GLMM standardisation of the commercial abalone CPUE for Zones A-D over the period 1980–2015

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 using a GLMM approach proposed by Brandão and Butterworth (2012), which adds new data for the 2014/2015 fishing "year" in Zones A and B. The standardised CPUE values for both Zones A and B for the 2015 model-year are slightly higher than the corresponding values for 2014; however these indices are still lower than those prior to 2011 and 2012 respectively.

.

Introduction

In this paper the GLMM described in Brandão and Butterworth (2012), and previously updated in Brandão and Butterworth (2015), has been applied to the commercial abalone data for Zones A-D to incorporate the further data now available for Zones A and B 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 the GLMM analysis 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 1980 to 2015. The covariates included in the GLMM analysis include the date (in terms of Model-year and season (3-monthly periods)), the divers, and the Zones that were dived. Zone C is split into subareas CNP (non-poached) and CP (poached). Records with a dive time less than 10 minutes were excluded as well as years which had too few records (less than eight) in a Zone/subarea, as were records for divers that had less than five dives in the whole database. A total of 44 176 data points remained for the analysis. Table 1 gives the number of records used in the final GLMM analysis per Model-year and per Zone/subarea.

General Linear Mixed Model (GLMM) to standardise the CPUE

The GLM used by Plagányi and Edwards (2007) and Brandão and Butterworth (2009) 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. In a GLM analysis only the mean (i.e. the fixed effects) of the data is modelled. A GLMM has the ability to model not only the mean of the data but also their variances. 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 ε. One of the covariates that was used in the GLM by Plagányi and Edwards (2007) is "divers" with 312 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:

$$\ln\left(CPUE)\right) = \mathbf{X}\alpha + \mathbf{Z}\beta + \varepsilon\ , \qquad (1)$$

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:

$\mu + \alpha_{year} + \beta_{season} + \gamma_{zone} + \eta_{year\times season} + \delta_{year\times zone}$, where

- μ is the intercept,
- *year* is a factor with 35 levels associated with the Model-years 1980–2015 (excluding 2009 during which 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),
- *zone* is a factor with 5 levels associated with the different zones/subareas (A, B, CNP, CP and D),
- *year×season* is the interaction between year and season, and
- *year×zone* is the interaction between year and zones/subareas, and

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 312 levels associated with the diver code, which includes both the entitlement holders coded in the database as well as "divers". Some divers not yet allocated a code were given a temporary code of 555 for the purposes of this analysis[1]),

[1] For the Model years 2006 to 2015 over which this code was used, such records comprise 0.04% of the total.

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(β) = E(ε) = 0, so that $\mathrm{E}\left(\ln\left(CPUE\right)\right) = X\alpha$. The variance-covariance matrix for the residual errors (ε) is denoted by **R** and that for the random effects (β) by **G**. The analyses undertaken here assume that the residual errors as well as the random effects are homoscedastic and are uncorrelated, so that both **R** and **G** are diagonal matrices given by:

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

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

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

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

The estimation of the variance components (**R** and **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 **R** and **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 for each zone/subarea is obtained from:

$$CPUE_{year,zone} = \left[\sum\nolimits_{season}\left(\exp\left(\mu + \alpha_{year} + \beta_{season} + \gamma_{zone} + \varphi_{diver} + \eta_{year\times season} + \delta_{year\times zone}\right)\right)\right]/4 \qquad (2)$$

where the standardisation is with respect to a diver code = 8, which contained the most observations as well as the longest period of operation in the fishery.

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 $\exp\left(\alpha_{year}\right)$ term alone is 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 graphical comparisons of the same. Broadly speaking, the standardisation makes relatively little difference to the nominal trends. Table 3 shows the parameter estimates, together with standard errors, obtained for the single fixed factors included in the GLMM model. The standardised CPUE values for both Zone A and Zone B for the 2015 model-year are slightly higher than the corresponding values for 2014; however these indices are still lower than those prior to 2011 and 2012 respectively.

