of the percent bias of legal abundance was 1.4% and 2.4% from the two- and three-stage models, respectively, a particularly small bias. The most extreme biases occurred during the period 1982-1985 (Fig. 4d,e,f) as with the uncorrelated observation errors, and for the same reasons.

With the process-error assumption, the three-stage output had slightly larger bias in the legal abundance estimates than the output from the two-stage model (Table 3). One possible reason for this is that the process error is compounded over two stages in the three-stage model, recruit and post-recruit. The difference was most apparent in the run with just process error with a standard deviation of 0.3 (Fig. 4h). Again, the three-stage model had the largest bias during the intermediate years when the population abundance was very low. The three-stage model also had a notably larger CV of legal abundance in this particular run.

### ***St. Matthew Island blue king crab***

The final runs were made with data from the period 1978-2003, with bootstrapped confidence intervals. Estimates of legal crab abundance from the three-stage model followed those from the two-stage CSA closely, although the three-stage model predicted slightly higher legal abundances for all but 5 years (Fig. 5a). As we set the instantaneous natural mortality



**Figure 5. Comparison of the 2- and 3-stage catch survey analyses for (a) St. Matthew and (b) Pribilof Islands blue king crab stocks.**



**Table 4. Comparison of the two- and three-stage CSA output for the final runs with all years.**

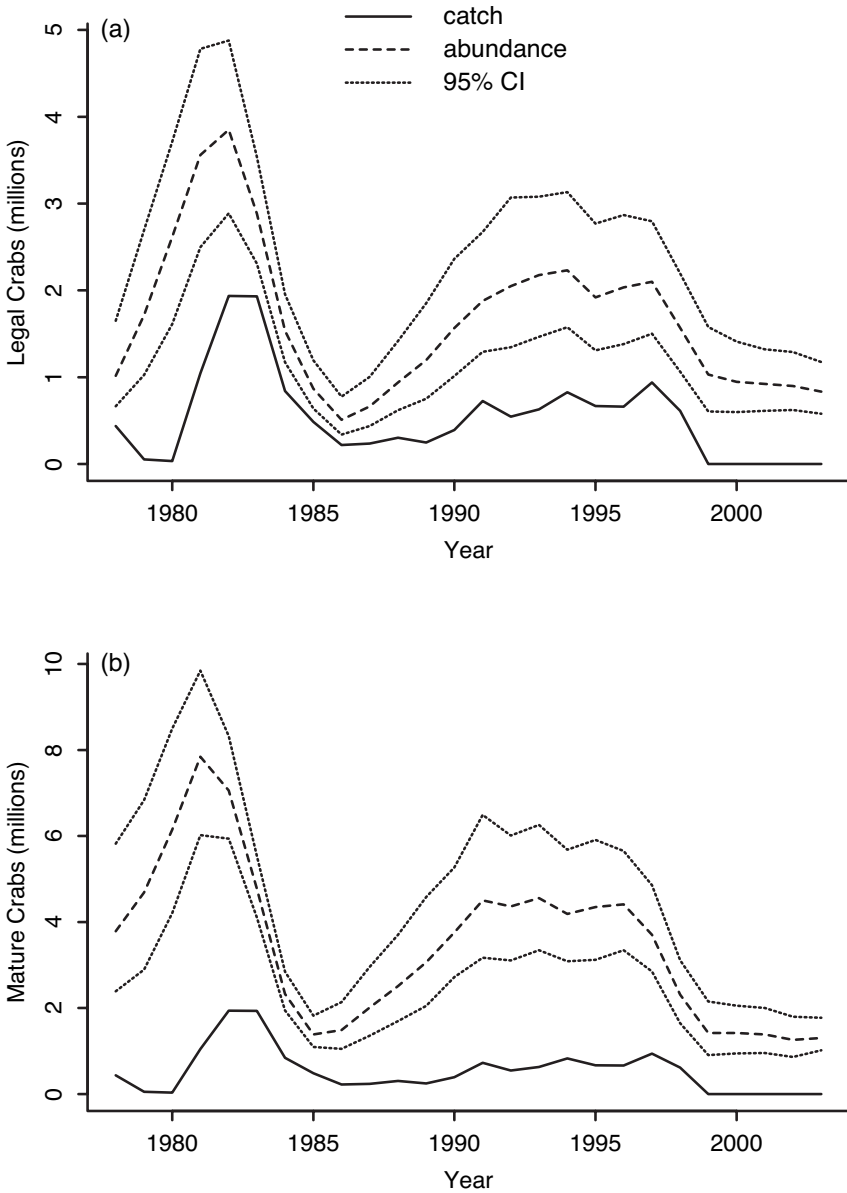
	Pribilof Islands		St. Matthew Island	
	Two-stage	Three-stage	Two-stage	Three-stage
Catchability coefficient	1.130	0.984	1.280	1.170
Standard error of $q$	0.488	0.968	0.436	0.428
Degrees of freedom	27	54	24	48
Sum of squared residuals	5.445	46.919	4.604	10.883
Mean square error	0.202	0.869	0.192	0.227

rate to 0.3 for all years, these results are consistent with the results of Zheng et al.'s (1997) two-stage CSA. In comparison with the two-stage model, the three-stage model smoothed and moderated abundance levels through time with the peak in 1982 reduced from about 4.5 to 3.9 million crabs. The estimated catchability of pre-recruits relative to post-recruits was estimated to be 0.60. The catchability coefficients and mean square errors from the two- and three-stage CSAs were comparable (Table 4).

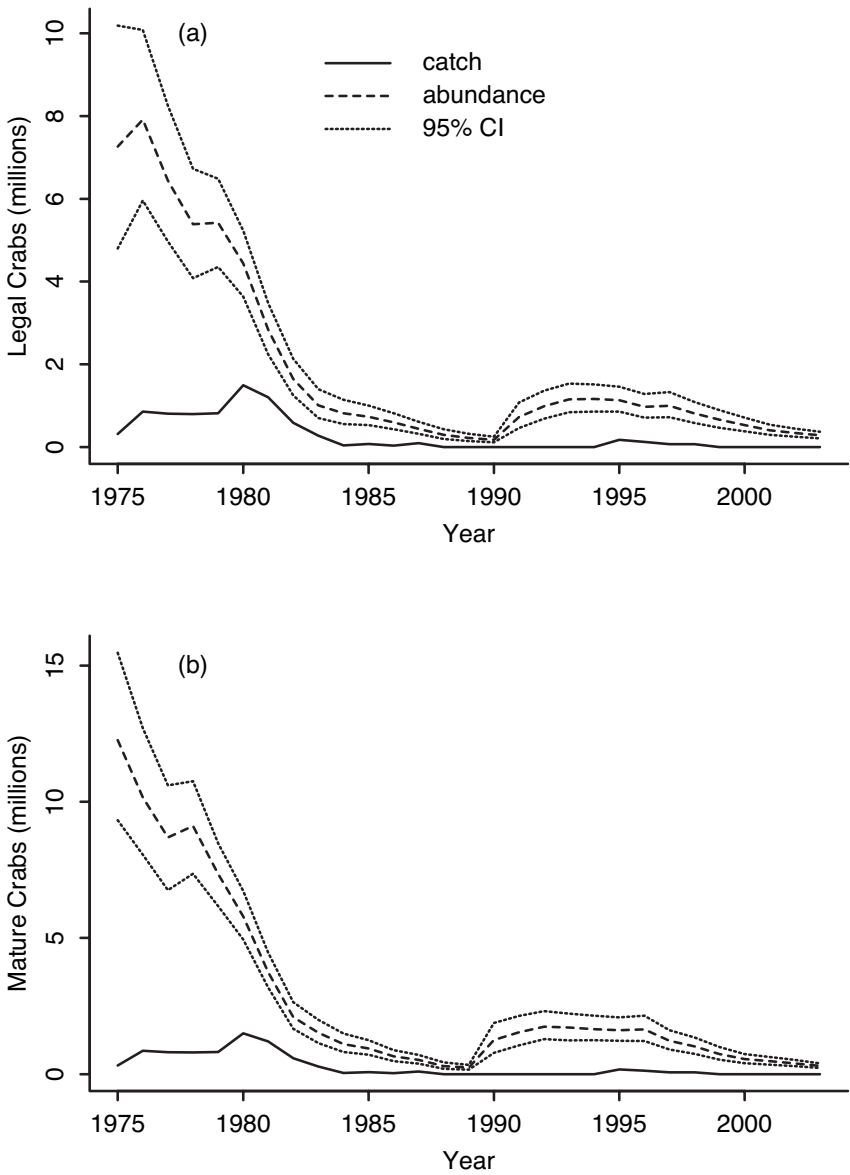
The bootstrap runs resulted in a very small (−0.073) bias in the catchability coefficient and a considerably smaller (0.004) bias in the mean estimate of the catchability of pre-recruits relative to post-recruits ( $\theta$ ). The bootstrapped average percent bias of mean legal abundance was 0.29 and the average coefficient of variation of legal abundance was 19.5. The bootstrapped abundance means coincided almost exactly with the model estimates of legal and mature crab abundance, indicating low bias in the estimation. The 95% confidence intervals are wider at high abundance because of the lognormal error structure and they are narrower during periods of high exploitation because the model more heavily constrains the abundance estimates (Fig. 6). After increasing in the 1990s, the St. Matthew stock decreased abruptly in 1999 and the fishery has been closed since then. Zheng and Kruse (2000) attributed this abrupt decrease to high natural mortality in 1999.

### ***Pribilof Islands blue king crab***

Our updated analysis used data for years 1975–2003 (Table 1b). The residual standard deviation from this fit was 0.93, which is quite high in comparison with the other stock. For the bootstrap simulation we used an input standard deviation of 0.36, which was calculated by removing the four largest residuals associated with the zero survey abundance estimates. The catchability coefficient was slightly negatively biased (−0.026) and the catchability of pre-recruits relative to post-recruits was slightly positively biased (0.010) in the bootstrap fits. The bootstrap average percent bias of mean legal abundance was −0.87 and the average coeffi-



**Figure 6. Legal and mature crab abundances and bootstrapped 95% confidence intervals from the 3-stage model for the St. Matthew Island blue king crab stock.**



**Figure 7. Legal and mature crab abundances and bootstrapped 95% confidence intervals from the 3-stage model for the Pribilof Islands blue king crab stock.**

cient of variation of legal abundance was 15.71. The model and bootstrap estimates of legal and mature crab abundance and the corresponding bootstrap 95% confidence intervals are plotted in Fig. 7a,b. After a modest recovery in the early 1990s the abundance of Pribilof Islands blue king crabs declined and the fishery has been closed since 1999.

There were a few large residuals from the three-stage model fit for the Pribilof Islands stock. These large residuals are associated with the cohorts with very low abundance as described in the residual analysis section of this paper. We found that the standard error of the residuals is sensitive to the offset value  $\delta$ , used in the objective function; offset values of 0.1, 0.01, 0.001, and 0.0001 yielded residual standard deviations of 0.45, 0.69, 0.93, and 1.29, respectively. Woodby (1994), in his estimation of the Southeast Alaska red king crab population size, used the smallest offset value that allowed the minimization subroutine to converge. Comparing the effect of the offset value on the estimation of population abundance, we found that for the wide range of offset values listed above there was a relatively small difference in population estimates.

For the Pribilof Islands stock, the three-stage CSA results were consistent with the results from the two-stage CSA except in 1975 and 1976 (Fig. 5b). The three-stage model predicted higher legal abundance because of the high pre-recruit index in 1975. During the periods from 1980 to 1989 and 1993 to 2003, the legal abundance estimates from the three-stage model were slightly higher than those from the two-stage model. The two-stage model predicted slightly higher abundances in 1990, 1991, and 1992. The catchability of pre-recruits relative to post-recruits at 0.81 was larger than the value estimated for St. Matthew, perhaps due to the larger size of this class. The catchability coefficient was estimated to be 0.98 in the three-stage model and was slightly lower than the  $q$  from the two-stage model (Table 4). The sum of squared residuals at 46.9 was quite high in the three-stage model, with the 1989 pre-recruit cohort contributing 36% of the sum of squared residuals and the 2000 pre-recruit cohort contributing 26%. In years with zero pre-recruit or recruit abundance indices, the model predicts a small, positive value to be consistent with non-zero indices in the previous and following year. The residuals were accentuated for these low indices due to the logarithm in the objective function.

## Discussion

The three-stage catch survey analysis appears to have several benefits over the two-stage CSA as it applies to king crabs. First, a smoothed value of legal abundance can be estimated for the most recent year with the three-stage model. The two-stage CSA is unable to estimate recruit abundance in this last year; thus the unsmoothed survey index must be used. The three-stage model can also be used to estimate the abundance

of mature male king crabs, because the size of pre-recruits corresponds approximately to the size at maturity of male king crabs (Somerton and MacIntosh 1983). This is important for fishery managers because the commercial quota is based upon estimates of both legal and mature crab abundance. However, the three-stage model cannot estimate pre-recruit abundance in the terminal year, so model estimates of legal crabs must be combined with the survey index of pre-recruits to estimate mature abundance in the last year.

The three-stage model uses more data than the two-stage model and therefore has a higher ratio of observations to parameters. For the three-stage model only two more parameters need to be estimated, one relative abundance estimate and  $\theta$ . The bootstrap simulations indicate that  $\theta$  is well defined in this model. The coefficient of variation of  $\theta$  from the simulations was 13.3% for St. Matthew Island and 9.5% for the Pribilof Islands, which resulted in quite narrow confidence intervals for  $\theta$ . For the St. Matthew Island stock, the mean square error of the three-stage model was comparable to the two-stage model (Table 4). In the Pribilof Islands case, the high mean square error resulted from the zero abundance indices.

The sensitivity analyses revealed that the three-stage model is able to smooth uncorrelated observation errors better than the two-stage formulation. As a result, the bootstrapped confidence intervals about the legal abundance estimates should be narrower with the three-stage than with the two-stage model. This was the case for Bristol Bay red king crab legal abundance confidence intervals when we compared the three-stage model (not shown) to the two-stage model of Collie and Kruse (1998). More detailed analysis will be necessary to diagnose the reason why the two-stage model with correlated observation errors performed better than the three-stage model with corresponding error structure. Also, the two-stage formulation is slightly more robust than the three-stage model with regard to process errors, which makes sense because the process errors are compounded over two time steps with the three-stage model as opposed to one time step with the two-stage model. Sensitivity analysis showed that estimates of pre-recruit molt probability and natural mortality are critical to the estimates of pre-recruit absolute abundance and, consequently, the mature crab abundance estimates.

The main purpose of this study was to develop and test a three-stage catch-survey analysis. The blue king crab examples were meant more to illustrate the method than as definitive assessments of these two stocks. However, CSA is used to set harvest guidelines for some Alaska king crab stocks. The harvest rate of blue king crabs is set on a sliding scale from 10% to 20% depending on mature crab abundance (Zheng and Kruse 2000). Below a minimum abundance threshold the fishery is closed, as has been the case for both stocks since 1999. This harvest strategy underscores the importance of accurate abundance estimates and of filtering out observation errors in the most recent year to the extent possible.

