General linear mixed model analyses of male west coast rock lobster somatic growth for each of three super-areas where mark-recapture programmes take place

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

Summary

Random effects model estimates of annual standardised somatic growth for male rock lobsters by super-area (Areas 8–14, 7 and 3–6) are presented in this paper. The "less data" selection has been used for these analyses. Given scant indication of significant differences in trend with year between the three super-areas, it is recommended that spatially disaggregated assessments be based upon a common trend in somatic growth with year estimated from analyses of data for all three super-areas combined.

Introduction

Three super-areas where a mark-recapture programme for male rock lobsters takes place have been defined here as the "Cape" (Areas 8–14), "Dassen Island" (Area 7) and "West Coast" (Areas 3–6) areas. Table 1 shows the macro and sub-areas which fall within these super-areas. The "less data" selection used in the General Linear Mixed Model (GLMM) analyses of Brandão and Butterworth (2003) has been split into the three super-areas. This same GLMM has been used to describe seasonal and spatial variation in somatic growth of the west coast rock lobster on a super-area scale.

The data

The "less data" selection which has **6 920** records was split into three data sets representing the three super-areas. The "Cape" area contains **4 080** records, the "Dassen Island" area **616** and the "West Coast" area **2 224** records.

The General Linear Mixed Model

The approach proposed in this paper is to treat the interaction between the year and the location factors as a random effect in the GLMM.

The GLMM applied to the growth data is of the form:

$$\Delta L = \boldsymbol{X}\alpha + \boldsymbol{Z}\beta + \varepsilon\,, \qquad (1)$$

where : ΔL is the annual growth increment (mm),

α is the unknown vector of fixed effects parameters (this vector includes all the parameters for the intercept (μ), the year effect (η_{year}), the location effect ($\varphi_{location}$) and the coefficient (ρ) for the initial length variate (L)),

X is the design matrix for the fixed effects,

β is the unknown vector of random effects parameters (here the year-location interactions),

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.

It is assumed that both the random effects and the error term have zero mean, i.e. E(β) = E(ε) = 0, so that E(*Incr*) = **X**α. We denote the variance-covariance matrix for the residual errors (ε) by **R** and the variance-covariance matrix for the random effects (β) by **G**. In the analyses of this paper it is assumed 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:

$$\mathrm{Cov}(Incr) = \mathbf{V} = \mathbf{ZGZ}^{\mathrm{T}} + \mathbf{R}\,,$$

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. 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). Variance component estimates are obtained by the method of residual maximum likelihood (REML) which produces unbiased estimates for the variance components as it takes the degrees of freedom used in estimating the fixed effects into account.

The super-area “Dassen Island” contains only one location, so that there is no year-location interaction which means that there is no random effects term in the model for this super-area. Also, the location factor in the fixed effect part of the GLMM falls away. Therefore for the “Dassen Island” super-area the following GLM model is fitted:

$$\Delta L = \mu + \eta_{year} + \rho L + \varepsilon \quad (2)$$

where:

μ	is the intercept,
year	is a factor with levels associated with the seasons from 68/69 to 02/03,
L	is a variate representing the pre-moult carapace length of a male lobster (mm),
ρ	is a constant coefficient estimated when this GLM model is fitted to the data, and
ε	is an error term assumed to be normally distributed with constant variance.

Results

The parameter estimates and 95% confidence intervals for the year effect when a GLMM is fitted to the "Cape" super-area data are shown in Figure 1a. Figure 1b shows these estimates and confidence intervals when the GLM of equation (2) is fitted to the "Dassen Island" location data and Figure 1c a GLMM fitted to the "West Coast" super-area data. Note that the confidence intervals for Dassen Island are "unrealistically" small compared to those for the other two super-areas. This is because, with data from only one location, the contributions of random effects and residual random variation cannot be separated, so that the former is reflected in random variability over time in the year-effect estimates, with the associated confidence interval estimates accordingly negatively biased. The estimate for the 96/97 season for Dassen Island is much lower than estimates for other seasons. However this estimate (as well as the estimate for the 84/85 season) is based on only four records and therefore not much reliability should be accorded to it. Figure 2 shows these estimates for the three super-areas in one plot for comparison.

