

Updated GLMM standardisation of the commercial abalone CPUE for Zone F

Anabela Brandão and Doug S. Butterworth

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

August 2016

Abstract

This paper presents an update of the standardisation of the abalone CPUE for Zone F using a GLMM approach proposed by Brandão and Butterworth (2012), which adds new data for the 2014/15 fishing “years” for that Zone. The standardised CPUE indices have been fairly stable since the 2007/08 fishing “year”, and show a slight increase in the latest year.

Introduction

A similar catch-per-unit-effort (CPUE) General Linear Mixed Model (GLMM) standardisation procedure described in Brandão and Butterworth (2012) for Zones A to D has been applied to the commercial abalone data for Zone F. This procedure which was first reported in Brandão and Butterworth (2012) and updated in Brandão and Butterworth (2015), is further updated in this paper to take into account further data that are now available for the 2015 Model-year (the 2014/15 fishing “year”), where a Model-year y runs from October of year $y-1$ to September of year y . The principle objective of these GLMM analyses is to obtain series of relative abundance indices that have been standardised by incorporating important covariates in the explanation of abalone CPUE variation.

The data

Commercial catch data (as kg whole mass), and effort data (as total duration of dives in minutes for each day dived) are available for the period covering Model-years 1981 to 2015. The covariates included in the GLMM analysis include the date (in terms of Model-year and season (3-monthly periods)), and the divers. Records with a dive time of less than 10 minutes were excluded, as were records for divers that had less than three dives in the whole database. A total of 1 928 data points remained for the analysis. Table 1 gives the number of records used in the GLMM analysis per Model-year.

General Linear Mixed Model (GLMM) to standardise the CPUE

The GLM (General Linear Model) used to standardise commercial CPUE indices assumes that all factors in the model are fixed effects with the variance of the response values being that of the error term ε . A GLM analysis models only the mean (i.e. the fixed effects) of the data. A GLMM has the ability to model not only the mean of the data but also its variance. In fact, a GLMM also allows for the presence of random variables (called random effects) which describe additional variability in the data apart from that reflected by the error term of equation (1) below. One of the covariates used in the GLM by Plagányi and Edwards (2007) is “divers” with 165 different levels (in the present analysis) associated with different divers, with some of the divers in the fishery having very few dives. The alternative approach proposed by Brandão and Butterworth (2012) is to treat “divers” as a random effect in a GLMM.

The GLMM applied to the abalone commercial CPUE data is of the form:

$$\ln(\text{CPUE}) = \mathbf{X}\alpha + \mathbf{Z}\beta + \varepsilon, \quad (1)$$

where :

CPUE is the catch-per-unit-effort defined as catch (kg) divided by dive time (minutes),

α is the unknown vector of fixed effects parameters which includes:

$$\mu + \alpha_{\text{year}} + \beta_{\text{season}} + \eta_{\text{year} \times \text{season}}, \text{ where}$$

μ is the intercept,

year is a factor with 31 levels associated with the Model-years 1981–2015 (excluding the years 1984, 1985, 1999 and 2009; in 2009 the fishery was closed),

season is a factor with 4 levels associated with the season effect (1 = Jan-Mar; 2 = Apr-Jun; 3 = Jul-Sep; 4 = Oct-Dec), and

$\text{year} \times \text{season}$ is the interaction between year and season,

\mathbf{X} is the design matrix for the fixed effects,

β is the unknown vector of random effects parameters (here diver which is a factor with 165 levels associated with the diver code, which includes both the entitlement holders coded in the database as well as "divers"),

\mathbf{Z} is the design matrix for the random effects,

ε is an error term assumed to be normally distributed and independent of the random effects.

This approach assumes that both the random effects and the error term have zero mean, i.e. $E(\beta) = E(\varepsilon) = 0$, so that $E(\ln(\text{CPUE})) = \mathbf{X}\alpha$. The variance-covariance matrix for the residual errors (ε) is denoted by \mathbf{R}

and that for the random effects (β) by \mathbf{G} . The analyses undertaken here assume that the residual errors as well as the random effects are homoscedastic and are uncorrelated, so that both \mathbf{R} and \mathbf{G} are diagonal matrices given by:

$$\mathbf{R} = \sigma_{\varepsilon}^2 \mathbf{I}$$

$$\mathbf{G} = \sigma_{\beta}^2 \mathbf{I}$$

where \mathbf{I} denotes an identity matrix. Thus, in the mixed model, the variance-covariance matrix (\mathbf{V}) for the response variable is given by:

$$\text{Cov}(\ln(\text{CPUE})) = \mathbf{V} = \mathbf{ZGZ}^T + \mathbf{R},$$

where \mathbf{Z}^T denotes the transpose of the matrix \mathbf{Z} .

The estimation of the variance components (\mathbf{R} and \mathbf{G}), the fixed effects (α) and the random effects (β) parameters in GLMM requires two steps. First the variance components are estimated by the method of residual maximum likelihood (REML), which produces unbiased estimates for the variance components as it takes into account the degrees of freedom used in estimating the fixed effects. Once estimates of \mathbf{R} and \mathbf{G} have been obtained, estimates for the fixed effects parameters (α) can be obtained as well as predictors for the random effects parameters (β).

For this model, because of interactions with year (which imply changing spatio-temporal distribution patterns), the standardised CPUE series is obtained from:

$$\text{CPUE}_{\text{year}} = \left[\sum_{\text{season}} \left(\exp(\mu + \alpha_{\text{year}} + \beta_{\text{season}} + \varphi_{\text{diver}} + \eta_{\text{year} \times \text{season}}) \right) \right] / 4 \quad (2)$$

where the standardisation is with respect to a diver code = 110, which contained the most observations.