Additional extensions to three-stage catch-survey analysis are possible. Jie Zheng (ADFG, pers. comm.) pointed out that pre-recruits that don't molt can survive as pre-recruits the following year. This information could add additional structure to the three-stage model, especially if the shell age were known for the pre-recruit size class. Even without observations of shell age, the predicted number of old-shell pre-recruits (surviving pre-recruits that didn't molt the previous year) should be less than or equal to the predicted number of total pre-recruits for each year.

We examined the output from the baseline runs of the three-stage model to see if this constraint was met. In every case but one the total pre-recruit estimate exceeded the prediction of old-shell pre-recruits. The exception occurred in 1988 for the Pribilof Islands stock, for which the pre-recruit index was 0.0 and the predicted old-shell pre-recruits was 0.0075, a very small difference. For these king crab stocks, the proportion of pre-recruits not molting and surviving is small, around 0.2. Thus the extra constraint is unlikely to be binding. However, partitioning the pre-recruit size class into new and old shell categories could provide additional model structure to help estimate confounded parameters.

In some applications it is necessary to convert the harvest rates  $h$  to the corresponding fishing mortality rates  $F$ , for comparison with biological reference points. Given the assumption of a discrete harvest, the most consistent approximation is  $F \approx -\ln(1 - h)$ . However, the reference point should be calculated in the same way, and for stocks that are assessed with CSA, it may be preferable to express the reference points as harvest rates.

In future examinations of the three-stage model, one might investigate the effect of autocorrelated process errors, as might be expected because of environmental shifts. Further, it would seem appropriate to explore lognormal observation errors for the aggregate survey index and multinomial classification of errors within each year. In their original formulation of CSA, Collie and Sissenwine (1983) estimated process and observation errors. Our Monte Carlo simulations with two- and three-stage models confirm that the all observation error version is robust to realistic levels of measurement and process error. Cadrin (2000) used Monte Carlo simulations to evaluate the performance of the mixed-error two-stage CSA; similar studies are needed for the three-stage CSA. More rigorous and realistic assumptions about the error distributions could be incorporated in a full likelihood model (Punt 2003).

Recently, Zheng and Kruse (2000) applied a four-stage CSA to the estimation of blue king crabs. The size of the fourth stage is one molt below the pre-recruit stage. By incorporating additional data into the estimation it is possible to obtain model estimates of mature abundance in the most recent year. Adding more stages to the model is a logical progression toward a full length-based model with stochastic growth (Zheng et al. 1995). Yet a full length-based model is not possible for these

blue king crab stocks owing to limited size-structured data stemming from difficulties conducting trawl surveys in these rugged areas. Optimal model complexity depends, in a large part, on how the results are used. For setting harvest guidelines, a simpler model can often perform as well as a more complicated model. In the case of king crabs, incorporating additional stages is clearly important for setting harvest guidelines. However, this comes at the cost of adding potentially noisy data that can decrease model precision. Therefore model performance needs to be evaluated with simulation studies, such as we have performed here, on a case-by-case basis.

## Acknowledgments

Trawl survey data for this study were provided by Bob Otto and Lou Rugolo (NMFS), and compiled by Jie Zheng (ADFG). We thank Bob Otto for providing unpublished data from tagging experiments and the two reviewers for their constructive comments. This study was partially funded through a cooperative agreement between the University of Rhode Island and the Alaska Department of Fish and Game (COOP-98-005).

## References

- Armstrong, D.A., J.L. Armstrong, R. Palacios, G. Williams, G.C. Jensen, and W. Pearson. 1985. Early life history of blue king crab, *Paralithodes platypus*, around the Pribilof Islands. In: Proceedings of the International King Crab Symposium. Alaska Sea Grant College Program, University of Alaska Fairbanks, pp. 211-229.
- Blau, S.F. 1996. The 1995 St. Matthew Island blue king crab survey. Alaska Department of Fish and Game, Commercial Fisheries Management and Development Division, Regional Information Report 4K96-27, Kodiak. 36 pp.
- Cadrin, S.X. 2000. Evaluating two assessment methods for Gulf of Maine northern shrimp based on simulations. J. Northw. Atl. Fish. Sci. 27:119-132.
- Cadrin, S.X., S.H. Clark, D.F. Schick, M.P. Armstrong, D. McCarron, and B. Smith. 1999. Application of catch-survey models to the northern shrimp fishery in the Gulf of Maine. N. Am. J. Fish. Manag. 19:551-568.
- Collie, J.S., and G.H. Kruse. 1998. Estimating king crab (*Paralithodes camtschaticus*) abundance from commercial catch and research survey data. In: G.S. Jamieson and A. Campbell (eds.), Proceedings of the North Pacific Symposium on Invertebrate Stock Assessment and Management. Can. Spec. Publ. Fish. Aquat. Sci. 125:73-83.
- Collie, J.S., and M.P. Sissenwine. 1983. Estimating population size from relative abundance data measured with error. Can. J. Fish. Aquat. Sci. 40:1871-1879.
- Conser, R.J. 1991. A DeLury model for scallops incorporating length-based selectivity of the recruiting year-class to the survey gear and partial recruitment to the commercial fishery. Northeast Regional Stock Assessment Workshop Report, Res. Doc. SAW 12/2, Woods Hole, Massachusetts. 18 pp.

- Conser, R.J., and J. Idoine. 1992. A modified DeLury model for estimating mortality rates and stock sizes of American lobster populations. Northeast Regional Stock Assessment Workshop Report, Res. Doc. SAW 14/7, Woods Hole, Massachusetts. 28 pp.
- Kennedy, W.J., and J.E. Gentle. 1980. Statistical computing. Statistics: Textbooks and monographs, Vol. 33. Marcel Dekker, Inc., New York.
- Mesnil, B. 2003. The catch-survey analysis (CSA) method of fish stock assessment: An evaluation using simulated data. *Fish. Res.* 63:193-212.
- Neter, J., M.H. Kutner, C.J. Nachtsheim, and W. Wasserman. 1996. Applied linear statistical models. 4th edn. Times Mirror Higher Education Group, Inc.
- Otto, R.S., and P.A. Cummiskey. 1990. Growth of adult male blue king crab (*Paralithodes platypus*). In: Proceedings of the International Symposium on King and Tanner Crabs. Alaska Sea Grant College Program, University of Alaska Fairbanks, pp. 245-258.
- Punt, A.E. 2003. Extending production models to include process error in the population dynamics. *Can. J. Fish. Aquat. Sci.* 60:1217-1228.
- Somerton, D.A., and R.A. MacIntosh. 1983. The size at sexual maturity of blue king crab, *Paralithodes platypus*, in Alaska. *Fish. Bull. U.S.* 81:621-628.
- Stevens, B.G., R.A. MacIntosh, J.A. Haaga, C.E. Armistead, and R.S. Otto. 2002. Report to industry on the 2002 eastern Bering Sea crab survey. National Marine Fisheries Service, Alaska Fisheries Science Center Processed Report 2002-05, Kodiak. 59 pp.
- Woodby, D. 1994. Population sizes and harvest rates of red king crabs in southeast Alaska, 1993. Alaska Department of Fish and Game, Commercial Fisheries Management and Development Division, Regional Information Report 1J94-20, Juneau. 10 pp.
- Zheng, J., and G.H. Kruse. 2000. Overview of stock assessment and recommended harvest strategy of St. Matthew Island blue king crabs. Alaska Department of Fish and Game, Division of Commercial Fisheries, Regional Information Report 5J00-06, Juneau. 19 pp.
- Zheng, J., G.H. Kruse, and M.C. Murphy. 1996. Comparisons of abundance estimation methods for red king crab in Bristol Bay and Kodiak. In: High latitude crabs: Biology, management, and economics. Alaska Sea Grant College Program, University of Alaska Fairbanks, pp. 284-293.
- Zheng, J., M.C. Murphy, and G.H. Kruse. 1995. A length-based population model and stock-recruitment relationships for red king crab, *Paralithodes camtschaticus*, in Bristol Bay, Alaska. *Can. J. Fish. Aquat. Sci.* 52:1229-1246.
- Zheng, J., M.C. Murphy, and G.H. Kruse. 1997. Application of a catch-survey analysis to blue king crab stocks near Pribilof and St. Matthew islands. *Alaska Fish. Res. Bull.* 4:62-74.
- Zheng, J., M.C. Murphy, and G.H. Kruse. 1998. Abundance estimation of St. Matthew Island blue king crabs using survey and commercial catch and effort data. In: Fishery stock assessment models. Alaska Sea Grant College Program, University of Alaska Fairbanks, pp. 575-588.



# Bias of Equilibrium-Based Estimators under Biological and Fishery Disequilibria

**Billy Ernst and Juan L. Valero**

*University of Washington, School of Aquatic and Fishery Sciences, Seattle, Washington*

## Abstract

Steady state or equilibrium assumptions have been recurrent temptations in fishery science to circumvent the problem of limited data and to provide managers with biological parameters and population status estimates in data-poor situations. Historically, fishery scientists and ecologists have used equilibrium approaches to produce estimates of key population parameters (e.g., mortality rates and abundance) with applications ranging from small-scale artisanal to large-scale industrial fisheries. Unfortunately, these methods rely on many restrictive hard-to-test assumptions, especially for those related to the lack of representation of the underlying dynamics. Size (mostly length) frequency data are among the basic statistics collected in most fisheries. Length-based cohort analysis and length-converted catch curve analysis (LCCCA) are part of the methodological toolkit available to modelers to obtain abundance estimates and mortality rates. Ingenious variations and extensions of these methods have been proposed but the uncertainty in model structure has virtually not been addressed. We evaluate the performance of equilibrium-derived estimators using two different methods. We implemented a dynamic operating model to show the lack of robustness of Jones' length-based cohort analysis to trends in fishing mortality rates. We also evaluated the performance of LCCCA in estimating natural mortality rates of the Patagonian scallop (*Zygochlamys patagonica*) by a cross comparison of these results to the estimates from a fully integrated statistical dynamic model. Substantial biases in parameter estimates suggest the importance of formally incorporating dynamics into the model structure. We recom-

mend that mortality rates and abundance estimates derived from these methods should be cautiously accepted or not used at all. Indeed, scientists and managers should take more proactive measures to collect new data and develop a better understanding of the status and productivity of the stocks.

## **Introduction**

Developing and small-scale artisanal fisheries often lack appropriate data time series to implement formal stock assessment analyses. These data constraints have historically tempted scientists to use simple alternative modeling approaches in an effort to provide fishery managers with reference points and/or stock status information (Sparre and Venema 1998). Many of these methods are based on restrictive assumptions that cannot be validated with the available data. Included among these methods are equilibrium-based techniques, which have a long tradition in ecological data analysis and fisheries stock assessment (Table 1).

Terms like “equilibrium,” “steady-state,” or “pseudo-equilibrium conditions” have been widely used in population biology to describe different concepts. Caddy (1996) points to alternative ways in which equilibrium assumptions can be introduced into stock assessment, for example by using stationary model parameters (e.g., natural mortality), implementing fitting procedures (e.g., Gulland-Fox equilibrium approximation methods) and using historical trend information to generate model predictions (most models). In a strict sense, many dynamic estimation procedures, which we commonly use from simple aggregated models to more complex age/length/stage structured models, can harbor equilibrium assumptions, i.e., in the initial conditions (Punt 1990, Hilborn and Walters 1992). This situation differs greatly from rigid equilibrium assumptions used to generate length-based equilibrium estimators, which are popular in tropical and artisanal fish stock assessments (Pauly and Morgan 1987).

Sparre and Venema justify the use of simple equilibrium methods through the following rationale: “We are often in a situation which forces us to make assumptions, which are known to be crude approximations to reality. It often happens that only by making such assumptions we are able to carry out an analysis of available data, and it is better to do a crude analysis than none at all.” (Sparre and Venema 1998, p. 133). This line of thinking reflects a proactive attitude toward resource management in data-poor situations, but can often create several additional problems: (a) produces very biased results and therefore bad management advice; (b) induces managers to develop an unrealistic confidence in those data and results (given the quantitative nature of these estimates); (c) leads managers to inappropriately address the problems of lack of data and research funding in one particular fishery, and therefore discourage the

**Table 1. Some equilibrium methods/assumptions traditionally used in ecological and fisheries data analysis.**

Type of analysis	Type of data	Estimation
Age-based catch curve analysis	<i>AFD</i>	Abundance-at-age/ <i>Z</i>
Size converted catch curve analysis	<i>SFD/GP</i>	Abundance-at-size/ <i>Z</i>
Van Sickle's method	<i>SFD/GP</i>	Total mortality
Jones' length cohort analysis	<i>SFD/GP/M</i>	Abundance-at-size/ <i>Z</i> - at-size
Yield-per-recruit	<i>GP/M</i>	Age of entry/ <i>F<sub>max</sub></i>
Dynamics age/length structured model	<i>AFD/SFD/GP/M/Idx</i>	Initial age and length structure
Fitting a biomass dynamic model	<i>C/f</i>	Stock productivity

*LFD* = length frequency data, *AFD* = age frequency data, *GP* = growth parameters, *Idx* = index of abundance, *M* = natural mortality, *Z* = total mortality, *C* = catch, *f* = effort.

implementation of more costly approaches (e.g., tagging); and (d) deters the development of alternative simpler robust management methods.

Analytical and computational problems of dealing with highly nonlinear models also historically promoted this venue of scientific inference (Table 1). Development of stock assessment techniques, aided by a large increase in computational power and availability of longer time series, has determined an increasing trend in the use of dynamic statistical nonlinear estimation techniques in many major fisheries (Megrey 1989, Hilborn and Walters 1992, Methot 2000, Ernst 2002). Nevertheless, stock assessments in small-scale artisanal fisheries often rely on equilibrium derived estimators, partially due to data limitations as well as the technicalities involved in the dynamic statistical model fitting approach.

