GLMM- AND GLM-STANDARDISED LOBSTER CPUE FROM THE TRISTAN DA CUNHA GROUP OF ISLANDS FOR THE 1997-2007 PERIOD

S.J. Johnston, A. Brandão, and D.S. Butterworth.

MARAM
Department of Mathematics and Applied Mathematics
University of Cape Town
Rondebosch, 7701

June 2009

ABSTRACT

The longline CPUE series for the three outer islands are GLMM standardised. Year, month, area, trap-type, soak time, depth and year-area interactions are treated as fixed effects, and year-month interactions treated as a random effect. For Tristan, for which the available powerboat data are more limited, a GLM with year and month as fixed effects is applied. After initial increase, the standardised CPUE indices show drops over the most recent years for all islands except Gough, for which there is a steep recent increase.

INTRODUCTION

The commercial CPUE series of a resource is often used as an index of population density and consequently to inform on population abundance when modelling the dynamics of the underlying population. It is known, however, that a number of other factors besides density may influence the recorded values of CPUE. Where sufficient data exist, General Linear Mixed Model (GLMM) standardisation is able to take some of these further effects into account, thereby producing a more reliable index of abundance. This document reports the application of a GLMM standardisation to *Jasus tristiani* lobster catch and effort data from around Inaccessible, Nightingale and Gough Islands for the period 1997-2007. For Tristan, for which the data are more limited, a simpler GLM approach is used.

For the outer islands, only longline CPUE data are considered (i.e. the powerboat data are ignored for reasons given below). For Tristan, where normally all fishing occurs using powerboats, the CPUE series relates to powerboats.

METHODOLOGY

Data

Raw Logsheet data
The logsheet data for all islands have been entered electronically into EXCEL spreadsheets. Logsheet data from the fishery are available for the Season-Years

between 1996 and 2007, where a Season-Year is taken to run from September until August the following year, i.e. Season-Year 2005 refers to the period from September 2005 to August 2006. Unfortunately logsheet data for 2006 have been misplaced (James Glass, pers. comm.). Logsheet data are also incomplete for Season-Year 1996 (Edwards and Glass, 2007) for the three outer islands (Gough, Nightingale and Inaccessible), and thus 1996 is also omitted from these analyses.

Summary sheet data
Data summary sheets recorded by the Agriculture and Natural Resources Department on Tristan da Cunha are available from Season-Years 1996 to 2007. These contain summary data from both the logsheets (total catch and total effort) and factory reports (Edwards, 2007).

Accounting for inaccurate records for the three outer islands

Although logsheet data are valuable as they record details of the catches, e.g. location and soak-time which are needed for standardisation, the logsheet entries are known to be inaccurate (Edwards, 2007). In particular, longline catch and powerboat effort are unreliable. Furthermore there is currently insufficient information concerning the different catch rates for longline monster and powerboat traps, thereby precluding the standardisation of the catch rate across different types of fishing. All powerboat data were therefore excluded from the analyses presented here for Inaccessible, Nightingale and Gough.

Because of inaccurate longline catch records, the total logsheet catch for each Season-Year differ from the actual catch taken. A more accurate (best) estimate of the total longline catch in Season-Year *y* (C_y) is provided by subtracting the total powerboat catch from the total packed weight (both recorded on the Summary sheets), where the packed weight is scaled upwards to account for weight lost during processing (Edwards, 2007). This catch estimate can then be used to adjust the longline catch records so that the total catches from both sources are equal. Unfortunately there are logsheets missing for some years. An adjustment coefficient k_y was therefore developed using the ratio of total recorded effort for the Summary sheets and logsheets, to scale adjustments.

Adjusted logsheet catches were calculated as follows:

$$c_{i,y} \rightarrow c^{*}_{i,y} = c_{i,y} k_y = c_{i,y} \frac{C_y}{C_y^{LS} \frac{E_y^{SS}}{E_y^{LS}}} \qquad (1)$$

where

$c_{i,y}$ is the *i*'th logsheet longline catch record for Season-Year *y*,

C_y^{LS} is the total logsheet longline catch for Season-Year *y*,

C_y is the best estimate of the total longline catch for Season-Year *y* (based on summary sheets),

E_y^{LS} is the total logsheet longline effort for Season-Year *y*, and

E_y^{SS} is the total Summary sheet longline effort for Season-Year *y*.