Table 2 reports the annual estimates of the mean moult increment of a 70mm male lobster from the CP2 location of super-area "Cape", from "Dassen Island" and from the EB location of the "West Coast" area. Table 3 shows the parameter estimates (together with their standard errors) of the variance components of the GLMM as well as the estimate of the slope coefficient (ρ) for the three super-areas.

Conclusions

The results in Table 3 do not suggest major differences in the variance structure or the slope parameter ρ between the super-areas. In principle the separate analyses presented for each super-area could be combined in a single estimation with a year-effects that differ by super-area, but common slope coefficient (ρ) and variance structure. However, the GenStat package used for this analysis does not allow for this option. Furthermore, the superposition of the year effects for the super-areas in Figure 2 does not suggest any difference in the trends over time for the super-areas. Indeed the further superposition in Figure 3 of the estimates plus confidence intervals for the "less data" set overall, without disaggregation by super-area, suggests that any differences in trend are not statistically significant (note that in this context the point estimates for Dassen Island need to be interpreted

allowing for the fact that they incorporate year-location random-effects and that the estimates for the 84/85 and 96/97 seasons are both based on only four records).

Accordingly it is recommended that super-area disaggregated assessments be based upon a common somatic growth rate trend time series obtained from analysis of the “less data” set as a whole. The choice of location (or some “averaged” location) within each super-area to which to normalise such growth, will however require further discussion.

References

Brandão, A. and Butterworth, D.S. 2003. Standardised male west coast rock lobster somatic growth trend using a mixed linear (“random effects”) model including data from the 2002/03 season. Marine and Coastal Management document: WG/07/03/WCRL19.

Table 1. Super-areas, macro-areas and locations which contribute to the “less data” set of male rock lobster growth information used in the GLMM analyses.

Recapture Location	Macro-area	Super-area
CP1	Cape Peninsula	Cape
CP2		
CP3		
CP4		
CP5		
RI	Robben Island	
HB	Knol	
WB2	Walker Bay	
DI1	Dassen Island	Dassen Island
EB	Elands Bay	West Coast
LB1	Lamberts Bay	
LB2		
LB3		
SB1	Saldanha Bay	
SB2		
ST1	St. Helena Bay	
ST2		
ST3		
ST4		

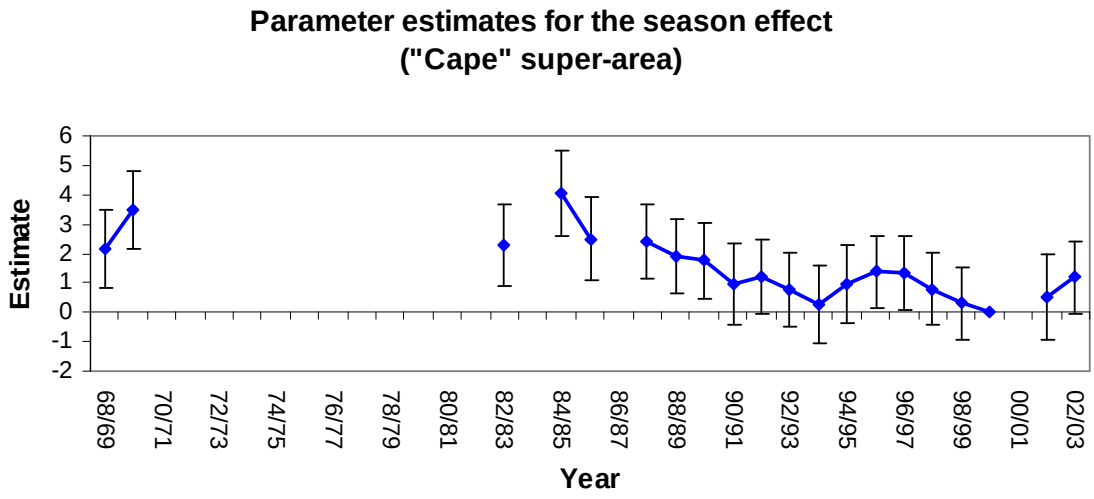
Table 2. Estimated mean moult increment (in mm) of a 70 mm male lobster when a GLMM (equation (1)) is fitted to the three super-areas for “less data” selection. Results are shown for the CP2 location of the “Cape” super-area and the EB location of the “West Coast” super-area.