There are several equilibrium methods readily available to estimate mortality rates and/or assessing stock status using size-frequency data and additional input parameters (Sparre and Venema 1998). Two of these approaches are Jones' length cohort analysis (JLCA) and length-converted catch curve analysis (LCCCA). Both methods assume that length-frequency data accurately describe a harvested population at steady state. The use of both techniques is still popular worldwide and implementations of these methods often appear in peer-reviewed and nontraditional journals (Table 2). Despite the generalized use of these techniques, especially in small-scale fisheries, few analyses have been aimed at evaluating the robustness of output statistics derived from these models to departure from equilibrium assumptions (Ernst 2002). Jones (1987) tested the performance of his estimator on cohort growth variability, but did not test the equilibrium assumption. Lai and Gallucci (1988) used analytical tech-

**Table 2. Applications of Jones' "length cohort analysis" or "length converted catch curve analysis" on finfish, crustacean, and mollusk fisheries worldwide.**

Species	Location	Source
<i>Nephrops norvegicus</i>	Mediterranean Sea	Sarda et al. 1998
<i>Paneus stylirostris</i>	Costa Rica	Gonzalez-Pajuelo et al. 1996
<i>Sardinella aurita</i>	Fujian Coast	Xu et al. 1995
<i>Dosichus gigas</i>	Gulf of California	Pierce and Guerra 1994
<i>Sprattus sprattus</i>	Bulgaria, Black Sea Coast	Daskalov and Prodanov 1994
<i>Aristeus antennatus</i>	Northwestern Mediterranean Sea	Demestre et al. 1993
<i>Panulirus argus</i>	Jamaica	Haughton and King 1992
<i>Gadus morhua</i>	Baltic Sea	Sullivan 1991
<i>Scomber japonicus</i>	Tsushima Current	Limbong et al. 1988
<i>Paneus indicus</i>	Tanzania	Wakwabi 1990
<i>Theragra chalcogramma</i>	Bering Sea	Honkalehto 1989
<i>Protothaca staminea</i>	Garrison Bay, Washington	Lai and Gallucci 1988
<i>Strongylocentrotus pallidus</i>	Northern Barents Sea	Bluhm et al. 1998
<i>Schilbe intermedius</i>	Cross River, Nigeria	Etim et al. 1999
<i>Periophthalmus barbarus</i>	Imo River estuary, Nigeria	Etim et al. 2002
<i>Crassostrea rhizophorae</i>	Santa Marta, Colombia	Mancera and Mendo 1996
<i>Pristipomoides filamentosus</i>	Mahe Plateau, Seychelles	Mees 1993
<i>Zygochlamys patagonica</i>	Argentinean Shelf	Lasta et al. 2001

niques to evaluate Jones' method for its sensitivity to input parameters, and indicated that abundance estimates were very sensitive to natural mortality estimates. Despite the relevance of local sensitivity analysis, it does not capture the effect of key model assumptions on potential biases of model outputs. Tests on other equilibrium length-based models have also been conducted. Somerton and Kobayashi (1991) implemented numerical experiments to test the Wetherall length-based regression estimator to departure from population equilibrium, and whose results indicated a serious lack of robustness. Ehrhardt and Ault (1992) presented a sensitivity investigation on Beverton and Holt's length-based total mortality model (Beverton and Holt 1956), which demonstrated a lack of robustness under biased input estimates of exploitable life span of the species considered in their analysis.

We think that the current published literature does not sufficiently warn fishery analysts about the potential risks associated with length-

based equilibrium models used for yield management parameters and/or stock status information. The objective of the present work is to evaluate the impact of departure from equilibrium assumptions on parameter estimates produced by (a) Jones' length cohort analysis (Jones 1984) and (b) length converted catch curve analysis (Pauly 1984, Sparre and Venema 1998) using two different approaches. First, Monte Carlo simulation methods were applied to assess the effect of trends in fishing mortality on abundance estimates from JLCA, and second, the bias in LCCCA's natural mortality estimates for the Patagonian scallop was compared with results of a statistical integrated dynamic age-structured model. These examples aim at illustrating problems associated with these equilibrium estimators under common non-equilibrium conditions.

## Material and methods

### ***Jones' length cohort analysis***

#### *Estimation model*

Jones' length cohort analysis (JLCA) was developed to produce stock size and total mortality estimates using very basic population data (Jones 1984). The method follows a formulation similar to standard age cohort analysis (Megrey 1989), but with the added complication of dealing with length as opposed to age. The main model assumptions can be summarized as follows: (a) an exponential decay function determines the demographic changes of a cohort, (b) continuous recruitment, (c) constant fishing mortality rate operating on the stock, (d) individual growth given by a deterministic von Bertalanffy growth model, and (e) a stationary set of fishing and biological parameters ( $F_{term}$ ,  $M$ ,  $L_{\infty}$ ,  $k$ ). A set of equations allows the computation of numbers-at-length based on fisheries' length frequency data (LFD) and four fishing and biological parameters (Jones 1984). The abundance in the last length interval is computed by a modification of the catch equation and estimation proceeds backward using a modified version of Pope's approximation (Jones 1984, 1987). Corrections are necessary to compute the average number of individuals in the population at each size interval. This procedure avoids biases based on the dependence of abundance estimates to the chosen length interval width of the LFD.

#### *Operating model and data*

The operating model was based on a dynamic age-structured model (see Appendix I for details of model equations and Appendix II for parameter values). The abundance at age was propagated using standard demographic equations and the simulation horizon was 20 years. This provided a reasonable temporal framework to test different exploitation trends. Fishing mortality was divided into temporal and age components

(Megrey 1989). Logistic models were used to describe the selectivity and maturity ogives, and recruitment was driven by a Beverton-Holt relationship. The operating model was parameterized by data from the squat lobster population (*Pleuroncodes monodon*) off central Chile (Appendix II). This is a temperate zone galatheid that inhabits the continental shelf off central Chile and has a medium life span of about 7 years. Other important parameters that were used are: natural mortality =  $0.5 \text{ yr}^{-1}$ ,  $k = 0.197 \text{ yr}^{-1}$ ,  $L_{\infty} = 5.045 \text{ cm}$ , number of size intervals = 35. Sampling from the simulated population was performed using Baranov's catch equation. Numbers-at-age from the catch were assigned to the different equally spaced length intervals using a probability of length-given-age matrix (Fournier et al. 1998; see Appendix I for details). After constructing the deterministic length distribution of the catch, numbers collected at each length category were obtained by independently sampling the length intervals in a lognormal fashion ( $\sigma_{obs}$ , observation error levels of 0, 0.5, and 0.2 were used for scenarios A, B, and C-F, respectively). All the length frequency samples were taken at the beginning of each month and averaged to obtain an annual length frequency distribution. This was implemented to mimic the sampling process in a year-round fishery.

Non-equilibrium experiments were implemented as trends in fishing mortality rates (Fig. 1). Different scenarios included an equilibrium base case situation with a constant fishing mortality of 0.1 (scenario A), linear increasing trends in fishing mortality rate (to mimic a developing fishery) and linear increase in the first 10 years followed by a linear decrease (scenarios C-F). These trends are calculated by

$$F_{t+1} = F_t + F_{Step}$$

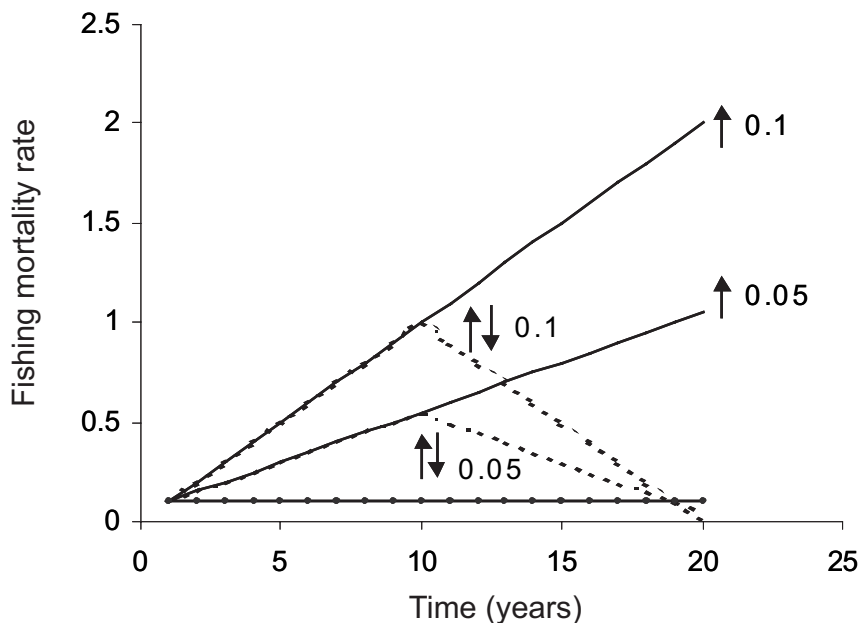
in which  $F_t$  is the fishing mortality rate in year  $t$ .

The simulation scheme is presented in Fig. 2 and considers 100 Monte Carlo trials for each case. Stock status was assessed by the equilibrium estimator in each time step using only LFD for that particular year; therefore there is no buildup of information along the time series.

Even though the mean square error is a more comprehensive statistic for evaluating estimator performance (Casella and Berger 1991), we focused on the bias component introduced by the equilibrium assumptions of the estimator. We used a relative error ( $\epsilon_{\theta}$ ) as a measure of performance represented by

$$\epsilon_{\theta} = \frac{(\theta_t^{Est} - \theta_t^{True})}{\theta_t^{True}}$$

where  $\theta_t^{Est}$  and  $\theta_t^{True}$  are the estimated (from JLCA) and true (from operating model) values respectively for the abundance at time  $t$  for one particular simulation run.

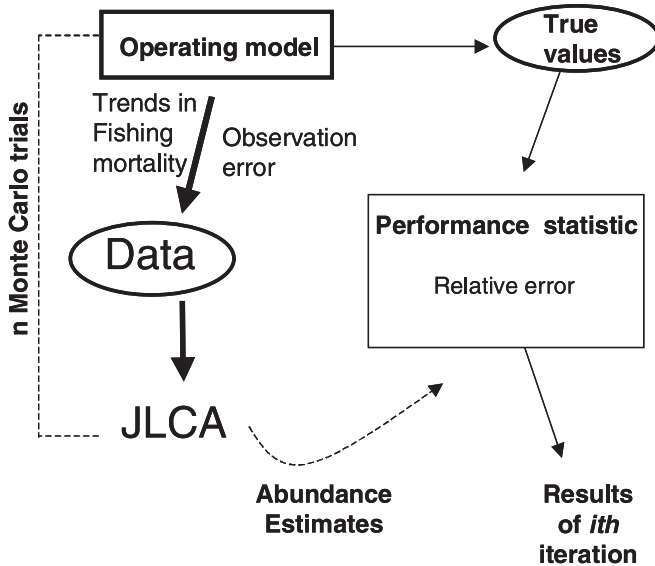


**Figure 1. Simulation of perturbation histories determined by trends in fishing mortality rates. Horizontal line corresponds to equilibrium scenarios with a constant fishing mortality rate of 0.1.**

### ***Length converted catch curve analysis***

Our second case study is presented to illustrate serious biases that arise in natural mortality estimation if we ignore the recruitment history and somatic growth dynamics of benthic resources. By reanalyzing and taking full advantage of the data we showed that equilibrium assumptions could be relaxed in favor of a statistical, integrated, and dynamic estimation procedure and we highlight bias levels in length-based estimates that arise under several data combinations aided also by simulation analysis.

Early uses of this approach are reviewed in Ricker (1975). This method assumes exponential mortality (constant), and therefore plotting the logarithms of the number of individuals caught (or sampled) against age results in a linear model with a slope equal to the total mortality ( $Z$ ). This method was extended to versions that use length data, for cases where age data are not available (Pauly 1984, see Sparre and Venema 1998). Basically, lengths are mapped into ages using available growth



**Figure 2.** Simulation scheme for assessing the performance of Jones' length cohort abundance estimator. One hundred simulation trials were performed and the observation error was incorporated on the catch-at-length data in a log-normal fashion ( $CV = 0.2$  for the last four panels of Fig. 3).

information. Key assumptions of this method include population at equilibrium, constant recruitment, and invariance of growth parameters across cohorts. Lasta et al. (2001) used this approach to estimate natural mortality using a combined size frequency distribution (SFD) from 1995 to 1998, by pooling bimonthly samples from an area closed to fishing. Here, we reconstructed their analysis and obtained further natural mortality estimates using size frequency distributions from each year independently (1995, 1996, 1997, and 1998) instead of pooling the data and discarding valuable information on the dynamics. In addition to the analysis of the raw data, we performed simulation analysis to assess the performance of an LCCCA equilibrium estimator. Simulated distributions were obtained from an age-structured operating model with known constant mortality, growth parameters, and either constant or variable recruitment (parameter estimates taken from Lasta et al. 2001 and Valero 2002). We organized the analysis into 11 scenarios (Table 3) that represent the estimation of natural mortality under different data conditions.



**Table 3. Mortality estimates and associated asymptotic 95% confidence limits using different scenario inputs.**

Scenario	Method	$M \text{ yr}^{-1}$	CI 95%
1	Integrated model (IM)	0.31	0.21-0.42
2	Simulated data equilibrium ( $M = 0.31$ )	0.29	0.22-0.37
3	1996	0.55	0.39-0.71
4	Pooled 1995-2000 SFD recruits predicted from IM	0.67	0.61-0.74
5	Lasta et al. (2001)	1.04	Not reported
6	Pooled 1995-1998 SFD data	1.04	0.94-1.14
7	1995	1.19	1.02-1.36
8	Pooled 1995-2000 SFD data	1.30	1.14-1.45
9	Pooled 1995-1998 SFD predicted from IM	1.34	1.20-1.49
10	1998	2.10	1.80-2.39
11	1997	2.43	2.15-2.71

SFD = length frequency data.

Scenario 1 corresponds to natural mortality estimation using real data and the statistical integrated age-structured size-based model described in the following section. In scenarios 3, 7, 10, and 11 only a single year of SFD was used and for scenarios 5, 6, and 8 averaging of SFD over several years was implemented (see Table 3). These seven scenarios were also based on real data. Simulation testing was introduced in scenarios 2, 4, and 9, where mortality estimates from the equilibrium estimator were based on simulated SFD data either under constant recruitment (scenario 2) or recruitment history as estimated by the integrated model (scenarios 4 and 9). Confidence bounds were estimated following Sparre and Venema (1998).

### *Statistical dynamic age-structured model and data*

We implemented an integrated age-structured size-based model largely following Fournier et al. (1998), in order to describe monthly dynamics of Patagonian scallop (*Zygochlamys patagonica*) abundance and shell growth. The model consisted of exponential survival equations, which account for natural mortality (the only source of mortality in this area), cohort-specific individual growth and variability of size-at-age and a logistic selectivity ogive of the sampling gear (Valero 2002). This model was fitted to data derived from bimonthly surveys in the *Reclutas* bed (central position: 39°24S, 55°56W, Argentina) in the northern Atlantic distribution of *Z. patagonica* from January 1995 to December 1998. Total bed area is approximately 1,020 km<sup>2</sup> (Valero 2002). The data set contains biological (i.e., size frequency distributions, total individual weights, and weights

of specified body fractions) as well as survey fisheries data (e.g., survey catch per unit effort, catch composition). Since this section of the *Reclutas* bed is closed to commercial fishing we assumed that the only source of mortality was related to natural events described as an exponential decay of cohorts as a function of a time invariant natural mortality parameter.