Reference

Brandão, A. and Butterworth, D.S. 2009. A summary of the General Linear Model analyses applied to the commercial abalone CPUE data for Zones A-D over the period 1980-2008. MCM/2009/OCT/SWG-AB/06.

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

Brandão, A. and Butterworth, D.S. 2015. Updated GLMM standardisation of the commercial abalone CPUE for Zones A-D over the period 1980-2014. FISHERIES/2015/AUG/SWG-AB/02.

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 Zone available for the final GLMM analysis to standardise the commercial abalone CPUE series are shown. Subarea CNP was closed during the 2001 fishing season and subarea CP during all of the 2001, 2002 and 2003 fishing seasons. 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	**Zone/subarea**					
	A	**B**	**CNP**	**CP**	**D**	**Total**
1980	257	555	73	753	535	2173
1981	192	578	147	622	383	1922
1982	311	610	109	594	608	2232
1983	327	690	144	466	301	1928
1984	334	696	274	364	373	2041
1985	359	620	158	366	583	2086
1986	340	763	222	445	205	1975
1987	443	586	106	494	144	1773
1988	457	434	96	498	147	1632
1989	448	414	91	504	184	1641
1990	525	410	138	458	140	1671
1991	446	404	161	539	167	1717
1992	348	302	98	396	142	1286
1993	299	238	110	334	75	1056
1994	345	290	155	287	162	1239
1995	441	238	137	333	171	1320
1996	508	324	402	428	206	1868
1997	720	248	249	117	194	1528
1998	599	472	207	71	291	1640
1999	686	418	57	8	301	1470
2000	448	321	23		305	1097
2001	391	289			133	813
2002	288	226	99		95	708
2003	415	128	54		26	623
2004	97	574	158		69	898
2005	63	599	170		56	888
2006	41	673	164		50	928
2007		483				483
2008		291				291
2009						
2010	176	229				405
2011	368	384				752
2012	285	351				636
2013	333	320				653
2014	197	213				410
2015	203	192				395

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) 1980 to 2014 and Zones/subareas A, B, CNP, CP and D. Both the nominal and the standardised values have been divided by the mean value of the respective series.

a) Nominal CPUE series

Model year	Zone/subarea				
	A	B	CNP	CP	D
1980	1.142	0.835	0.875	0.841	0.908
1981	1.114	0.837	0.900	0.834	0.834
1982	0.978	0.847	0.884	0.834	0.804
1983	0.965	0.826	0.942	0.875	0.724
1984	1.047	0.882	0.965	0.891	0.797
1985	0.982	0.892	0.922	0.965	0.811
1986	1.088	0.974	1.026	1.104	0.771
1987	1.119	0.954	1.137	1.057	0.869
1988	1.201	1.048	1.189	1.149	1.035
1989	1.097	1.065	1.158	1.116	0.895
1990	1.260	1.299	1.422	1.215	1.247
1991	1.259	1.343	1.226	1.103	1.233
1992	1.388	1.413	1.268	1.234	1.165
1993	1.505	1.671	1.093	1.292	1.911
1994	1.429	1.451	1.223	1.324	1.711
1995	1.329	1.565	1.256	1.131	1.469
1996	1.315	1.479	0.979	0.902	1.441
1997	1.221	1.573	0.895	0.722	1.497
1998	1.263	1.412	0.977	0.737	1.545
1999	1.077	1.240	0.985	0.673	1.027
2000	1.123	1.252	1.096		0.946
2001	1.113	1.156			0.849
2002	1.120	1.188	1.244		0.762
2003	0.909	1.128	0.805		0.484
2004	0.873	0.824	0.559		0.445
2005	0.552	0.742	0.512		0.393
2006	0.586	0.628	0.460		0.428
2007		0.541			
2008		0.535			
2009					
2010	0.677	0.800			
2011	0.482	0.523			
2012	0.443	0.530			
2013	0.387	0.489			
2014	0.450	0.504			
2015	0.507	0.554			