Adjusted catches were then used to calculate Adjusted CPUE values (I_y^*) for each Season-year:

$$I_y = \frac{1}{n_y}\sum_i \frac{c_{i,y}}{e_{i,y}} \rightarrow \frac{1}{n_y}\sum_i \frac{c^*_{i,y}}{e_{i,y}} = I_y^* \qquad (2)$$

where

I_y is the nominal CPUE for Season-Year *y*,

$e_{i,y}$ is the *i*'th logsheet longline effort record for Season-Year *y*, and

n_y is the number of logsheet records for Season-Year *y*.

The General Linear Mixed Model for the three outer islands

A GLMM which includes both fixed and random effects is used to standardise the lobster CPUE data, where catches are the adjusted logsheet catches of Equation (1) and effort is logsheet effort. (Note that this approach assumes that the logsheet data represent an unbiased sample of all the fishery in each Season-Year.). This model allows for possible annual differences in the areal distribution of the lobsters (which is considered to be a fixed effect) and for annual differences in each month (considered as a random effect). This model is given by:

$$\ln(CPUE + \delta) = \mathbf{X}\alpha + \mathbf{Z}\beta + \varepsilon \qquad (3)$$

where:

α is the unknown vector of fixed effects parameters (in this case this consists of the factors given by equation (4) below),

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

β is the unknown vector of random effects parameters (which in this application consists of a year-month interaction),

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

δ is a small constant added to the rock lobster CPUE to allow for the occurrence of zero CPUE values (0.1 kg/trap in this case, being about 10% of the average nominal values), and

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 $E(\ln(CPUE+\delta)) = \mathbf{X}\alpha$. 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 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 sum of the factors that are considered as fixed effects (i.e. $\mathbf{X}\alpha$ in equation (1)) in the GLMM is given by the following:

$$\ln(CPUE + \delta) = \mu + \alpha_{year} + \beta_{month} + \gamma_{area} + \eta_{trap-type} + \lambda_{soaktime} + \theta_{depth} + \tau_{year\ x\ area} \quad (4)$$

where:

μ	is the intercept,
year	is a factor with 10 levels associated with the years (i.e. the Season-Years: 1997-2007, omitting 2006),
month	is a factor with levels associated with the fishing month (1-12 for Gough, 1-3 and 9-12 for Nightingale, 1-3 and 8-12 for Inaccessible),
area	is a factor with levels associated with groupings of fishing areas (Gough = 5 areas, Nightingale = 6 areas, Inaccessible = 9 areas),
trap type	is a factor with levels associated with the trap type (monster and plastic pots for Inaccessible, and Monster only for Gough and Nightingale),
soak time	is a factor with 3 levels associated with the soak time period ("1"=0.0–0.49 days, "2"= 0.5–1.9 days and "3" for 2 or more days),
depth	is a factor with 4 levels associated with fishing depth ranges ("1" for depths < 10m, "2" for 10–39.9m, "3" for 40–89.9m, and "4" for depths ≥ 90 m),
year x area	is the interaction between year and area.

In this application the CPUE has been standardised on the year 1997, month of *September*, trap type *Monster*, soak time "1" , and depth category "1".

For this model, because of the fixed effect interaction of area with year (which implies changing spatio-temporal distribution patterns), an index of overall abundance needs to integrate the different trends in density in each area over the size of these areas. Accordingly the standardised CPUE series is obtained from:

$$CPUE_{year} = \left[\sum_{area}\left(\left(\exp\left(\mu + \alpha_{year} + \gamma_{area} + \tau_{year\ x\ area}\right) - \delta\right) * A_{area}\right)\right] / A_{total} \quad (5)$$

where:

A_{area} is the surface size of the area concerned,

A_{total} is the total size of the fishing ground considered (the division by A_{total} is to keep the units and size of the standardised CPUE index comparable with those of the nominal CPUE).

Table 1 provides the A_{area} values for Inaccessible, Nightingale and Gough Islands.