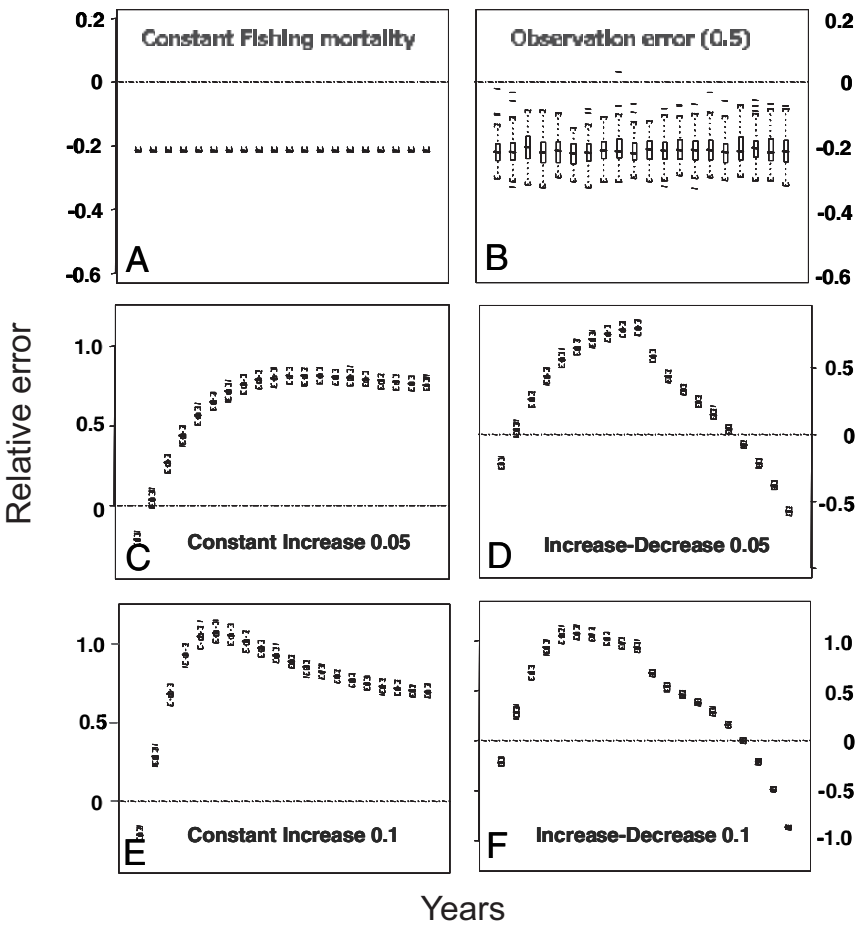
The integrated modeling approach allowed the estimation of cohort-specific growth parameters and an overall natural mortality rate in a dynamic and integrated way. Maximum likelihood theory (Fournier et al. 1998) was employed to derive model parameter estimates and their associated uncertainty using nonlinear optimization procedures. This model simulates population dynamics and allows for simultaneous estimation of natural mortality, research vessel selectivity, cohort specific shell growth parameters, and year class strength. The maximum likelihood estimate of natural mortality constituted scenario 1, and the associated overall parameterization (model and parameter values) was used to simulate SFD for scenarios 2, 4, and 9.

## Results

Jones' length cohort analysis and length converted catch curve analysis yielded biased results in cases that incorporated dynamic systems. JLCA was tested under trends in fishing mortality rates, simulating a developing fishery, and failed to accurately describe the abundance dynamics. On the other hand the performance of LCCCA, tested under variable recruitment (characteristic of sessile, highly spatially structured benthic populations), was poor and the size-based estimator overestimated natural mortality rates in all non-steady state scenarios. All these results assume a perfect knowledge of additional input parameters required to perform the analysis.

A 21% negative relative error was detected between the JLCA estimates and the operating model under equilibrium conditions (upper two panels of Fig. 3). This was attributed to underlying differences in model structure between the simulator and the equilibrium estimator, such as variability in length-at-age considered by the operating model versus deterministic growth assumed by JLCA. Observation error did not induce additional bias in the abundance estimates and only increased the dispersion around median values. Non-equilibrium scenarios mediated by fishing mortality rate trends strongly affected JLCA's ability to produce reliable abundance estimates. Increasing trends (Fig. 3, C and E) in fishing mortality induced a severe positive bias in the abundance estimates, and decreasing trends (Fig. 3, D and F) caused a negative bias in the abundance estimates. These changes were detected for low and moderate changes in fishing mortality rates (Fig. 3).

Natural mortality estimates derived from the integrated model and LCCCA are summarized in Table 3 and Fig. 4. Table 3 shows mortality es-



**Figure 3.** Time series (20 years) of relative error of abundance estimates (in numbers) derived from JLCA. First two panels correspond to equilibrium situations with (right) and without (left) observation error. Bottom four panels represent different perturbation histories induced by positive and/or negative trends in fishing mortality rates.

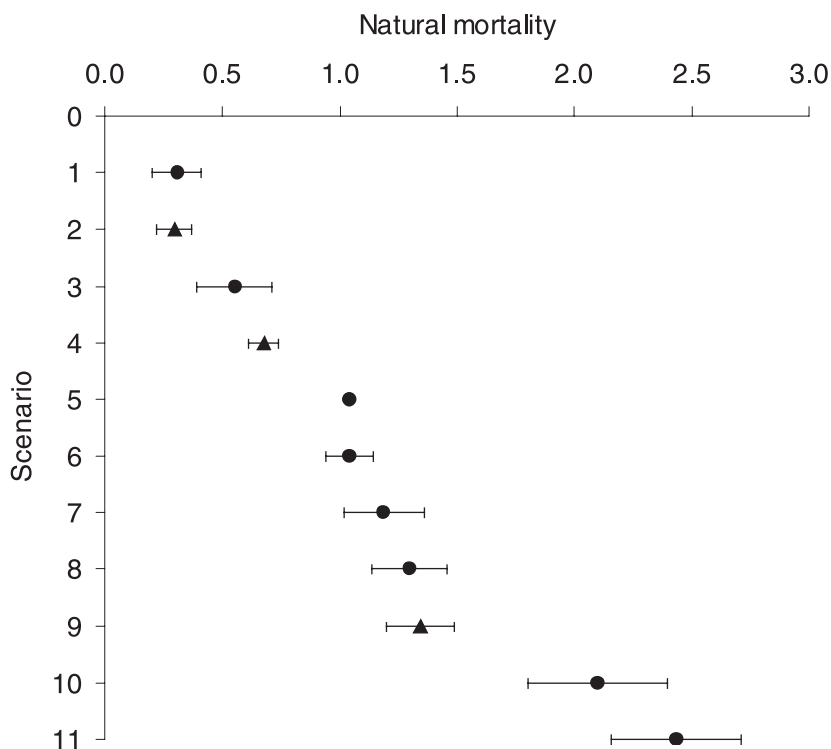
timates and associated asymptotic 95% confidence limits under different scenarios and data inputs. Only the scenario with constant recruitment (scenario 2) yielded results comparable with those obtained with the integrated model (scenario 1, Table 3 and Fig. 4). The mortality estimate of the integrated model is lower by a factor of three than previously reported by Lasta et al. (2001) (scenario 5, Table 3, Fig. 4). Mortality estimates obtained through LCCCA are greater than estimated by the integrated model by as much as a factor of 6 (scenarios 2-11, Table 3, Fig. 4). Simulation analysis showed that LCCCA lacks robustness to recruitment disequilibria and the confidence bounds derived from this approach are unrealistically small, not containing the true parameter value (Fig. 4).

## Discussion

Equilibrium-based estimates provided by both methods were seriously biased under realistic departures from model assumptions. Validations of these methods have traditionally involved mostly local sensitivity analysis to input parameters, therefore largely have not evaluated the fundamental model assumptions from these analyses. Non-equilibrium conditions are clearly expected in incipient fisheries, due to the characteristic increase in fishing effort during the developing phase. The results reported for JLCA can be generalized to other species. Ernst (2002) showed similar trends for two finfish species.

We showed that LCCCA produced serious biases when estimating natural mortality for the Patagonian scallop. This method is broadly used worldwide to provide total mortality estimates which are often used as proxies of stock productivity during the quota recommendation process. This fishery, as with many other fisheries in the world (i.e., Zheng et al. 2002), uses natural mortality to approximate target reference points, and therefore the serious bias levels reported in this work could have a negative impact in the appropriate fishery management.

Model testing using complex operating models has proven useful in model comparisons (Punt 1988), or to assess the effect of strong restrictive model assumptions. Evaluation of model performance and robustness under non-equilibrium scenarios calls for the implementation of simulation analysis (e.g., Monte Carlo approach). Parrack (1990) indicated that the list of unknown parameters might be so lengthy for some specific statistical models that determination might be a problem, and therefore Monte Carlo tests should be implemented to analyze an estimator's performance. Simulation analysis has been implemented in a number of stock assessment applications, ranging from models of similar structural complexity to comparative analysis across different models or methodologies (e.g., Hilborn 1979, Ludwig and Walters 1985, Butterworth 1988, Punt 1997, Maunder 2001, Ernst 2002).



**Figure 4.** Natural mortality estimates derived from LCCCA for all scenarios described in Table 3. MLE was obtained from the statistical calibration of the integrated dynamic age-structured model. Solid triangles represent mortality estimates derived from simulated data; solid circles were obtained from observed data; horizontal lines represent confidence bounds.

Jones (1987) used an equilibrium simulator to assess the performance of his method, and therefore ignored the effect of steady-state assumptions on parameter estimates. Our approach includes a more generic formulation of the state of nature and addresses deficiencies in model performance. Evaluation of model performance of other equilibrium size-based estimators through simulation analysis has been conducted in the past. Somerton and Kobayashi (1991) developed a simulation model to test the robustness of the Wetherall length-based method for recruitment variability and fishing effort trends.  $Z/K$  (ratio of total mortality rate to individual growth rate) estimates derived from this approach presented

large bias levels throughout their 15 year time series. Their results indicated that the Wetherall size-based estimator was very sensitive to recruitment and fishery disequilibria, a situation that is also present in both of our model testing results, and only recommended this approach under strict equilibrium conditions.

Temporal and/or spatial variability in recruitment is characteristic of small-scale, sedentary stocks, and spatially structured systems, nowadays called S-fisheries (Orensanz et al. 2005). Equilibrium methods are often used to assess the status of these populations due to lack of data or appropriate knowledge of alternative methods. It is unlikely that steady-state assumptions hold for these particular systems due to underlying dynamics or the sampling process scale. Caddy (1996) warned about the increasing use of equilibrium length-based methods and questioned the validity of some of their results, given that the underlying untested equilibrium assumptions probably are not met.

In addition to potential biases in abundance or mortality estimates associated with a model structure simplification, there is also an underutilization and lack of integration of input data. In order to match the data requirements of these methods, valuable auxiliary information is often ignored (Valero 2002). Additionally, the quantitative nature of the parameter estimates often creates overconfidence in managers, and discourages the development of further research. We disagree with Caddy (1996) and do not believe that managers are always critical and skeptical about the results of these assessments and their results often permeate directly into management.

Consequently scientists and managers can be misled and let the already available equilibrium models wrongfully dictate the kind of data collection or analysis required, instead of letting the available data dictate the nature of the models to be used to capture the dynamics and underlying uncertainty. Due to biases inherent to these methods, alternative robust approaches should be encouraged in data-limited situations.

Equilibrium assumptions will continue to play a crucial role for exploring and understanding dynamic systems behavior and this practice should be differentiated from attempting to estimate key population parameters under steady state assumptions (Caddy 1996). In light of the current results we do not advocate the use of either of these two equilibrium methods to assess stock status and/or estimate key demographic rates (perhaps only when equilibrium can be demonstrated); but rather encourage scientists to critically review the conditions under which these models are applied (Somerton and Kobayashi 1991).

## Acknowledgments

BE was supported by a Chilean government scholarship (Presidente de la Republica) and JLV by scholarships of Fulbright, OAS (Organization of American States), and Fundación Antorchas (Argentina). April Wilkinson, Jeanne Simon, and Heather Honeycutt helped in editing the manuscript. Two anonymous reviewers and the invited editor of this volume helped to improve the quality and readability of this manuscript.

## References

- Beverton, R.J.H., and S.J. Holt. 1956. A review of methods for estimating rates in exploited fish populations, with special reference to sources of bias in catch sampling. Rapp. P.-V. Reun. Cons. Int. Explor. Mer 140:67-83.
- Bluhm, B.A., D. Piepenburg, and K. von Juterzenka. 1998. Distribution, standing stock, growth, mortality and production of *Strongylocentrotus pallidus* (Echinodermata: Echinoidea) in the northern Barents Sea. Polar Biol. 20:325-334.
- Butterworth, D.S. 1988. Some recommendations regarding the assessment methodologies used by ICSEAF. Collect. Sci. Pap. ICSEAF Recl. Doc. Sci. CIPASE colec. Doc. Cient. CIPASO 15(1):107-155.
- Caddy, J.F. 1996. Regime shifts and paradigm changes: Is there still a place for equilibrium thinking? Fish. Res. 25:219-230.
- Casella, G., and R.L. Berger. 1991. Statistical inference. Duxbury Press. 650 pp.
- Daskalov, G., and K. Prodanov. 1994. Stock assessment of *Sprattus sprattus* L. off Bulgarian Black Sea coast, using length cohort analysis. Marine population dynamics. Report on the first meeting of the CIESM working group DYNPOP, Tunisia. 10-14 September 1994, Zaragoza, Spain. CIHEAM 1995 10:71-72.
- Demestre, M., and J. Leonart. 1993. Population dynamics of *Aristeus antennatus* (Decapoda: Dendrobranchiata). Sci. Mar. (Barc.) 57:183-189.
- Ehrhardt, N., and J. Ault. 1992. Analysis of two length-based mortality models applied to bounded catch length frequencies. Trans. Am. Fish. Soc. 121:115-122.
- Ernst, B. 2002. An investigation on length-based models used in quantitative population modeling. Ph.D. thesis, University of Washington, Seattle. 150 pp.
- Etim, L., P.E. Lebo, and R.P. King. 1999. The dynamics of an exploited population of a siluroid catfish (*Schilbe intermedius* Reupell 1832) in the Cross River, Nigeria. Fish. Res. 40:295-307.
- Etim, L., R.P. King, and M.T. Udo. 2002. Breeding, growth, mortality and yield of the mudskipper *Periophthalmus barbarus* (Linnaeus 1766) (Teleostei: Gobiidae) in the Imo River estuary, Nigeria. Fish. Res. 56:227-238.
- Fournier, D.A., J. Hampton, and J.R. Sibert. 1998. MULTIFAN-CL: A length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. Can. J. Fish. Aquat. Sci. 55:2105-2116.

