Investigation of possible redesign of the FIMS of the Rock Lobster resource of South Africa, taking additional variance into account

A. Brandão and D.S. Butterworth

Marine Resource Assessment & Management Group (MARAM)
Department of Mathematics and Applied Mathematics
University of Cape Town
Rondebosch 7701, Cape Town

January 2010

Abstract

The additional (inter-leg) variance for FIMS is calculated for each of the four Zones covered by the survey. This allows evaluation of how the overall CV for these surveys would change if the number of survey legs was increased, with a compensating decrease in the number of stations sampled per leg. This would achieve appreciable improvements in precision for surveys of the Dassen Island and Lambert's Bay Zones, but the benefits for the Cape Point and Saldahna Bay Zones are only slight.

Introduction

Possible ways to redesign the FIMS survey in order to be able to decrease the variance of the survey abundance indices of rock lobster are investigated by looking at different number of stations sampled and number of survey legs carried out in a year, each of which will affect the total variance the CPUE indices. The estimated sampling variances of the FIMS weighted mean Catch-Per-Unit-Effort (CPUE) indices reported in Brandão and Butterworth (2009) are likely to be negatively biased as they do not take into account any correlation within each leg of the FIMS survey. In this paper an additional variance component of the FIMS CPUE indices is estimated for each Zone that takes this into account. The total variance for the CPUE indices consists of these two components of variance, the sampling variance and the additional variance.

Data

The weighted mean CPUE for each leg in a particular Zone (together with the associated sampling standard error) for the period 1992/93 to 2008/09 is used to estimate the additional variance for each Zone. The methodology to obtain the weighted mean CPUE and the standard errors is described in Brandão and Butterworth (2009). The mean CPUE for the second leg of the Dassen Island survey in 2007 is close to zero and much lower than the index for the first leg as well as lower than indices for other years. It was thus decided to consider this value as an outlier. Results are reported both including and excluding this outlier when estimating additional variance.

Methodology

The calculation of the standard errors of the weighted mean CPUE series described in Brandão and Butterworth (2009) does not take into account any correlation within each leg of the survey. Since the repeated sampling at some stations occurs in a relatively short time period (two weeks), it is likely that catch rates will be positively correlated as environmental conditions change over a larger time period. This in turn leads to negatively biased estimates of the sampling variance. To be able to determine whether the FIMS design should be altered in some way to be able to decrease the variance of the survey abundance indices of rock lobster, both sources of variance need to be investigated.

Estimation of additional variance

The FIMS survey abundance indices of rock lobster for each Zone (Dassen Island, Lambert's Bay, Saldanha Bay and Cape Point) and each leg (together with the associated standard errors) are used to estimate the additional variance present in the FIMS survey over and above the sampling variance alone. The model used assumes that the observed abundance indices are lognormally distributed about their expected values:

$$I_{y,i} = qe^{\eta_{y,i}} B_y e^{\varepsilon_{y,i}}, \quad (1)$$

where

$I_{y,i}$ is the observed FIMS abundance index year y and leg i,

B_y is the corresponding model estimate of abundance of the resource for year y,

q is the expected catchability coefficient for lobsters,

$\varepsilon_{y,i}$ is normally distributed with mean zero and standard deviation $\sigma_{y,i}$, whose values are given by the sampling CV for each $I_{y,i}$ and are input, and

$\eta_{y,i}$ is normally distributed additional variance arising from fluctuations of *q* with mean zero and standard deviation σ_{add}.

Equation (1) can be re-written as:

$$\ln\left(I_{y,i}\right) = \alpha_y + \eta_{y,i} + \varepsilon_{y,i} \tag{2}$$

Maximum likelihood estimates (MLE) of α_y and σ_{add} can be obtained by minimising the negative of the log-likelihood function (ignoring constants):

$$-\ln L = \sum_y \sum_i \left\{ \frac{1}{2(\sigma_{add}^2 + \sigma_{y,i}^2)} \left(\ln I_{y,i} - \alpha_y\right)^2 + \ln\left(\sqrt{\sigma_{add}^2 + \sigma_{y,i}^2}\right) \right\}, \tag{3}$$