Simple GLM (for Tristan data)

The model used here is given by:

$$\ln(CPUE + \delta) = \mu + \alpha_{year} + \beta_{month} \quad (5)$$

where:

C	is the catch in kg,
E	is the effort in powerboat days,
μ	is the intercept,
year	is a factor with 11 levels associated with the years (i.e. the Season-Years: 1997-2007),
month	is a factor with levels associated with the fishing month (1-12), and
δ	is taken to be 20 kg/powerboat-day (about 10% of the nominal average values).

For Tristan Island the CPUE has been standardised on the month of *September*. Further, as no *area*year* interactions are included, the standardised CPUE series is obtained from:

$$CPUE_{year} = \exp\left(\mu + \alpha_{year} + \beta_{September}\right) - \delta \quad (6)$$

RESULTS

Table 1 provides standardised CPUE values derived from the GLMM/GLM considered. For comparison, the adjusted nominal CPUE values are also reported. Figure 1 compares the adjusted nominal CPUE with the standardised CPUE series – both series have been renomalised for comparative purposes. Figure 2 shows the month effects for each island, and Figure 3 shows the area effects for each of Inaccessible, Nightingale and Gough Islands (no area data are reported with the Tristan CPUE datasheets).

DISCUSSION

From the analyses of this paper, the GLMM/GLM standardised CPUE series shown in Table 1 are put forward as the best upon which to base assessment of the resource.

The CPUE series for Tristan for the 2003+ period is not comparable with the pre-2003 period because at the start of the 2003 season changes were made to the method of fishing which included effectively reducing the length of the day spent fishing by a powerboat each day. This was done to reduce catches delivered to the factory on Tristan to a manageable level, and so improve the delivery of strong live lobster for whole frozen product.

Note that care should be taken in interpreting the post 2002 increase in standardised CPUE at Nightingale Island as entirely an abundance-related effect. Before that time with two vessels fishing, catching was near continuous. Subsequently only one vessel

fished for series of short periods. This allowed the lobster to redistribute into the limited fishable areas, thus inflating catch rates.

These results were discussed with one of the skippers working in the fishery (Clarence) who commented that the resulting standardised CPUE trends, as well as the trends in CPUE by month and area shown in Figure 2 and 3, were consistent with his experiences on the grounds.

FUTURE WORK

Time constraints have precluded the further analyses of these data that would be desirable in a fuller investigation. Factors which will be investigated further in the future include the choice of distributions other than the log-normal and the choice of the value for δ if the log-normal is used, and attempting to take explicit account of the post 2002 fishing strategy change at Nightingale Island. Future work will also include examining stratifying the existing area by depth for a better representation of density patterns prior to integrating over areas. It would also be useful if a record of the specific location fished could be kept. This could be done by recording, say, the shooting point for each line. Using these positions, one would be able to see more clearly the pattern of fishing in each area, and hence refine the extent of the area considered lobster habitat for use for A_{area} in equation 5.

ACKNOWLEDGEMENTS

Prior contributions of Charlie Edwards to work underlying these analyses is gratefully acknowledged.

REFERENCES

Edwards, C.T.T. 2007. Sources of data from the lobster fisheries on Inaccessible, Nightingale, Gough and Tristan da Cunha. Technical Report MARAM/Tristan/07/Dec/05, Ovenstone Fisheries.

Edwards, C.T.T. and Glass, J.P. 2007. Reconciliation of data from the lobster fisheries on Inaccessible, Nightingale, Gough and Tristan da Cunha. Technical Report MARAM/Tristan/07/Dec/06, Ovenstone Fisheries.

Table 1a: The size (km^2) of each fishing area at **Inaccessible** Island.

Area	Name	Size
1	Bank	53.58
2	North point	5.88
3	Salt beach	1.10
4	East Point	10.14
5	Toms beach and Black spot	3.60
6	South Hill	3.60
7	Pyramid rock and Blinder	5.23
8	West point	5.04
9	Blendon Hall	4.32

Table 1b: The size (km^2) of each fishing area at **Nightingale** Island.

Area	Name	Size
1	North	12.13
2	North East	3.29
3	South East	3.02
4	South	9.00
5	West	5.87

Table 1c: The size (km^2) of each fishing area at **Gough** Island.