- Gonzalez-Pajuelo, J.M., J.M. Lorenzo-Nespereira, and M. Mendez-Villamil Mata. 1996. Determination of the state of exploitation of the common pandora *Pagellus erythrinus* (Linnaeus, 1758) in Gran Canaria waters using length cohort analysis. *Bol. Inst. Esp. Oceanogr.* 12(2):115-130.
- Haughton, M.O., and D.P.F. King. 1992. Population biology of the spiny lobster (*Panulirus argus*) in Jamaican waters. In: M.H. Goodwin, S.M. Kau, and G.T. Waugh (eds.), *Proceedings of the Forty-first Annual Gulf and Caribbean Fisheries Institute*, St. Thomas, U.S.V.I., November 1988. 41:528-542.
- Hilborn, R. 1979. Comparison of fisheries control systems that utilize catch and effort data. *J. Fish. Res. Board Can.* 36:1477-1489.
- Hilborn, R., and C.J. Walters 1992. *Quantitative fisheries stock assessment: Choice, dynamics and uncertainty*. Chapman and Hall, New York. 570 pp.
- Honkalehto, T. 1989. A length cohort analysis for walleye pollock based on empirical estimation of cannibalism and predation by marine mammals. *Proceedings of the International Symposium on the Biology and Management of Walleye Pollock*. Alaska Sea Grant College Program, University of Alaska Fairbanks, pp. 651-665.
- Jones, R. 1984. Assessing the effect of changes in exploitation pattern using length composition data (with notes on VPA and cohort analysis). *FAO Fish. Tech. Pap.* 256.
- Jones, R. 1987. An investigation of length composition analysis using simulated length compositions. In: D. Pauly and G.R. Morgan (eds.), *Length-based methods in fisheries research*. International Center for Living Aquatic Resource Management (ICLARM), Kuwait Institute for Scientific Research. 468 pp.
- Lai, H.L., and V.C. Gallucci. 1988. Effects of parameter variability on length-cohort analysis. *J. Cons. Int. Explor. Mer* 45:82-91.
- Lasta, M., J. Valero, T. Brey, and C. Bremec. 2001. *Zygochlamys patagonica* beds on the Argentinian shelf. Part II: Population dynamics of *Z. patagonica*. *Arch. Fish. Mar. Res.* 49:1-13.
- Limbong, D., K. Hayashi and Y. Matsumiya. 1988. Length cohort analysis of common mackerel, *Scomber japonicus*, Tsushima Warm Current stock. *Bull. Seikai Reg. Fish. Res. Lab. Seisui Kenpo* 66:119-133.
- Ludwig, D., and C.J. Walters. 1985. Are age structured models appropriate for catch-effort data? *Can. J. Fish. Aquat. Sci.* 42:1066-1072.
- Mancera, E., and J. Mendo. 1996. Population dynamics of the oyster *Crassostrea rhizophorae* from the Cienaga Grande de Santa Marta, Colombia. *Fish. Res.* 26:139-148.
- Maunder, M.N. 2001. A general framework for integrating the standardization of catch per unit of effort into stock assessment models. *Can. J. Fish. Aquat. Sci.* 58:795-803.
- Mees, C.C. 1993. Population biology and stock assessment of *Pristipomoides filamentosus* on the Mahe Plateau, Seychelles. *J. Fish Biol.* 43:695-708.



- Megrey, A. 1989. A review and comparison of age-structured stock assessment models from theoretical and applied points of view. *Am. Fish. Soc. Symp.* 6:8-48.
- Methot, R.D. 2000. Technical description of the stock synthesis assessment program. NOAA Tech. Memo. NMFS-NWFSC-43. 46 pp.
- Orensanz, J.M., A.M. Parma, G. Jerez, N. Barahona, M. Montecinos, and I. Elias. 2005. What are the key elements for the sustainability of "S-fisheries"? Insights from South America. In: N. Ehrhardt (ed.), *The scientific bases for the sustainability of fisheries*. *Bull. Mar. Sci. (Special issue)* 76:527-556.
- Parrack, M.L. 1990. Estimating stock abundance from size data. Ph.D. thesis, University of Washington, Seattle. 116 pp.
- Pauly, D. 1984. Length converted catch curves: A powerful tool for fisheries research in the tropics. Part 2. *Fishbytes* 2:17-19.
- Pauly, D., and G.R. Morgan (eds.). 1987. Length-based methods in fisheries research. International Center for Living Aquatic Resources Management, ICLARM Conference Proceedings 13, Manila. 468 pp.
- Pierce, G.J., and A. Guerra. 1994. A review of the stock assessment methods used for cephalopod fisheries. *ICES-CM* 1993/K:33.
- Punt, A. E. 1988. Model selection for the dynamics of southern African hake resources. M.S. thesis, University of Cape Town, Department of Applied Mathematics, South Africa.
- Punt, A.E. 1990. Is  $B1 = K$  an appropriate assumption when applying an observation error production model estimator to catch-effort data? *S. Afr. J. Mar. Sci.* 9:249-259.
- Punt, A.E. 1997. The performance of VPA-based management. *Fish. Res.* 29:217-243.
- Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. *Can. Dep. Env., Fish. Mar. Ser., Bull.* 191, Ottawa. 382 pp.
- Sarda, F., J. Lleonart, and J.E. Cartes. 1998. An analysis of the population dynamics of *Nephrops norvegicus* (L.) in the Mediterranean Sea. *Sci. Mar. (Barc.)* 62:135-143.
- Somerton, D., and D. Kobayashi. 1991. Robustness of the Wetherall length-based method to population disequilibria. *Fish. Bull. U.S.* 89:307-314.
- Sparre, P., and S.C. Venema. 1998. Introduction to tropical fish stock assessment. Part I: Manual. FAO Tech. Pap. 306/1 Rev. 2. 470 pp.
- Sullivan, K.J. 1991. The estimation of parameters of the multispecies production model. In: N. Daan and M.P. Sissenwine (eds.), *Multispecies models relevant to management of living resources*. *ICES Mar. Sci. Symp.* 193:185-193.
- Valero, J.L. 2002. Temporal and spatial growth variation and natural mortality estimation with an integrated dynamic model for the Patagonian scallop (*Zygochlamys patagonica*). M.S. thesis, University of Washington, Seattle. 154 pp.

- Wakwabi, E. 1990. Stock assessment from applying length based cohort and catch curve analyses. Proceedings of the Workshop on the Management of the Shallow Water Shrimp Fishery of Tanzania, July 18-27, 1990, Victoria Seychelles. October 1990, SWIOP Document 0150, RAF 87/008. FAO 58:60-68.
- Xu, X., and S.Y. Chiu. 1995. Population estimates with variances of Spanish sardine along the southern Fujian Coast, China, from length-based cohort analysis. *Trans. Am. Fish. Soc.* 124(3):437-452.
- Zheng, J., S. Siddeek, D. Pengilly, and D. Woodby. 2002. Overview of recommended harvest strategy for snow crabs in the eastern Bering Sea. Alaska Department of Fish and Game, Division of Commercial Fisheries, Juneau, Regional Information Report 5J02-03.

## Appendix I. The population dynamics model

The simulation model considered in this appendix is based on a dynamic age-structured model. A probability matrix of length-given-age was used to represent the variability around the mean length-at-age. Yearly samples were collected and passed to the JLCA abundance estimator to assess stock size during a simulation framework of 20 years.

### A. Basic population dynamics

The abundance at age was propagated using the following equations

$$N_{a+1,t+1} = N_{a,t} \exp\{-(M + F_t s_a)\} \quad (I.1)$$

$$\begin{aligned} N_{A,t} = & N_{A-1,t-1} \exp\{-(M + F_{t-1} s_{A-1})\} \\ & + N_{A,t-1} \exp\{-(M + F_{t-1} s_A)\} \end{aligned} \quad (I.2)$$

where

$N_{a,t}$  is abundance at age  $a$  and time  $t$ ,

$F_t$  is instantaneous fishing mortality rate at time  $t$ ,

$M$  is instantaneous natural mortality rate,

$A$  is the age of the plus group,

$s_a$  is selectivity at age  $a$ , represented by the following logistic model:

$$s_a = \frac{1}{1 + \exp\left(-\left[\frac{\ln(19)(a - A_{50})}{(A_{95} - A_{50})}\right]\right)} \quad (I.3)$$

$A_{50}$  and  $A_{95}$  are age of 50 and 95% of recruitment to the fishing gear.

### B. Recruitment

Population renewal was modeled after a Beverton-Holt recruitment model

$$R_{t+1} = \frac{S_t}{\alpha + \beta S_t} \quad (I.4)$$

where

$R_t$  is recruitment in year  $t$ ,

$S_t$  is spawning biomass at time  $t$ ,

$$S_t = \sum_a w_a \phi_a N_{a,t} \quad (I.5)$$

$w_a$  is average weight of an individual at age  $a$  computed from the mean length-at-age and allometric parameters [ $= \exp(\delta_1)(L_a)^{\delta_2}$ ],

$\phi_a$  is the maturity at age  $a$  computed by the following logistic model:

$$\phi_a = \left\{ \frac{1}{1 + e^{a_1 + a_2 a}} \right. \quad (1.6)$$

$\alpha_1$  and  $\alpha_2$  are parameters of the logistic maturity ogive.

The parameters of the Beverton-Holt model were re-parameterized as virgin recruitment ( $R_0$ ), virgin spawning stock size ( $S_0$ ) and steepness ( $z$ ) by using the following equations:

$$\alpha = S_0 \frac{1 - z}{4z R_0} \quad (1.7)$$

$$\beta = \frac{5z - 1}{4z R_0} \quad (1.8)$$

### C. Initial conditions

Initial conditions allowed for a historic equilibrium exploitation rate

$$N_{a,1} = R_{2-a} \prod_{i=1}^{a-1} \exp[-(M + F_{init} s_i)] \quad a > 1 \quad (1.9)$$

$$N_{A,1} = \frac{R_{2-A} \prod_{i=1}^{A-1} \exp[-(M + F_{init} s_i)]}{\{1 - \exp[-(M + F_{init} s_A)]\}} \quad (1.10)$$

### D. Length frequency samples

Baranov's catch equation was used to compute numbers-at-age in the catch for each time interval,

$$C_{a,t} = N_{a,t} \frac{F_t S_a}{F_t S_a + M} (1 - \exp[-(F_t S_a + M)]) \quad (1.11)$$

Numbers-at-age were assigned to the different equally-spaced length intervals by using a probability of length given age matrix, parameterized with a mean length-at-age ( $\mu_a$ ) and a variance of length-at-age ( $\sigma_a^2$ ):

$$P(l|a) = \frac{(2\pi\sigma_a^2)^{-1/2} \exp\left\{-\left[\frac{(l-\mu_a)^2}{2\sigma_a^2}\right]\right\}}{\sum_{l=1}^l (2\pi\sigma_a^2)^{-1/2} \exp\left\{-\left[\frac{(l-\mu_a)^2}{2\sigma_a^2}\right]\right\}} \quad (I.12)$$

$$\mu_a = L_{\infty} (1 - e^{-k(a-a_0)})$$

$L_{\infty}$ ,  $k$  and  $a_0$  are standard von Bertalanffy growth parameters. In order to constrain the parameter space,  $\sigma_a^2$  were computed based on  $\sigma_1^2$  and  $\sigma_A^2$  and a linear relationship with age.

Numbers collected at each length category were obtained by independently sampling the length intervals in a lognormal fashion

$$N_{l,t} = [C_{a,t} P(l|a)] \exp\left\{\varepsilon_{l,t}^{obs} - \frac{\sigma_{obs}^2}{2}\right\} \quad (I.13)$$

where

$\varepsilon_{l,t}^{obs}$  is a normal deviate with mean zero and variance  $\sigma_{obs}^2$ .

## Appendix II. Model parameters

**The following list contains all model parameters used in the equations of Appendix I.**

Simulation parameter	Data set
Upper length interval	4.4
Lowest length interval	1.0
Step size for length interval	0.1
$M$	0.5
$k$	0.197
$L_{\infty}$	5.045
$a_0$	-0.51
$A$	7
$\delta_1$	-8.642
$\delta_2$	3.31
$R_0$	$10^7$
$Z$	0.7
$\alpha_1$	13.65
$\alpha_2$	-0.502
$\sigma_1$	0.17
$\sigma_A$	0.27
$A_{50}$	3.1
$A_{95}$	4.0
$F_{init}$	0.1
Number of length categories	35

# Using a Spatially Structured Model to Assess the Tasmanian Fishery for Banded Morwong (*Cheilodactylus spectabilis*)

**Malcolm Haddon, Philippe Ziegler, Jeremy Lyle, and Paul Burch**

*Tasmanian Aquaculture and Fisheries Institute (TAFI), University of Tasmania, Hobart, Tasmania, Australia*

## Abstract

Fisheries can be regarded as data poor if insufficient information is available to produce a defensible stock assessment. Data-limited fisheries can be (1) new fisheries with no time-series of information; (2) low-value fisheries for which little data are collected; (3) bycatch fisheries, which are often ignored; and (4) spatially structured fisheries where data collected may not be representative of the whole stock. This last category is especially problematic because when little data are available, stock assessments are generally limited to simple models. Unfortunately, the impact of fishing and our ability to assess a fishery can be greatly affected when the dynamics of a stock or stocks are influenced by some spatial structure. In the instance studied here, there are similarities to studies of fished areas adjacent to marine protected areas closed to fishing. A simulation study of a spatially structured fishery was undertaken, using the Tasmanian banded morwong fishery to mimic such a data-limited situation. It was found that the characteristics of the catch (age-structure and sex ratio), which were available through a relatively focused research program, could only be fitted using a model that included a spatially structured population, with fished and unfished components (based on depth) and movement in between. Based on such a fishery assessment, the stock was seen to be declining. This outcome contrasted greatly with the outcome of considering only the catch and effort data typically available for such inshore reef fisheries in Australia. The catch and catch rate information, when modeled as coming from a spatially homogeneous

single population, gave the appearance of relative stability and sustainability. Fisheries for species likely to have some spatial structuring, such as widely distributed reef dwelling species, may possibly be far more complex and difficult to manage than the classic theory of homogeneous fish populations would suggest. Some form of explicit spatial management would be a more appropriate approach to ensuring long-term sustainability in such fisheries.