b) GLMM-standardised CPUE series

Model year	Zone/subarea				
	A	B	CNP	CP	D
1980	1.151	0.893	0.952	0.916	0.949
1981	1.152	0.902	1.022	0.925	0.885
1982	1.002	0.928	1.026	0.915	0.867
1983	0.976	0.878	0.997	0.930	0.755
1984	1.042	0.913	0.987	0.940	0.829
1985	0.989	0.922	0.941	0.977	0.845
1986	1.017	1.001	1.050	1.123	0.906
1987	1.069	0.945	1.090	1.025	0.962
1988	1.089	1.017	1.117	1.082	1.080
1989	1.108	1.082	1.144	1.127	1.006
1990	1.168	1.197	1.281	1.137	1.229
1991	1.185	1.125	1.158	1.075	1.060
1992	1.334	1.365	1.292	1.235	1.215
1993	1.280	1.483	1.102	1.263	1.698
1994	1.191	1.118	1.198	1.205	1.269
1995	1.275	1.306	1.216	1.123	1.173
1996	1.305	1.433	1.056	1.000	1.341
1997	1.255	1.572	0.872	0.717	1.432
1998	1.249	1.439	0.956	0.698	1.467
1999	1.129	1.369	1.035	0.587	1.083
2000	1.121	1.223	1.063		0.957
2001	1.142	1.165			0.865
2002	1.206	1.159	1.031		0.792
2003	0.988	1.118	0.790		0.568
2004	0.981	0.907	0.608		0.668
2005	0.666	0.757	0.528		0.492
2006	0.739	0.724	0.488		0.607
2007		0.647			
2008		0.620			
2009					
2010	0.735	0.913			
2011	0.532	0.604			
2012	0.491	0.617			
2013	0.462	0.575			
2014	0.470	0.518			
2015	0.502	0.566			

Table 3. Parameters estimates and standard errors for the single fixed factors *Year*, *Season* and *Zone* included in the GLMM to obtain standardised indices of abundance for abalone.

Year	Parameter estimate	Year	Parameter estimate
1980	0.000	**2008**	-0.470 (0.035)
1981	0.034 (0.027)	**2009**	—
1982	0.042 (0.027)	**2010**	-0.058 (0.038)
1983	0.061 (0.028)	**2011**	-0.424 (0.034)
1984	0.074 (0.027)	**2012**	-0.362 (0.036)
1985	0.109 (0.029)	**2013**	-0.397 (0.039)
1986	0.099 (0.028)	**2014**	-0.433 (0.057)
1987	0.079 (0.030)	**2015**	-0.359 (0.042)
1988	0.175 (0.033)		
1989	0.172 (0.032)		
1990	0.409 (0.032)		
1991	0.316 (0.035)		
1992	0.491 (0.043)		
1993	0.535 (0.052)		
1994	0.348 (0.044)		
1995	0.319 (0.046)		
1996	0.557 (0.032)		
1997	0.548 (0.047)		
1998	0.600 (0.028)		
1999	0.467 (0.031)		
2000	0.438 (0.035)		
2001	0.401 (0.036)		
2002	0.366 (0.037)		
2003	0.308 (0.048)		
2004	0.313 (0.049)		
2005	0.058 (0.031)		
2006	-0.117 (0.031)		
2007	-0.208 (0.049)		
Season			
Jan-Mar	0.000		
Apr-Jun	-0.016 (0.022)		
Jul-Sep	0.115 (0.022)		
Oct-Nov	0.114 (0.061)		
Zone			
A	0.391 (0.031)		
B	0.000		
CNP	-0.034 (0.052)		
CP	-0.056 (0.024)		
D	0.155 (0.026)		