Area	Name	Size
1	Cave Cove	6.48
2	Hawkins Bay	8.53
3	SE pt	8.01
4	SW pt	9.11
5	Gaggins pt	10.38
6	N pt	3.69

Table 2a: Standardised longline CPUE series for **Inaccessible** Island using the GLMM model detailed in the text. The number of data records for each Season-Year (*N*) is provided, along with the adjusted nominal CPUE series for comparison.

Season-Year	*N*	**Adjusted Nominal CPUE (kg/trap)**	**Standardised CPUE**
1997	259	2.583	0.294
1998	381	2.896	0.649
1999	371	2.846	0.661
2000	652	2.902	0.781
2001	575	5.219	0.682
2002	427	5.688	1.210
2003	245	5.902	0.624
2004	425	5.947	2.187
2005	467	5.667	1.703
2007	709	4.747	1.263

Table 2b: Standardised longline CPUE series for **Nightingale** Island using the GLMM model detailed in the text. The number of data records for each Season-Year (*N*) is provided, along with the adjusted nominal CPUE series for comparison.

Season-Year	*N*	**Adjusted Nominal CPUE (kg/trap)**	**Standardised CPUE**
1997	772	1.476	1.224
1998	482	3.130	2.223
2000	201	4.052	2.580
2001	584	3.098	2.695
2002	500	3.249	2.807
2003	232	6.115	4.836
2004	512	5.942	5.084
2005	414	7.238	4.640
2007	320	5.659	3.995

Table 2c: Standardised longline CPUE series for **Gough** Island using the GLMM model detailed in the text. The number of data records for each Season-Year (*N*) is provided, along with the adjusted nominal CPUE series for comparison.

Season-Year	*N*	**Adjusted Nominal CPUE (kg/trap)**	**Standardised CPUE**
1997	1204	2.388	1.057
1998	1220	1.947	0.975
1999	2012	1.848	1.109
2000	2110	1.484	0.741
2001	1579	1.251	0.808
2002	1905	1.350	0.713
2003	1684	1.383	0.831
2004	1071	1.616	0.747
2005	753	2.714	1.420
2007	401	5.841	2.990

Table 2d: Standardised powerboat CPUE series for **Tristan** Island using the GLM model detailed in the text. The number of data records for each Season-Year (*N*) is provided, along with the adjusted nominal CPUE series for comparison. Note the series from 2003 is not comparable with the pre-2003 series as a result of operational changes.

Season-Year	*N*	**Adjusted Nominal CPUE (kg/powerboat-day)**	**Standardised CPUE**
1997	311	265	314
1998	446	576	469
1999	337	806	627
2000	323	858	791
2001	333	850	795
2002	334	881	853
2003	433	373	325
2004	367	414	452
2005	312	480	453
2006	292	642	541
2007	318	553	415

Figure 1a: Comparative plot of the adjusted nominal and GLMM standardised longline CPUE series for **Inaccessible** Island. Both series have been renormalised to a mean of 1 for easier comparison of trends.

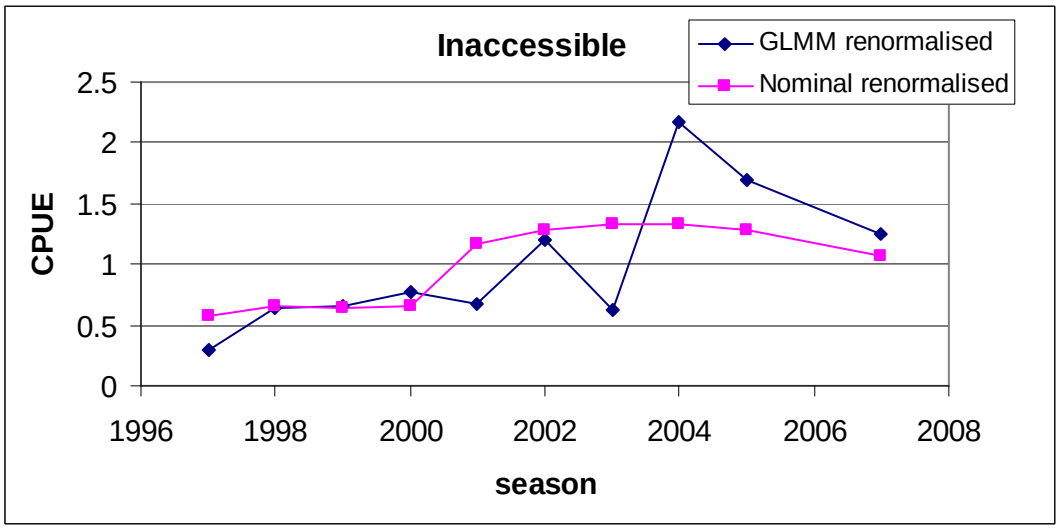

Figure 1b: Comparative plot of the adjusted nominal and GLMM standardised longline CPUE series for **Nightingale** Island. Both series have been renormalised to a mean of 1 for easier comparison of trends.