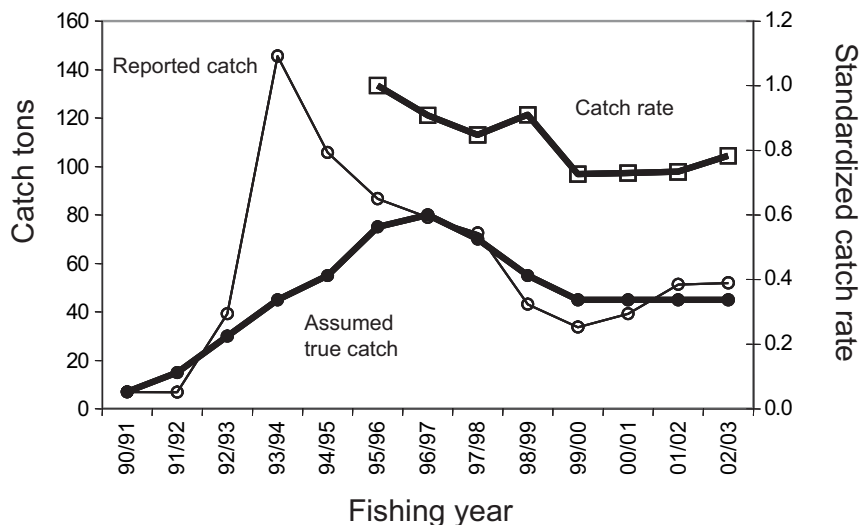
## Introduction

A fishery is data poor if insufficient information is available to produce a defensible stock assessment. This definition moves away from a simple statement about the quantity of data and focuses instead on the purpose for which data are collected. Insufficient information for an assessment can arise because (1) a fishery is new or developing and a time series of information has yet to be collected; (2) sufficient specific information is unavailable in low-value fisheries irrespective of how long they have operated (research is expensive and extensive data collection is difficult to justify in such circumstances); (3) data collection tends to focus on the target species so bycatch fisheries constitute another common data-poor category, even though they may constitute a significant proportion of a target fishery's catch; and (4) fisheries that are spatially or otherwise structured where an assumption of spatial homogeneity has been adopted. Where a stock has spatial structure, simple data collections may not be representative of the whole and thus could not be used validly to assess the whole (e.g., Orensanz and Jamieson 1998, Parma et al. 2003). A spatially structured fishery seems likely to cause extra problems in data-poor circumstances because when little data are available, stock assessments are generally limited to simple models. Unfortunately, our ability to assess a fishery can be greatly affected when the dynamics of a stock or stocks are influenced by some spatial structure (Punt 2003). Spatial structuring of a stock can occur in many ways, making it difficult to produce a simple solution to the complexities present in the instance studied here. We present an example from a spatially structured (both along coast and onshore and offshore), low-value temperate reef fish fishery. The dynamics of the fishery, which involves onshore fished areas and offshore unfished areas, imply there are similarities to studies of fished areas adjacent to marine protected areas closed to fishing (Beverton and Holt 1957, Botsford and Hobbs 1986, Quinn et al. 1993, Haddon et al. 2003, Gerber et al. 2003). The objectives of this present work were to illustrate problems that can arise when one assumes that available data are representative of the dynamics of a spatially homogeneous stock when it is in reality spatially structured (Punt 2003), and to suggest a simple strategy that can be adopted in data limited situations to reduce the risk of generating inappropriate management advice.



Many reef fish populations demonstrate spatial structuring, with limited movement between reefs after settlement (e.g., McCormick and Choat 1987, Jones 1988, Newman and Williams 2001). Banded morwong (*Cheilodactylus spectabilis*), is such a species in New Zealand and south-eastern Australia. Principally in Tasmania, banded morwong supports a small-scale fishery that services domestic live-fish markets. The fishery has a short history, being first developed in the early 1990s and producing a current annual catch of about 50 t (see Ziegler et al. 2005). The only data generally available for the scalefish fisheries in Tasmania are the legally required commercial catch and effort data. Fortunately, banded morwong has, in addition, been the subject of a previous research project (Murphy and Lyle 1999), and of recent fishery-independent catch samples, which have provided sex ratio, age-structure, and location data relating to the catch. Banded morwong exhibit an unusual combination of very fast initial growth, an early age of maturity and a long life expectancy, with maximum ages greater than 80 years for both sexes (Murphy and Lyle 1999). In addition, growth rates and maximum sizes are distinctly different for the two sexes. Because of the availability of these research data, banded morwong was chosen as an example reef fish fishery with which to test a range of fishery performance indicators for data-poor fisheries that usually only have commercial catch and effort statistics, sometimes of questionable integrity, plus limited availability of biological information.

To test a range of fishery performance measures (see Ziegler et al. 2005), a detailed model description of the stock was developed—an operating model—in order to generate artificial data sets for the comparisons. However, in the process of developing this detailed description, it proved difficult to fit a relatively simple model to all the observed biological characteristics of the catch. Catch rates (CPUE) and verbal reports concerning the fishers' impressions of the status of the fishery, did not exhibit any recent trend (Fig. 1). However, the observed age-structure of the catch, which included many relatively old fish of both sexes of 50 years and above, suggested that while some stock depletion had occurred since the advent of the fishery, there was still an accumulation of older fish available. The numbers of animals above 15 years of age had been reduced but the age-structure of these older fish was still relatively flat. Despite this, the observed change to the sex ratio from 70% at the start of the fishery to just below 50% females in the catch in the most recent years of the fishery implied significant stock depletion, at least in some areas. The change in sex ratio came about both through a reduction in the proportion of older females (>15 years old) from about 65% to about 30% in the catch and an increase in the proportion of younger fish in the catches. Whether the increases in younger fish were a result of stock depletion allowing more younger fish to be seen or were due to increases in absolute recruitment success was unknown. Whatever the explanation, a change

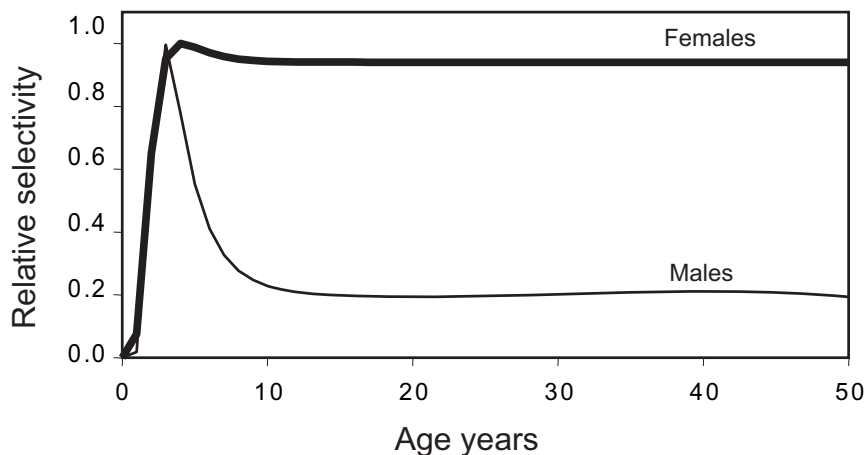


**Figure 1.** The reported catch compared to the catch time series used in all the simulations. The extremes reported in 1993/1994 and 1994/1995 were reduced to remove false catch returns. The time series of assumed true catches is meant to represent the typical development of a fishery. The reported catch rates (CPUE) were standardized relative to 1995/1996.

in the dynamics of the fishery had occurred and a simple model was unable to provide an adequate description of the available data. Before the structure of the operating model could be fully developed, these apparent inconsistencies between the various data sources had to be resolved.

The quality of the catch and effort statistics from the fishery's early period is highly questionable (Ziegler et al. 2005) and catch rate data are only available from 1995/1996 onward, so the early part of the fishery is poorly known (Fig. 1). Limited entry by license was introduced in 1998 with reported catch history used in the license allocation process. This use of catch history had been anticipated by fishers prior to allocation and led to some of them generating large amounts of fictitious "paper" fish, artificially inflating the catch statistics (Fig. 1).

The fishery is currently managed by limited entry licenses, a two-month spawning season closure (March and April), and a keyhole size limit currently between 360 and 460 mm fork length. Given the selectivity of the usual fishing gear (gillnets), and sex-based growth rate characteristics (Murphy and Lyle 1999), this size limit range implies that almost all female fish remain in the fishery after they reach legal size, whereas



**Figure 2.** A comparison of the relative selectivity by age for male (fine line) and female (thick line) banded morwong from the east coast of Tasmania (in 137 mm gillnets). Lines are derived from translating selectivity by length into selectivity by age.

the males mainly grow beyond the upper size limit by the age of 11 to 12 (Fig. 2). This keyhole size-limit regulation combined with the growth characteristics of the species enable the use of sex ratio as a performance indicator for this fishery, with an expectation of a predominance of females in the catch, which certainly occurred early in the fishery. The keyhole size limit gives rise to discarding with the potential for some discard mortality. However, discard mortality is assumed to be trivially low because fishers set their nets in shallow waters and continuously retrieve fish soon after they are caught in an effort to maintain high quality for the live fish market.

The changes to the age-structure and changes to the sex ratio, observed in research sampling, suggested different levels of depletion. Finding a simple explanation or model that fitted or was consistent with the observations did not appear possible when only a homogeneous population was considered. A somewhat more complex explanation and model was that a proportion of the fish population lived in an unfished depth refuge and there was some movement between the two sub-populations. Thus, older individuals were supplied to the open population through leakage out of the unfished area. This seems plausible because the known depth distribution of banded morwong extends well below the depths that are fished. Fishers generally restrict their activities to depths less than 20 to 25 meters to avoid problems of barotrauma in this live-fish

trade. Exploration was made of the implications of subdivided populations for simple performance indicators.

## Methods

A standard age-structured population model was constructed (Quinn and Deriso 1999, Haddon 2001) that subdivided the available biomass between a fished portion and an unfished portion with movement occurring between. This was used to investigate the effect of subdivision and movement on simple catch statistics and other fishery performance measures.

For the simulations, the proportion of the total suitable habitat used by each of the sub-populations was used as a proxy for the initial distribution of the virgin biomass. Thus, 50:50 implies that half the initial biomass lies in each of the two sub-populations, while 40:60 implies 40% is fished and 60% is unfished. These proportions were also used to determine the distribution of the recruits, and thus, effectively, of the initial spawning biomass. In addition, the relative area available to each sub-population influenced the movement dynamics. Each of the two sub-populations had its own dynamics but recruitment was generated as a function of the total spawning biomass combined. Movement between the sub-populations was a combination of mobility rate and the relative proportion of suitable habitat into which the animals can move. A further assumption was that the age-structure and sex ratio was identical in the fished and unfished areas when the fishery began. The base case, against which everything else was compared, was one with a mobility of 1.0 (completely mobile) or of habitat ratio of 100:0. This is equivalent to the special case where the fished population is equivalent to the whole population, as with a non-spatial model.

Growth is described in terms of length at age and weight at length by the standard von Bertalanffy equation:

$$L_{s,t} = L_{s,\infty} \left( 1 - e^{-K_s(t-t_0^s)} \right) + \epsilon \quad (1)$$

where the  $L_{s,t}$  is the length of sex  $s$  at age  $t$ ,  $L_{s,\infty}$  is the average maximum length for the species for each sex  $s$ ,  $K_s$  is the Brody growth coefficient by sex  $s$ ,  $t_0^s$  is the age at a hypothetical length of zero for each sex  $s$ , and  $\epsilon$  is a normal random residual. The weight at length relationship for each age  $t$ , for each sex  $s$ , is described by:

$$W_t^s = aL_{s,t}^b \quad (2)$$

where the  $a$  and  $b$  coefficients define the relationship between length and weight.

A logistic model is used to describe maturity at age for females, which is used to determine the mature or spawning biomass:

$$\mu_t = (1 + e^{c+dt})^{-1} \quad (3)$$

where  $\mu_t$  is the proportion of age class  $t$  that is sexually mature, and  $c$  and  $d$  are the maturity parameters. Given these, and a knowledge of numbers in each population and sex, then the mature or spawning biomass in year  $y$  can be determined using:

$$B_{s,y} = \sum_{t=0}^{t_{\max}} \sum_{p=1}^2 \mu_t W_t N_t^{p,s} \quad (4)$$

where  $B_{s,y}$  is the spawning biomass in year  $y$ ,  $N_t^{p,s}$  is the number of fish in population  $p$  of age  $t$  where the sex  $s$  is female,  $t_{\max}$  is the maximum age (plus-group). The numbers at age 0 at the start of year  $y$  are a function of the total mature or spawning biomass across both fished and unfished populations at the start of year  $y$ :

$$R_{0,y} = \frac{B_{s,y}}{(\alpha + \beta B_{s,y})} e^{\epsilon_y} \quad (5)$$

where  $\alpha$  and  $\beta$  are the parameters of the Beverton-Holt stock recruitment relationship,  $B_{s,y}$  is the total spawning biomass in both the fished and unfished populations in year  $y$ , and  $\epsilon_y$  is the recruitment residual for year  $y$ . The values for  $\alpha$  and  $\beta$  are determined from the steepness of the stock recruitment relationship ( $h$ ) and the pre-exploitation biomass ( $B_0$ ) using Francis' (1992) method (Haddon 2001). For the purposes of this work, which was primarily about searching for parameter combinations that gave rise to realistic outputs, recruitment variability was set to zero so that recruitment was deterministic.

Numbers at age under equilibrium conditions of no fishing mortality and constant average recruitment are described by the standard equation modified by an expression summarizing movement between populations:

$$N_t^{p,s} = \begin{cases} \pi_p \bar{R}_{s,0} & t = 0 \\ N_{t-1}^{p,s} e^{-M} & 1 \leq t \leq t_{\max} - 1 \\ N_{t_{\max}-1}^{p,s} e^{-M} / (1 - e^{-M}) & t = t_{\max} \end{cases} \quad (6)$$

where  $N_t^{p,s}$  is the numbers of fish in population  $p$ , of sex  $s$ , and of age  $t$ ,  $\pi_p$  is the proportion of available habitat and hence of recruits to be found in population area  $p$ ,  $\bar{R}_{s,0}$  is the average recruitment of sex  $s$  and age 0,  $t_{max}$  is the maximum age modeled (the plus-group), and  $M$  is the instantaneous rate of natural mortality, assumed constant across all ages. Recruitment variability has been omitted from equation 6. The symbol  $t_{max}$  is referred to as the plus age-group because it combines age's  $t_{max}$  and all older ages that are not modeled explicitly. Equation 6 is combined with the movement dynamics to initiate the populations.

Movement is modeled as occurring at the end of each year and is restricted to mature fish (all females mature before entering the fishery). Movement between populations is a combination of mobility  $m$ , defined as the proportion of the mature population that becomes vagrant or mobile, and is capable of shifting from each population to adjoining populations, combined with  $\pi_p$ , the proportion of habitat/biomass in each population  $p$  into which the animals can move (also used to subdivide the recruits between populations). Thus, the movement rate from population 1 into population 2 can be represented as  $m\pi_1$ . The dynamics of movement can be represented formally for each age class as:

$$N_t^{s,p} = (1 - m\pi_{p+1})\mu_t N_t^{s,p} + m\pi_p \mu_t N_t^{s,p+1} \quad (7)$$

Thus, population  $p$  retains  $1 - m\pi_{p+1}$  of its total but gains  $m\pi_p$  of the neighboring population,  $p + 1$ . With only two populations,  $p$  and  $p + 1$  are the only possibilities. If the proportion of habitat is equal (i.e.,  $\pi_p = 0.5$ ) then the movement rate equals the mobility, however, if the proportional distribution of the population deviates from 50:50 then the movement rates will become asymmetric. This approach to describing movement was selected because it represents the degree of mobility in the population and determines where the fish will finally settle by the total area over which it can spread (which includes the area in which it began to move). Thus, a fish may begin to move and its probability of settling in one of the available areas is related to the relative area inhabited by the two populations.

The keyhole legal size range combined with the early rapid growth has a strong effect on selectivity at age. The gillnet mesh selectivity is best described using a gamma distribution (Millar and Holst 1997):

$$V_t^s = \left( \frac{l_t}{\gamma kn} \right)^\gamma e^{\left( \gamma - \frac{l_t}{kn} \right)} \quad (8)$$

where  $V_t^s$  is the selectivity of age class  $t$  for sex  $s$ ,  $l_t$  is the length of age class  $t$ ,  $n$  is the mesh size of the nets used, and  $\gamma$  and  $k$  are the se-

lectivity parameters (different for each sex). This gillnet mesh selectivity is modified at the legal limits for each sex. Together with the different growth pattern of the males and females this leads to a large difference in their selectivity characteristics (Fig. 2). The females remain vulnerable to the fishery for much longer than the males.