Year	“Cape” super-area	“Dassen Island” location	“West Coast” super-area
1968	3.76		
1969	5.10		
1970			
1971		8.11	
1972			
1973			8.14
1974			6.81
1975			5.87
1976			5.42
1977			5.27
1978			6.38
1979			5.93
1980		4.58	
1981			6.28
1982	3.90		
1983			4.33
1984	5.67	3.80	
1985	4.10		
1986			
1987	4.01	5.67	5.56
1988	3.52	6.22	4.66
1989	3.36	4.92	3.49
1990	2.59		2.93
1991	2.81		4.34
1992	2.37		3.55
1993	1.85	5.24	3.75
1994	2.54	4.04	3.40
1995	3.00	5.76	3.48
1996	2.95	-0.39	4.95
1997	2.39	4.37	3.27
1998	1.91	3.62	3.28
1999	1.61	3.61	3.32
2000		4.77	3.24
2001	2.11	3.78	4.11
2002	2.79	4.34	

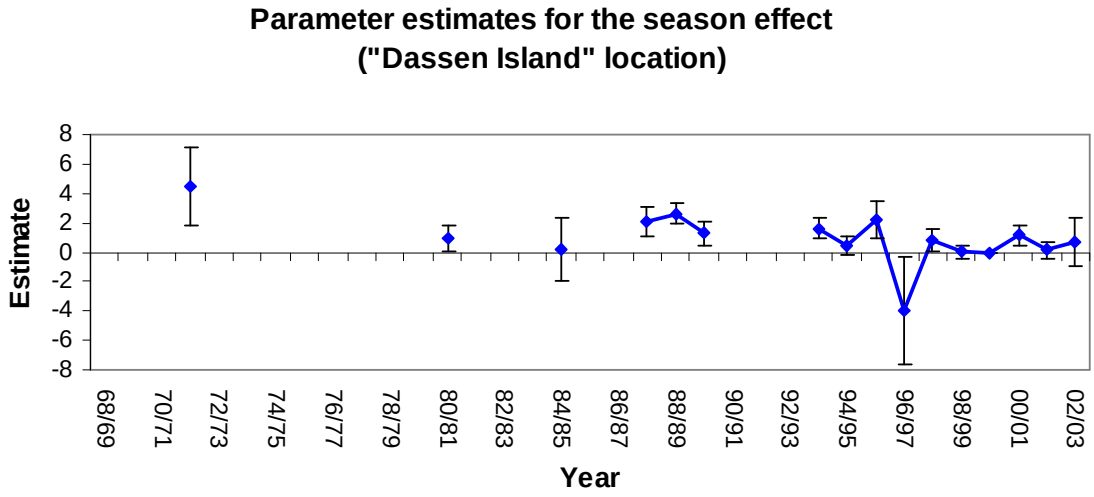
Table 3. Estimates (standard errors) for the variance components of the GLMM as well as the estimate (standard error) for the slope coefficient (ρ).