However, MLEs for variance are generally biased and the method of restricted (or residual) maximum likelihood (REML), which provides estimates which are unbiased, is more appropriate. REML estimates are also more reliable when dealing with small samples, as is the case here. REML minimises the negative likelihood of linear combinations of the observed values whose expectations are zero. This has the effect of adjusting for the available degrees of freedom. In this simple linear model, the REML negative log-likelihood function is given by:

$$-\ln L = \sum_y \sum_i \left\{ \frac{1}{2(\sigma_{add}^2 + \sigma_{y,i}^2)} \left(\ln I_{y,i} - \alpha_y\right)^2 + \ln\left(\sqrt{\sigma_{add}^2 + \sigma_{y,i}^2}\right) \right\} + \sum_y \ln\left(\sum_i \left(1 \Big/ \left(\sigma_{add}^2 + \sigma_{y,i}^2\right) \right) \right), \tag{4}$$

Possible re-designs of the FIMS survey

The FIMS survey design at present consists of annual surveys each with two legs where each leg has *n* sampling stations. To investigate possible alterations to the FIMS design, the effect on the CVs for the abundance indices by changing the number of legs in a year and the intensity of survey sampling is considered. A change in the number of legs in a year will affect the impact of the additional variance $\left(\sigma_{add}^2\right)$ on the average over legs of the survey index each year, while a change in the number of sampling stations will affect the size of the sampling variance $\left(\sigma_{y,i}^2\right)$. Changes in the number of surveys and the number of sampling stations investigated in this paper are restricted to options when the "effort" in performing the FIMS surveys remains the same (at least in terms of the combined number of legs and stations).

For the purposes of the exercise performed in this paper, a general sampling variance over all years is used in the computations described below. The value for this general sampling variance $\left(\bar{\sigma}^2\right)$ for the Zone considered is taken to be given by the geometric mean of all sampling variances over all years and all legs (i.e. all $\sigma_{y,i}^2$). The total variance (i.e. sampling variance and additional variance) for an

abundance index for a single leg under different sampling stations / number of survey legs scenarios is then given by:

$$\begin{aligned} \text{normal/normal} &\Rightarrow \left(\sigma^*\right)^2 = \bar{\sigma}^2 + \sigma_{add}^2 \\ \text{half/double} &\Rightarrow \left(\sigma^*\right)^2 = 2\bar{\sigma}^2 + \sigma_{add}^2 \\ \text{quarter/4 times} &\Rightarrow \left(\sigma^*\right)^2 = 4\bar{\sigma}^2 + \sigma_{add}^2 \end{aligned} \quad . \tag{5}$$

As it is assumed that the observed abundance indices are lognormally distributed, the estimates of variance referred to in equation (5) are on a logarithmic scale. Estimates of CV for the mean CPUE for each leg on the original scale (which take additional variance into account) are then of the form:

$$CV = \sqrt{\exp(\sigma^*)^2 - 1}\,. \tag{6}$$

The mean CPUE index for the survey in a particular year is the average of the mean CPUE in each leg, and so the CV for the mean CPUE (under the different sampling stations / number of legs scenarios) is given by:

$$\begin{aligned} \text{normal/normal} &\Rightarrow CV^* = \sqrt{CV^2/2} \\ \text{half/double} &\Rightarrow CV^* = \sqrt{CV^2/4} \\ \text{quarter/4 times} &\Rightarrow CV^* = \sqrt{CV^2/8} \end{aligned} \quad . \tag{7}$$

Results

Table 1 reports the FIMS CPUE indices for each individual Zone for rock lobsters measuring more than 60 cm together with their standard errors (in round brackets) as presented in Brandão and Butterworth (2009), together with standard errors that take into account the additional variance in these estimates due to the fact that samples within each leg of the survey will be positively correlated (in square brackets). Table 2 shows CV values for the mean CPUE estimate under the different sampling stations / number of legs scenarios. For all Zones except for Saldanha Bay, CVs can be steadily reduced by increasing the number of legs performed in a year, with a corresponding decrease in the number of stations sampled so that the "effort" involved in carrying out the FIMS surveys remains constant[1]. For Saldanha Bay, halving the number of stations and doubling the number of legs in a year decreases the mean CPUE CV but not when a quarter of the stations are sampled and the number of legs is increased four times; this difference from the behaviour for the other Zones arises because of the larger sampling variance for Saldahna Bay.

1 In practice the half / double option would increase effort somewhat, as the approach taken here assumes that only the time required to take a sample need be considered, whereas time needed to travel between stations also needs to be taken into account.