The fishing mortality rate for each age class is defined in terms of the fully selected instantaneous fishing mortality rate  $F_y$  in year  $y$ , combined with the selectivity  $V_t^s$  for each age class  $t$  and sex  $s$ :

$$F_{y,t}^s = V_t^s F_y \quad (9)$$

The catch in each year  $y$  is defined as the sum of the predicted catch at age multiplied by the weight at age:

$$C_y = \sum_{s=1}^2 \sum_{t=0}^{t_{\max}} W_t \frac{F_{y,t}^s}{F_{y,t}^s + M} N_t^{p,s} \left( 1 - e^{-(M+F_{y,t}^s)} \right) \quad (10)$$

while the dynamics of the two populations are described by the combination of:

$$N_{t,y+1}^{p,s} = \begin{cases} \pi_p \bar{R}_{s,0,y+1} & t=0 \\ N_{t-1,y}^{p,s} e^{-(M+F_{y,t}^s)} & 1 \leq t \leq t_{\max}-1 \\ N_{t_{\max}-1,y}^{p,s} e^{-(M+F_{y,t}^s)} + N_{t_{\max},y}^{p,s} e^{-(M+F_{y,t}^s)} & t=t_{\max} \end{cases} \quad (11)$$

remembering that fishing mortality only acts on one population  $p$ . Equation 11 is combined with the equation for movement, with movement acting at the end of each year to generate the final numbers at age in each population.

$$N_{t,y+1}^{s,p} = (1 - m\pi_{p+1}) \mu_t N_{t,y+1}^{s,p} + m\pi_p \mu_t N_{t,y+1}^{s,p+1} \quad (12)$$

The exploitable biomass is defined as the fishable biomass before any mortality in a given year had been imposed:

$$B_y^E = \sum_{s=1}^2 \sum_{t=0}^{t_{\max}} W_t V_t^s N_t^{p,s} \quad (13)$$

The model was fit to the available data by setting the initial unfished biomass, the mobility rate, and the relative proportion of fished and unfished

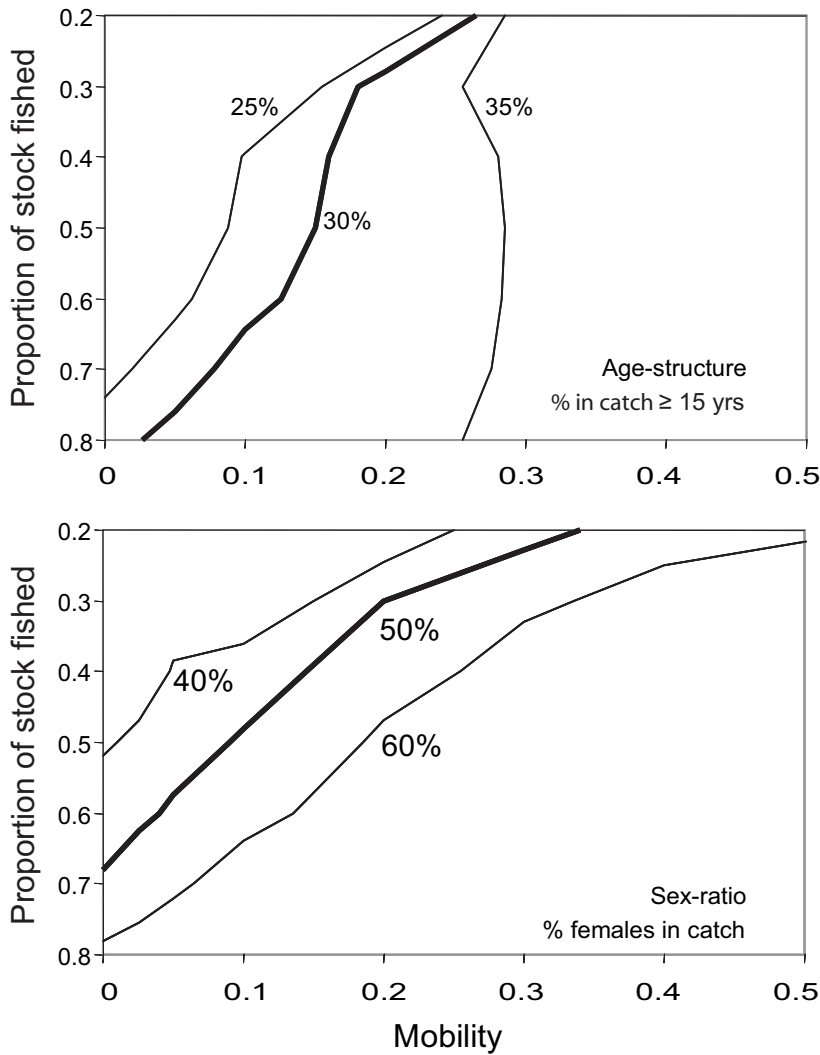
populations and then pursuing an iterative search for the set of fully selected fishing mortality rates that generated the observed catches. The sex ratio, proportion of fish in the plus group, the exploitable biomass and total spawning biomass in each populations, through the years of the fishery were saved, graphed, and inspected.

The effect of spatially distributing the spawning biomass was investigated by generating sub-populations at different distribution ratios (fished to unfished ratios) and growing the population for 13 years (the approximate age of the fishery) with different mobility levels. The outputs of interest, in effect representing fishery performance indicators, were the sex ratio of the catch and the proportion of the female catch of 15 years and older (a simple proxy for age-distribution). These were suggested by the changes observed in both sex ratio and the proportion of older fish in the research sampling conducted on the east coast of Tasmania.

Because the model was over-parameterized for the available data, instead of directly fitting the model to data, a single catch history was used for all simulations and many different combinations of initial biomass, mobility rates, and fished : unfished biomass ratios were examined. The catch history used in all simulations (assumed catch in Fig. 1) was similar to that observed in the fishery (reported catch in Fig. 1) with the extreme catches, most likely to be based on false data, removed. In each case, the model was conditioned on catch by searching for the schedule of annual fishing mortality rates required with each combination of parameters that would give rise to the catch time-series used in each simulation. This conditioning on catch, under different combinations of initial biomass, mobility, and fished : unfished ratios, led to different degrees of depletion, different age-structures, and different sex ratios of the catch following the 13 years of the fishery prosecuted upon the open population. The model outcomes for the final year were plotted as contour maps relative to the mobility and proportion of the stock fished (e.g., Fig. 3). The combined contour maps for both the measure of sex ratio and of age-structure enabled the optimal combinations of mobility and proportion of stock fished to be identified for each level of initial biomass (e.g., Fig. 5). The predicted values from each combination of model parameters were also compared (using minimal squared residuals) across all years where observed values were available. The combinations that led to the closest fit to the final year's values were also those that fitted best to all available data.

As no estimate of initial or unfished biomass was available the simulations were run with three different levels: 1,000 t, 1,200 t, and 1,500 t. These bracketing values were chosen because they enabled the model to predict sex ratios of 70:30 in the first year as well as the predicted levels (~65%) of females in the 15+ age group. The levels of mobility that were included ranged from zero to 1.0 (no mobility to completely mobile).





**Figure 3. Simulations with 1,200 t of unfished biomass. Contours of the percent of females  $\geq 15$  years of age (top panel) and of the percent of females in the catch or sex ratio (bottom panel). The approximate levels observed in research sampling are indicated by heavy lines.**

**Table 1. Proportion of females (sex ratio) in the catch after 13 years of fishing under different combinations of mobility and fished proportion of the total stock.**

Fished proportion	Mobility						
	0.00	0.05	0.10	0.20	0.30	0.40	1.00
0.20					0.484	0.532	0.593
0.30				0.480	0.543	0.566	0.599
0.40		0.410	0.469	0.533	0.563	0.580	0.602
0.50	0.386	0.460	0.507	0.555	0.577	0.588	0.604
0.60	0.459	0.509	0.539	0.571	0.587	0.594	0.604
0.70	0.517	0.545	0.563	0.583	0.593	0.598	0.606
0.80	0.557	0.573	0.581	0.593	0.599	0.602	0.606

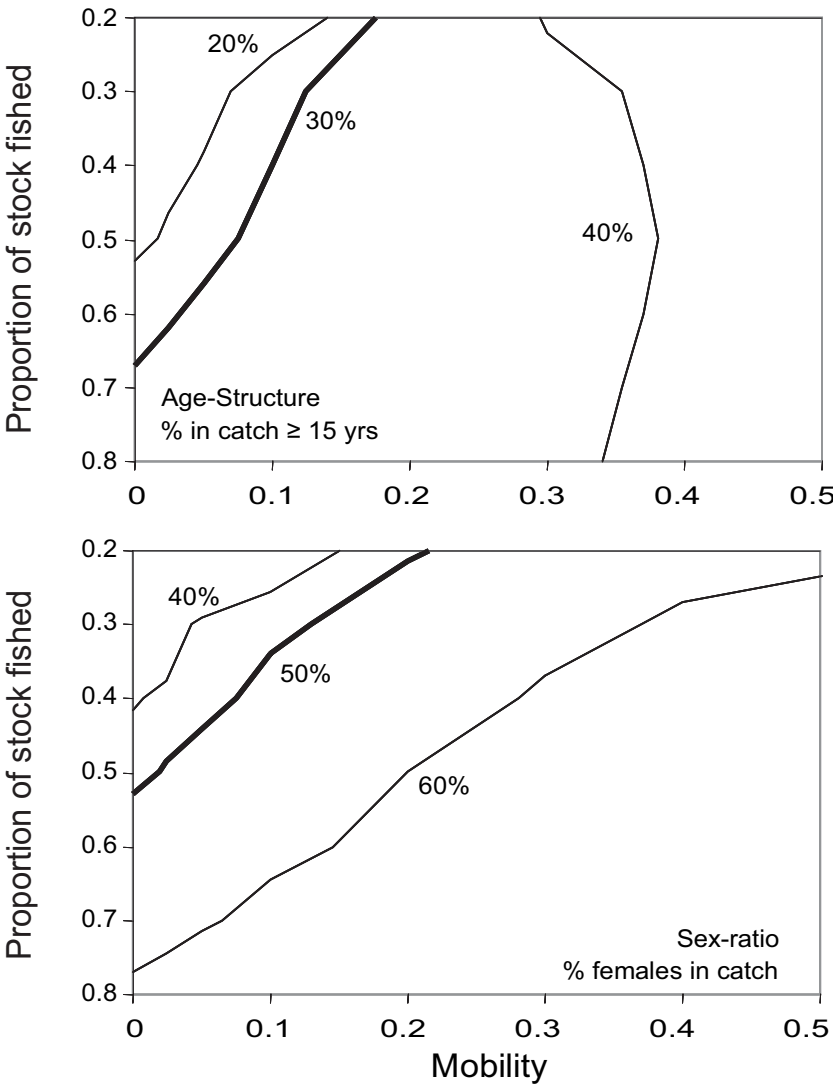
The observed sex ratio was approximately 50%. The unfished biomass in all cases was 1,200 t. Empty cells were combinations for which realistic solutions were not possible.

Finally, the proportion of the available biomass exposed to fishing ranged from 0.2 to 0.8 (20-80%).

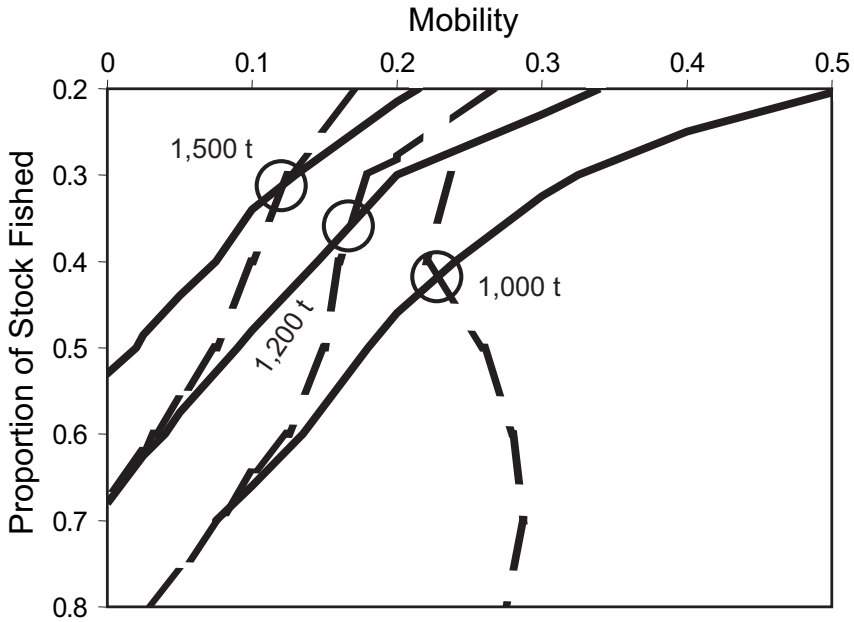
## Results

By conducting trials of many combinations of the parameters (initial biomass, mobility, relative size of the two populations) it was possible to tabulate predicted values of two performance indicators, the sex-ratio and the proportion of the catch aged 15 years or greater (e.g., Table 1). From such tables, contour diagrams of different levels of the two performance indicators were constructed (Figs. 3 and 4). In each case, the proportion of the stock that was fished had less effect upon the proportion of the females  $\geq 15$  years of age than upon the sex ratio. This is shown by the contours being more near to vertical in the age-structure graph than with the sex ratio (Figs. 3 and 4).

The contour lines relating to the final values in the research sampling after the 13-year fishery, i.e., 50% females in the sex ratio and 30% females  $\geq 15$  years of age, are highlighted in the graphs. Combinations of unfished biomass, mobility, and proportion of the stock exposed to fishing that would give rise to the observed population characteristics of sex ratio and age-structure (Fig. 5) were identified by where the contour lines overlap. As the unfished biomass was increased to 1,500 t, the overlap of the two types of contours occurred at decreasing levels of both the mobility and the proportion of the stock fished. Which particular combination is the most realistic would require more information. Under no combinations



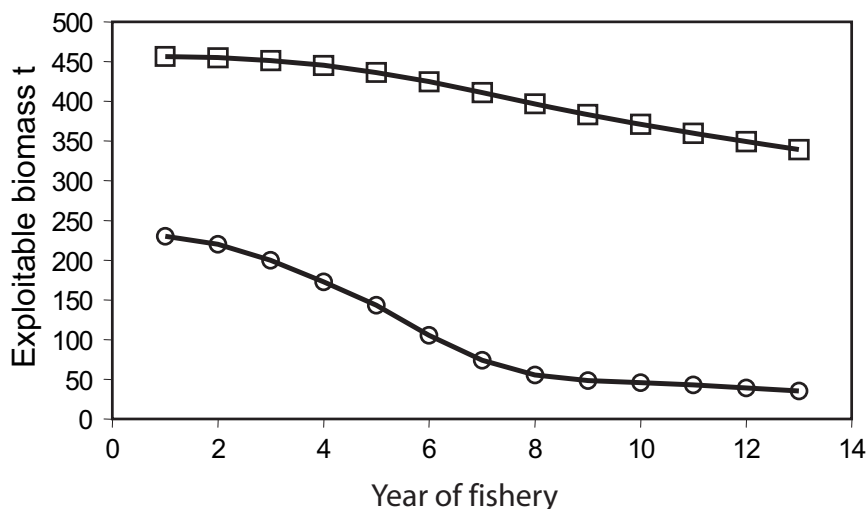
**Figure 4.** Simulations with 1,500 t of unfished biomass. Contours of the percent of females  $\geq 15$  years of age (top panel) and of the percent of females in the catch or sex ratio (bottom panel). The approximate levels observed in research sampling are indicated by the heavy lines.