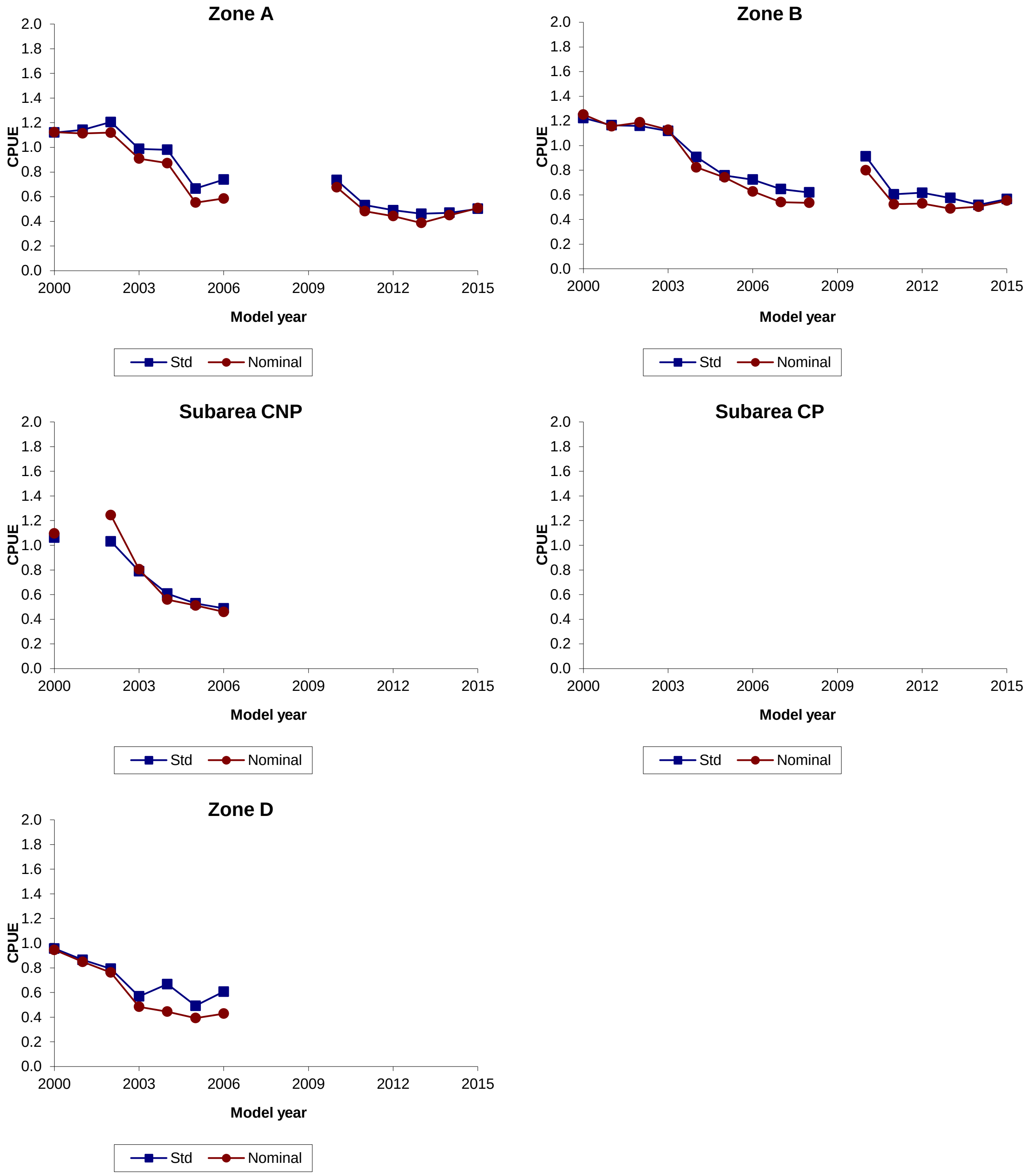

Figure 1. GLMM-standardised CPUE trends (normalised to their means over the 35 year period from 1980 to 2015, excluding 2009 during which the fishery was closed) for Zones/subareas A, B, CNP, CP and D. For comparison, the nominal series (also normalised to their means over that same period) are also shown. Plots commence from 2000 to provide greater clarity regarding more recent trends.