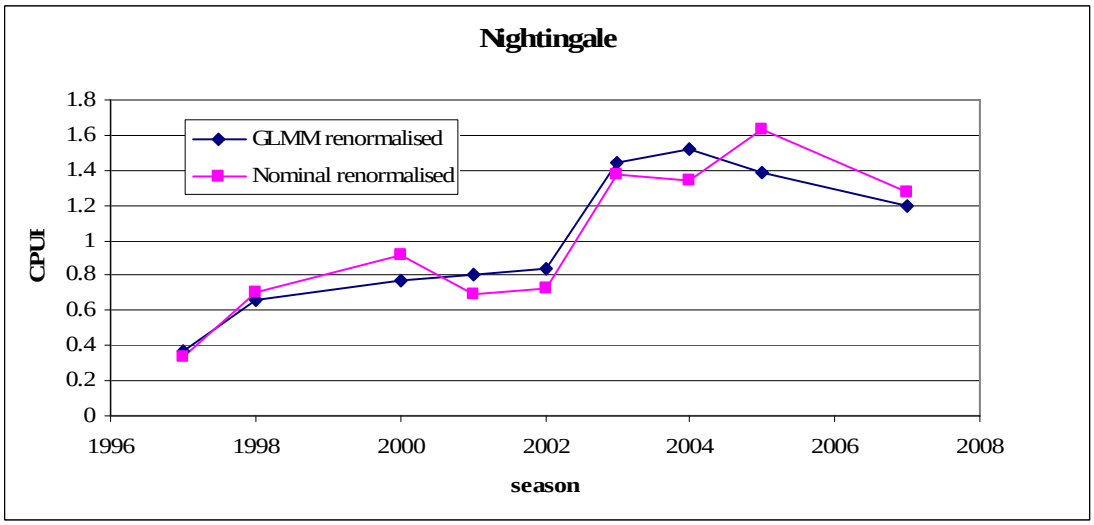

Figure 1c: Comparative plot of the adjusted nominal and GLMM standardised longline CPUE series for **Gough** Island. Both series have been renormalised to a mean of 1 for easier comparison of trends.

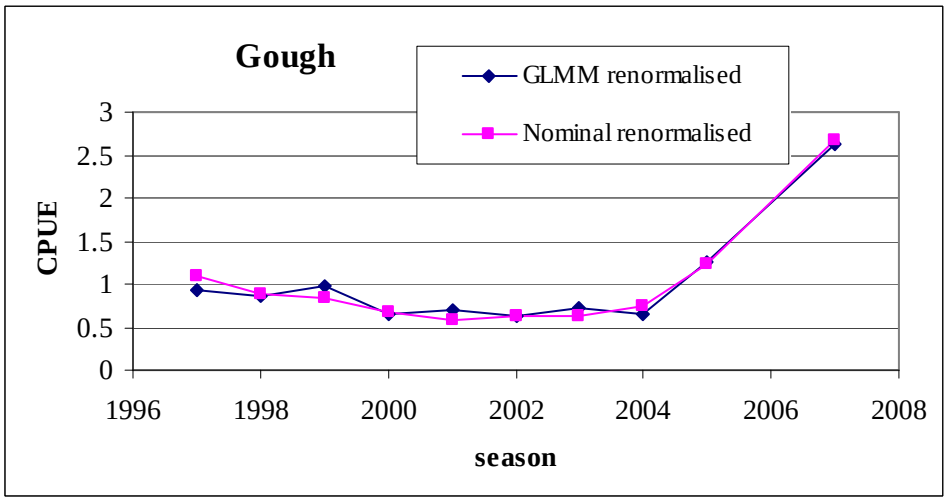

Figure 1d: Comparative plot of the adjusted nominal and GLM standardised powerboat CPUE series for **Tristan** Island. Both series have been renormalised to a mean of 1 for easier comparison of trends. Note the series from 2003 is not comparable with the pre-2003 series as a result of operational changes.