**Figure 5.** The overlap of the 30% contours for the percent of females  $\geq 15$  years of age (dashed lines) and of the 50% contours of females in the catch or sex-ratio (solid lines) for three different levels of unfished biomass (1,000 t, 1,200 t, and 1,500 t). The circles are centered over the points of overlap.

were the required conditions through time of sex ratio and of age-structure obtained without some degree of spatial structuring of the population and some degree of mobility between them.

In the case where the initial unfished biomass was 1,200 t and with the mobility level and proportion of stock fished that gave rise to the observed values of sex ratio and age-structure, the exploitable biomass in the fished population steadily declined until it appeared to be relatively stable at a low level after the 8th year of the fishery. The exploitable biomass in the unfished population only began to decline in the 5th year of the fishery after which it declined at a steady rate (Fig. 6). Assuming a constant catchability in the fishery, this pattern of depletion would also be reflected in accurately reported catch rates.



**Figure 6.** Exploitable biomass in metric tons against year of the fishery, in the population open to fishing (circles) and that unfished (squares). This simulation was with an unfished biomass of 1,200 t, a fished proportion of 34% and a mobility of 0.1575, conditions that lead to a sex-ratio of approximately 50% females and 30% females of age 15 years of greater.

## Discussion

Data-poor fisheries are not uncommon. Unfortunately, because the effects of spatial structuring of fished stocks are rarely taken into account when collecting data, many such stocks could also be considered to be data poor. In this context, even if extensive data collections are available, data-poor fisheries may thus be even more common than generally believed. In essence, the assumption that any data that are collected are representative of a complete population is called into doubt when a stock is spatially structured. Walters (2003) pointed out that large biases (up or down) may arise in stock assessments based on catch rate (CPUE) data when the spatial distribution of effort and of the stock is ignored. The fishery for the Tasmanian banded morwong is an extreme example of spatial structuring of a fished stock and of the information about that stock. Walters's (2003) comments, when applied to spatially structured data-poor fisheries, are not encouraging. The fishery for the Tasmanian banded morwong, being a site attached reef dwelling species, provided an opportunity for demonstrating the effects of spatial structuring on

our ability to assess an exploited population. Even though insufficient fishery data were available to properly fit a spatially explicit assessment model, it was possible to explore the impacts of parameters such as the mobility of the species, the initial unfished biomass, and the proportion of the stock exposed to fishing.

It was found that spatially structured populations, with movement between the populations, can generate fisheries data (catch rates, age structure, and sex ratios) that appear to be inconsistent with each other when using a non-spatially structured assessment model. This would appear to be a general conclusion to any spatially structured stock with fished and unfished areas. Significant leakage of larger, older animals into a fished area can obscure the depletion of older age-classes. In the case of banded morwong, significant numbers of 50-plus-year-old female fish among numerous young fish being caught, led initially to the idea that the stock was not depleted. It appeared to suggest that there had been a large accumulation of biomass that was in the early stages of being fished down. However, this did not tally with the expectation of low productivity from such a long-lived species or with a closer inspection of biological data (large changes to the sex-ratio). If a fished stock is spatially structured such that a significant proportion is effectively unfished, but there is some movement between fished and unfished populations, then clearly, standard data collection and assessment methods based on simple catch statistics that ignore spatial structuring would lead to an invalid assessment (Walters 2003, Walters and Bonfil 1999). Orensanz and Jamieson (1998) emphasize that for relatively sessile species the assumptions of most assessment models are inappropriate. In particular the assumptions that a unit stock is closed to immigration and emigration and that the effects of the fishing process and the distribution of the fished stock are both homogeneous within the region occupied by the stock, are all inappropriate. Such models and assumptions preclude the possibility of serial or localized depletion, which, in a spatially structured stock is a real possibility. Unfortunately, many published studies of spatially structured fisheries (Booth 2000, Berkeley et al. 2004, Smith and Rago 2004) are clear in the requirement for large amounts of detailed data from across the range of the fishery. We explored the implications of spatial structuring in the absence of such information.

In the model developed here, the amount of movement out of the fished and unfished areas at the start of fishing, before any depletion had occurred, was in proportion to their relative areas (hence in equilibrium). However, if the overall mobility is low enough, the fished part of the stock is quickly depleted and the fishery soon becomes dependent upon the growth of recruits in the fished part of the stock and the movement of larger fish out of the unfished area and into the fished area. In this case, the apparent relative stability arising from a consideration of catch and effort data obscures continued depletion of the total stock, at least in the

short-term. This result is not novel as it has similarities to the outcomes of modeling marine protected areas where there is a fished and an unfished structure to a stock (Gu  nette et al. 1998, Gu  nette and Pitcher 1999, Walters and Bonfil 1999). However, this is a novel result in terms of being expressed on a reef-based fishery. If a stock assessment was conducted in ignorance of the spatial structuring and its consequences, the present effort levels and their associated catch might be deemed sustainable. If projected forward in time, this situation would continue until, from the point of view of an assessment assuming a homogeneous fished population, the fishery would inexplicably collapse. If reef-based fish stocks have a greater tendency to be spatially structured this problem seems likely to be relatively common.

The simplest conclusion to draw from these simulations is that ignorance of the biology and behavior of a fished species can lead to risk-prone fisheries management. If insufficient weight is given to the implications of the biology of a fished species, then any assessment is likely to produce bad management advice. It would appear to be stating the obvious to suggest that a review of the known biology and other fisheries for a particular species be undertaken in data limited situations. However, care needs to be taken in such meta-analyses to ensure that there are at least some close similarities between the species being compared. King and McFarlane (2003) explore the possibility of categorizing species into five different life history strategies in a manner analogous to the classic *r*- and *K*-strategies but identifying strategies that relate more to the marine than the terrestrial environment. This approach holds the promise of being more reliable than simply assuming a related species will be similar, but it would require at least some biological investigation. Unfortunately, it would appear that investigating spatial structuring is essential for the stock assessment of all fisheries, including data-poor situations. Tagging studies in combination with even crude habitat mapping have great value in providing insights into the potential distribution of fished species for comparison with reports of where fishing occurs (habitat maps could be of use to many fisheries in an area). Of course, if fisheries are data poor because of a lack of resources then habitat mapping is unlikely to occur. Even with such information available the question remains how to include spatial information into an assessment. In the particular case of spatially structured populations some form of precautionary spatial management suggests itself. While managers learn about the fishery and its effects, which might take decades, it would be risk averse to require the fishery to be focused in a delineated area leaving significant proportions of any known habitat unfished for the species in question. This is not a recommendation for the creation of marine protected areas as fishing for other species could occur in the areas closed to, say, banded morwong fishing. Rather, it is a way of implementing the precautionary approach (Garcia and Grainger 1997) while allowing some fishing to occur, thus enabling

managers to learn about the productivity of the species and how it reacts to fishing. In addition, management plans need to be developed that account for the potential spatial structuring of the fishery. Some form of adaptive management is required, especially in a relatively new fishery (Collie and Walters 1991). Using the fishery to learn about the stock implies having some form of data gathering available, at least about landings and possibly effort.

Management objectives concerning sustainability are commonly so general that they provide little guidance in cases where there is spatial structuring in fisheries. Like most current stock assessment models, such objectives often reflect an implicit assumption of spatial homogeneity. For example, does sustainable exploitation mean sustainable in each sub-population or sustainable in the context of the complete species distribution? These two options imply different things and would lead to different management. If it is desired to exploit small-scale, low valued, spatially structured fish stocks, then management objectives that recognize the possibility of stocks being spatial structured must evolve that reflect the risk management strategies being implemented.

## References

- Berkeley, S.A., M.A. Hixon, R.J. Larson, and M.S. Love. 2004. Fisheries sustainability via protection of age structure and spatial distribution of fish populations. *Fisheries* 29:23-32.
- Beverton, R.J.H., and S.J. Holt. 1957. On the dynamics of exploited fish populations. *Fish. Investig. Ser. II G.B. Minist. Agric. Fish.* 19:1-533.
- Booth, A.J. 2000. Incorporating the spatial component of fisheries data into stock assessment models. *ICES J. Mar. Sci.* 57:858-865.
- Botsford, L.W., and R.C. Hobbs. 1986. Static optimization of yield per recruit with reproduction and fishing costs. *Fish. Res.* 4:181-189.
- Collie, J.S., and C.J. Walters. 1991. Adaptive management of spatially replicated groundfish populations. *Can. J. Fish. Aquat. Sci.* 48:1273-1284.
- Francis, R.I.C.C. 1992. Use of risk analysis to assess fishery management strategies: A case study using orange roughy (*Hoplostethus atlanticus*) on the Chatham Rise, New Zealand. *Can. J. Fish. Aquat. Sci.* 49:922-930.
- Garcia, S.M., and R. Grainger. 1997. Fisheries management and Sustainability: A new perspective or an old problem? In: D.A. Hancock, D.C. Smith, A. Grant, and J.P. Beumer (eds.), *Developing and sustaining world fisheries resources: The state of science and management*. 2nd World Fisheries Congress Proceedings. CSIRO Publishing, Collingwood, Australia, pp. 631-654.
- Gerber, L.R., L.W. Botsford, A. Hastings, H.P. Possingham, S.D. Gaines, S.R. Palumbi, and S. Andelman. 2003. Population models for marine reserve design: A retrospective and prospective synthesis. *Ecol. Appl.* 13(1):S47-S64.



- Gu  nette, S., and T.J. Pitcher. 1999. An age-structure model showing the benefits of marine reserves in controlling overexploitation. *Fish. Res.* 39:295-303.
- Gu  nette, S., T. Lauck, and C. Clark. 1998. Marine reserves: From Beverton and Holt to the present. *Rev. Fish Biol. Fish.* 8:251-272.
- Haddon, M. 2001. Models and quantitative methods in fisheries. CRC/Chapman & Hall, Boca Raton. 406 pp.
- Haddon, M., C. Buxton, C. Gardner, and N. Barrett. 2003. Modelling the effect of introducing MPAs in a commercial fishery: A rock lobster example. In: J.P. Beumer, A. Grant, and D.C. Smith (eds.), *Aquatic protected areas. What works best and how do we know?* University of Queensland, St Lucia, pp. 428-436.
- Jones, G.P. 1988. Ecology of rocky reef fish of north-eastern New Zealand: A review. *N. Z. J. Mar. Freshw. Res.* 22:445-462.
- King, J.R., and G.A. McFarlane. 2000. Marine fish life history strategies: Applications to fishery management. *Fish. Manag. Ecol.* 10:249-264.
- McCormick, M.I., and J.H. Choat. 1987. Estimating total abundance of a large temperate-reef fish using visual strip-transects. *Mar. Biol.* 96:469-478.
- Millar, R.B., and R. Holst. 1997. Estimation of gillnet and hook selectivity using log-linear models. *ICES J. Mar. Sci.* 54:471-477.
- Murphy R.J., and J.M. Lyle 1999. Impact of gillnet fishing on inshore temperate reef fishes, with particular reference to banded morwong. University of Tasmania, Marine Research Laboratories, Tasmanian Aquaculture and Fisheries Institute, Report for the Fisheries Research Development Corporation 95/145. 136 pp.
- Newman, S.J., and D.M. Williams. 2001. Spatial and temporal variation in assemblages of Lutjanidae, Lethrinidae and associated fish species among mid-continental shelf reefs in the central Great Barrier Reef. *Mar. Freshw. Res.* 52:843-851.
- Orensanz, J.M.L., and G.S. Jamieson. 1998. The assessment and management of spatially structured stocks: An overview of the North Pacific Symposium on Invertebrate Stock Assessment and Management. In: G.S. Jamieson and A. Campbell (eds.), *Proceedings of the North Pacific Symposium on Invertebrate Stock Assessment and Management*. Can. Spec. Publ. Fish. Aquat. Sci. 125:441-459.
- Parma, A.M., J.M.L. Orensanz, I. El  as, and G. Jerez. 2003. Diving for shellfish and data: Incentives for the participation of fishers in the monitoring and management of artisanal fisheries around southern South America. In: S.J. Newman, D.J. Gaughan, G. Jackson, M.C. Mackie, B. Molony, J. St. John, and P. Kailola (eds.), *Towards sustainability of data-limited multi-sector fisheries*. Australian Society for Fish Biology Workshop Proceedings, Bunbury, Western Australia, 23-24 September 2001. Department of Fisheries, Perth, Western Australia, Fisheries Occasional Publications 6:8-29.
- Punt, A.E. 2003. The performance of a size-structured stock assessment method in the face of spatial heterogeneity in growth. *Fish. Res.* 65:391-409.

- Quinn, T.J., and R.B. Deriso 1999. Quantitative fish dynamics. Oxford University Press. 542 pp.
- Quinn, J.F., S.R. Wing, and L.W. Botsford. 1993. Harvest refugia in marine invertebrate fisheries: Models and applications to the red sea urchin, *Strongylocentrotus franciscanus*. Am. Zool. 33:537-550.
- Smith, S.J., and P. Rago. 2004. Biological reference points for sea scallops (*Placopecten magellanicus*): The benefits and costs of being nearly sessile. Can. J. Fish. Aquat. Sci. 61:1338-1354.
- Walters C.J. 2003. Folly and fantasy in the analysis of spatial catch rate data. Can. J. Fish. Aquat. Sci. 60:1433-1436.
- Walters, C.J., and R. Bonfil. 1999. Multispecies spatial assessment models for the British Columbia groundfish trawl fishery. Can. J. Fish. Aquat. Sci. 56:601-628.
- Ziegler, P.E., J.M. Lyle, M. Haddon, and P. Burch. 2005. Developing assessments and performance indicators for a small-scale temperate reef fish fishery. In: G.H. Kruse, V.F. Gallucci, D.E. Hay, R.I. Perry, R.M. Peterman, T.C. Shirley, P.D. Spencer, B. Wilson, and D. Woodby (eds.), Assessment and management of new and developed fisheries in data-limited situations. Alaska Sea Grant College Program, University of Alaska Fairbanks. (This volume.)