Parameter	**"Cape" super-area**	**"Dassen Island" location**	**"West Coast" super-area**
Random effects	0.248 (0.101)	—	0.323 (0.147)
Residuals (ε)	2.526 (0.056)	3.443 (0.199)	3.286 (0.100)
ρ	-0.083 (0.004)	-0.087 (0.016)	-0.083 (0.008)

a)



b)



c)

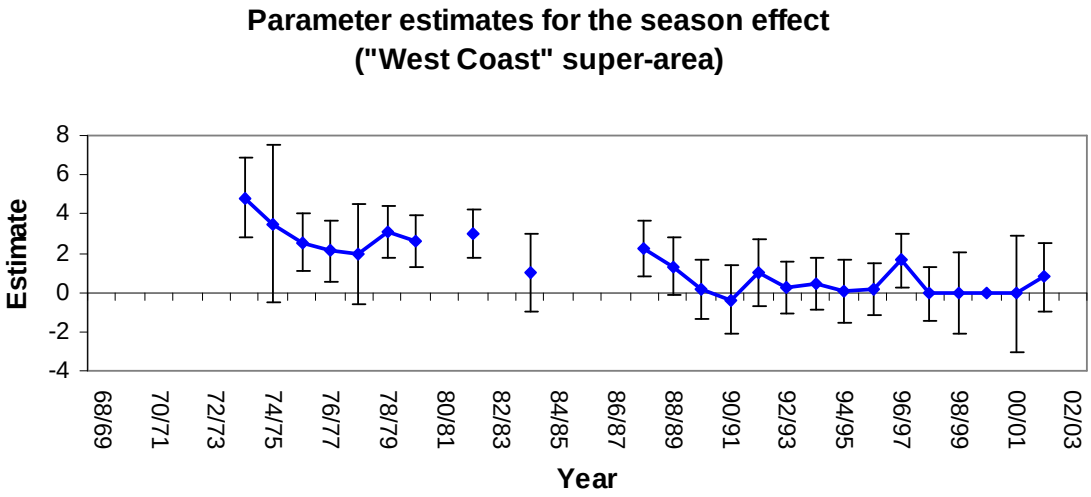


Figure 1. Parameter estimates (with 95% confidence intervals) for the year effect relative to the 1999/00 estimate when a GLMM is fitted to each individual data set for super-areas a) “Cape” b) “Dassen Island” and c) “West Coast”. For b) the GLMM simplifies to a GLM as data are available for one location only.