The reason for standardising in this way when year interactions are present, is that the standardised CPUE is to be used as an index of relative abundance when input to assessment models. CPUE itself is assumed to be proportional to local density, so that averaging over season is necessary to provide a quantity representative of a consistently calculated average over each year. This averaging is unnecessary in the absence of such interactions, because then the $\exp(\alpha_{\text{year}})$ term alone would then be proportional to abundance.

Results and Discussion

Table 2 lists the nominal and the GLMM-standardised CPUE indices provided by the model and Figure 1 shows these graphically. There is not much difference between the GLMM standardised CPUE and the nominal series. Both show a slight increase for the latest year.

Reference

Brandão, A. and Butterworth, D.S. 2012. GLM and GLMM standardization of the commercial abalone CPUE for Zones A-D. FISHERIES/2012/AUG/SWG-AB/04.

Brandão, A. and Butterworth, D.S. 2012. GLM and GLMM standardization of the commercial abalone CPUE for Zone F. FISHERIES/2012/AUG/SWG-AB/09.

Brandão, A. and Butterworth, D.S. 2015. Updated GLMM standardization of the commercial abalone CPUE for Zone F. FISHERIES/2015/AUG/SWG-AB/07.

Plagányi, É. and Edwards, C. 2007. Summary of the GLM used to standardise abalone catch-per-unit-effort data for Zones A-D over the period 1980-2006. Marine and Coastal Management document: WG/AB/07/Aug/19.

Table 1. The number of data entries per Model year available for the final GLMM analysis to standardise the commercial abalone CPUE series for Zone F are shown. The abalone fishery was closed in February 2008 and reopened in 2010. Some sample sizes were considered too small and were not included in the analysis (see text). Model-years are defined as the period October of the preceding year to September of the year indicated.

Model year	Number of records
1981	24
1982	16
1983	31
1984	
1985	
1986	65
1987	48
1988	48
1989	63
1990	61
1991	42
1992	63
1993	23
1994	32
1995	13
1996	16
1997	29
1998	32
1999	
2000	33
2001	37
2002	53
2003	57
2004	24
2005	24
2006	22
2007	89
2008	249
2009	
2010	70
2011	125
2012	100
2013	87
2014	191
2015	161

Table 2. Nominal and GLMM-standardised commercial CPUE series for abalone for Model-years (October of the preceding year to September of the year indicated) 1981 to 2015 for Zone F. Both the nominal and the standardised values have been divided by the mean value of the respective series.

Model year	Nominal CPUE	GLMM standardised CPUE
1981	1.591	1.224
1982	1.591	1.140
1983	1.295	1.216
1984		
1985		
1986	1.404	1.250
1987	1.297	1.189
1988	1.599	1.825
1989	0.773	0.860
1990	1.530	1.611
1991	1.189	1.151
1992	0.827	0.901
1993	1.018	1.151
1994	1.960	2.159
1995	1.363	1.522
1996	1.578	1.724
1997	1.381	1.401
1998	0.938	1.005
1999		
2000	0.579	0.448
2001	0.767	0.669
2002	0.859	0.682
2003	0.723	0.573
2004	1.099	1.034
2005	1.092	1.139
2006	1.169	1.236
2007	0.388	0.347
2008	0.502	0.440
2009		
2010	0.397	0.424
2011	0.418	0.434
2012	0.410	0.417
2013	0.415	0.492
2014	0.391	0.337
2015	0.460	0.431

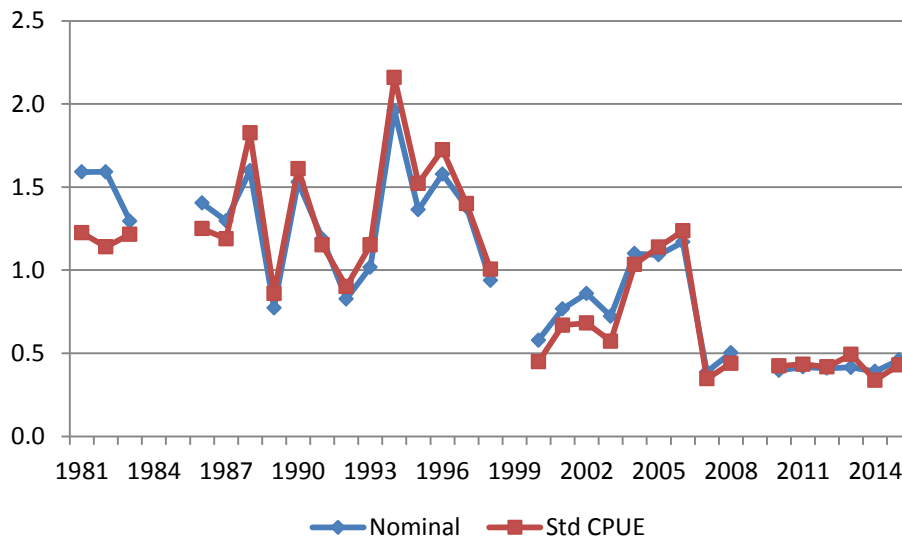


Figure 1. GLMM-standardised CPUE trends (normalised to their means over the 31 year period from 1981 to 2015, excluding the years 1984, 1985, 1999 and 2009 for which no or inadequate data are available (in 2009 the fishery was closed) for Zone F. For comparison, the nominal series (also normalised to its mean over the same period) is also shown.