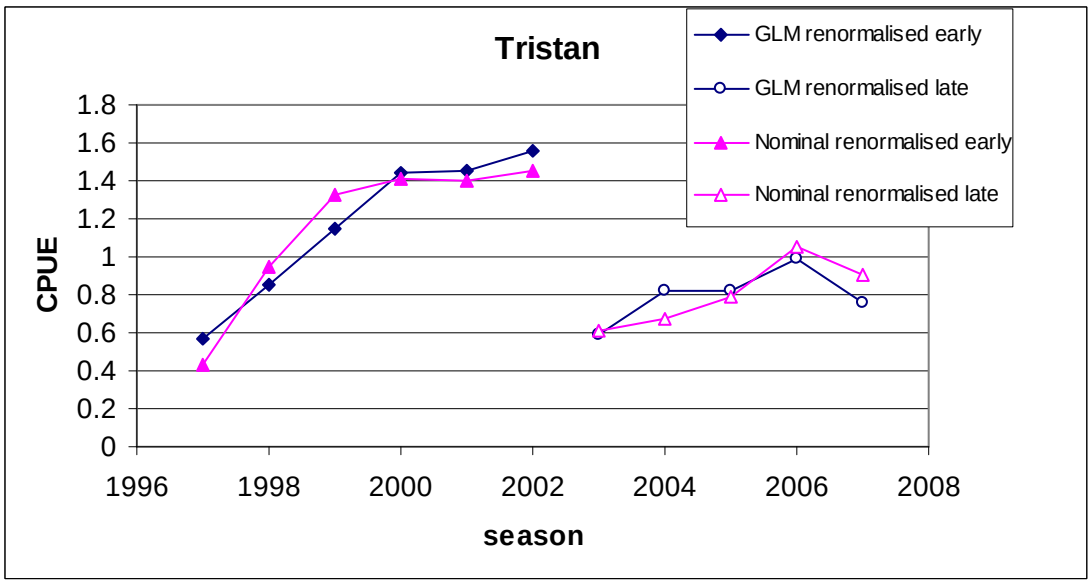

Figure 2a: GLMM month effects for the **Inaccessible** Island.

Figure 2b: GLMM month effects for the **Nightingale** Island.

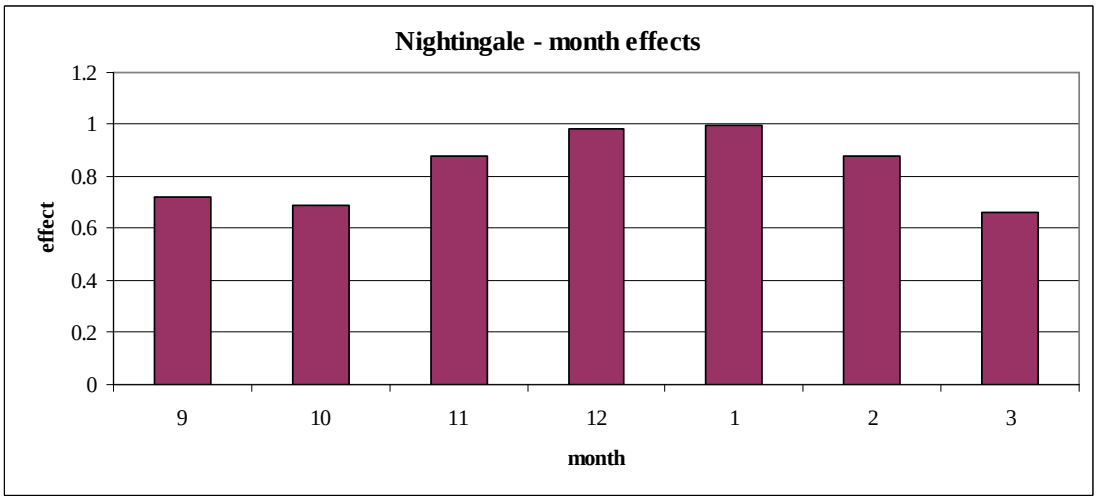

Figure 2c: GLMM month effects for the **Gough** Island.