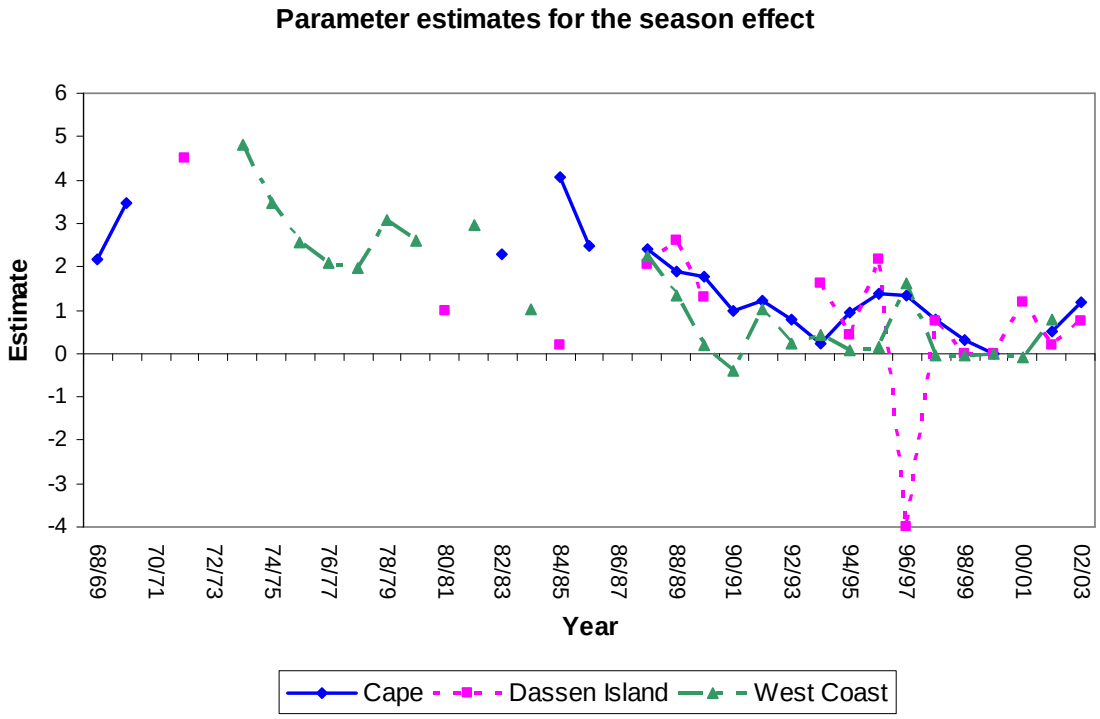


Figure 2. Comparative parameter estimates for the year effect relative to the 1999/00 estimate obtained for each of the three individual super-areas “Cape”, “Dassen Island” and “West Coast”.

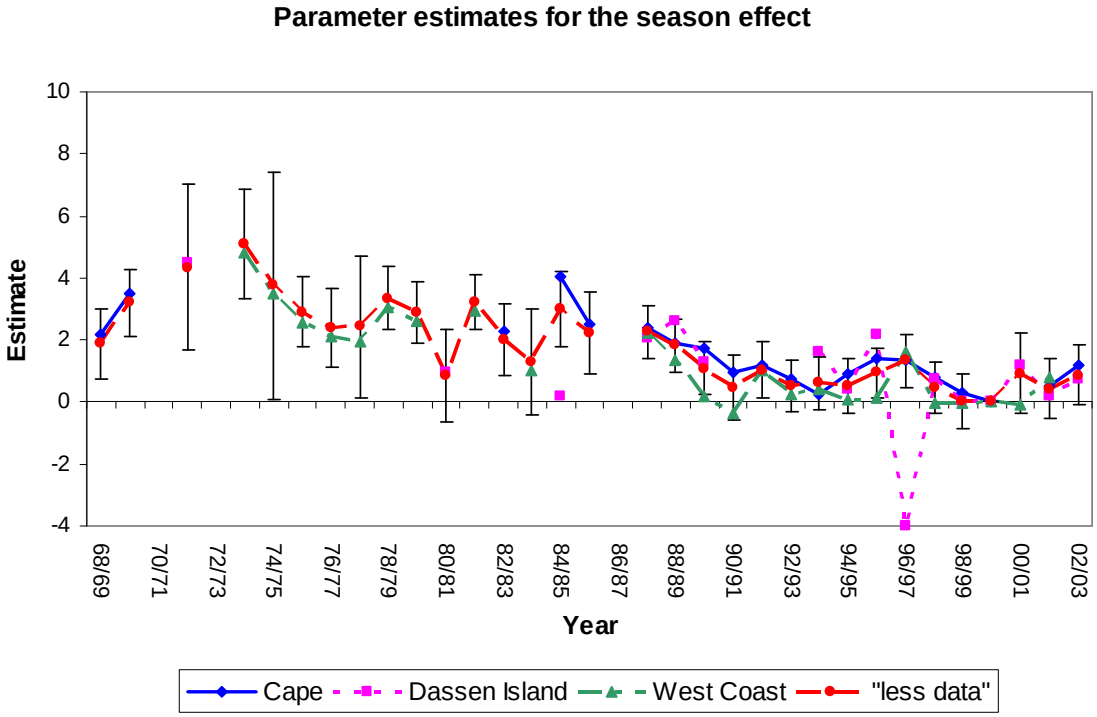


Figure 3. Repetition of the information in Figure 2 together with the results of the GLMM (including 95% confidence intervals) applied to the whole “less data” set without distinguishing super-areas.