Reference

Brandão, A. and Butterworth, D.S. 2009. Re-analysis of the Fisheries Independent Monitoring Survey of the rock lobster resource of South Africa. Marine and Coastal Management Document: MCM/2009/AUG/SWG/WCRL/11b.

Table 1. FIMS CPUE series for each individual Zone and their corresponding standard errors given within round brackets when no additional variance is estimated and within square brackets when this variance is estimated. Estimates of additional variance for Dassen Island are based on omitting the outlier in 2007.

Year	Zone			
	Cape Point	**Dassen Island (outlier omitted)**	**Saldanha Bay**	**Lambert's Bay**
1992/93	140.75 (17.30) [25.83]	24.89 (4.370) [15.21]	2.720 (0.871) [3.722]	3.228 (1.233) [13.25]
1993/94	128.18 (13.47) [22.28]	13.16 (3.435) [10.27]	0.615 (0.673) [1.647]	0.137 (0.061) [0.515]
1994/95	112.43 (20.97) [26.24]	6.057 (1.730) [4.043]	0.821 (0.443) [1.375]	0.204 (0.067) [0.907]
1995/96	120.07 (17.61) [24.46]	2.543 (1.196) [2.392]	0.185 (0.058) [0.267]	4.341 (1.042) [18.16]
1996/97	75.50 (9.572) [14.28]	9.295 (2.733) [6.839]	0.647 (0.471) [1.264]	9.855 (2.205) [42.06]
1997/98	132.26 (19.17) [26.38]†	12.84 (3.382) [8.601]	0.106 (0.047) [0.184]	0.068 (0.046) [0.296]
1998/99	141.64 (16.32) [25.26]	22.97 (4.019) [14.64]	3.403 (0.997) [6.334]	1.495 (0.571) [6.827]
1999/00	86.60 (20.02) [26.18]*			
2000/01	100.71 (16.60) [21.64]	4.809 (1.119) [3.041]	0.176 (0.100) [0.304]	1.344 (0.193) [5.497]
2001/02	105.01 (18.17) [23.30]	58.66 (7.127) [33.81]	0.075 (0.058) [0.150]	0.214 (0.097) [0.970]
2002/03	52.02 (10.43) [13.05]	14.49 (2.623) [9.725]	0.192 (0.174) [0.456]	0.473 (0.236) [2.186]
2003/04	98.67 (14.48) [20.06]	35.78 (6.696) [21.76]	0.276 (0.386) [0.860]	0.420 (0.223) [1.645]
2004/05	89.05 (12.35) [17.39]	25.36 (3.935) [14.96]	0.071 (0.030) [0.123]	0.375 (0.243) [1.570]
2005/06	62.71 (35.89) [37.63]†	15.79 (3.969) [12.04]	0.241 (0.063) [0.347]	1.725 (0.722) [6.149]
2006/07	79.18 (21.90) [24.90]	13.96 (3.393) [11.75]	0.119 (0.144) [0.341]	0.238 (0.098) [1.047]
2007/08	106.65 (29.10) [33.26]	21.88 (4.212) [13.47]	1.267 (1.343) [3.477]	0.277 (0.193) [1.4558]
2008/09	101.43 (33.20) [36.43]	9.665(1.974) [5.368]	0.756 (0.310) [1.174]	1.207 (0.536) [4.519]

* Based on only one leg of the survey.

† Standard error based on an estimate because only one station was sampled on a leg for a particular Hotspot.

TABLE 2. CV values under different scenarios for the number of survey legs conducted and different amounts of sampling on each leg. The square root of the additional variance (σ_{add}) and the general sampling standard error $(\bar{\sigma})$ used in the computation of these CVs are also shown.

	Square root of additional variance (σ_{add})	**General sampling standard error $(\bar{\sigma})$**	**Normal number of sampling stations and normal number of legs**	**Half the number of sampling stations and double the number of legs**	**Quarter the number of sampling stations and four times the number of legs**
Cape Point	0.187	0.237	0.219	0.200	0.193
Dassen Island	1.282	0.288	1.519	1.129	0.880
Dassen Island (outlier omitted)	0.693	0.288	0.615	0.476	0.395
Saldanha Bay	1.190	0. 638	1.612	1.441	1.582
Lambert's Bay	1.721	0.509	3.466	2.802	2.582