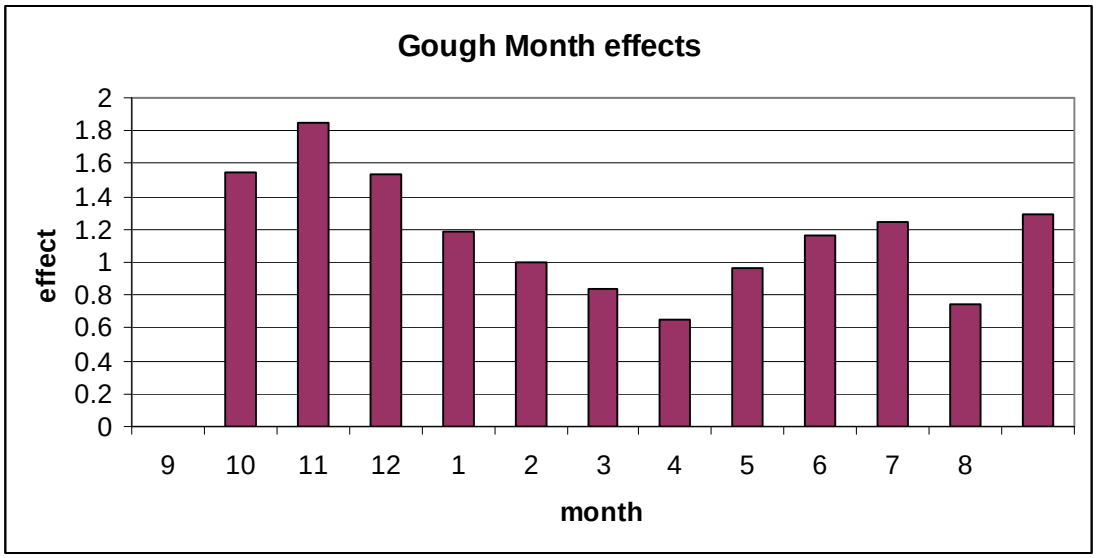

Figure 2d: GLM month effects for the **Tristan** Island.

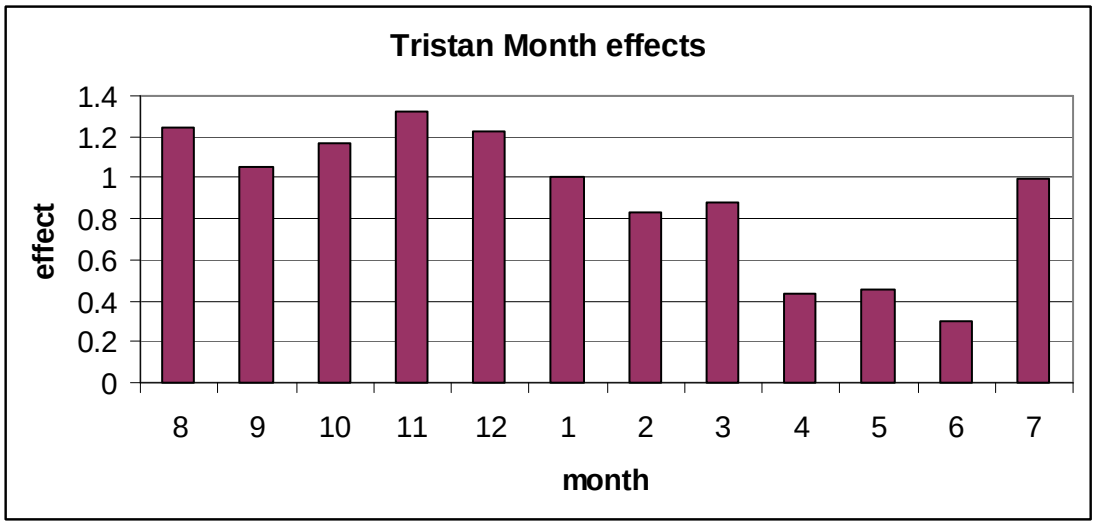

Figure 3a: GLMM area effects for **Inaccessible** Island.

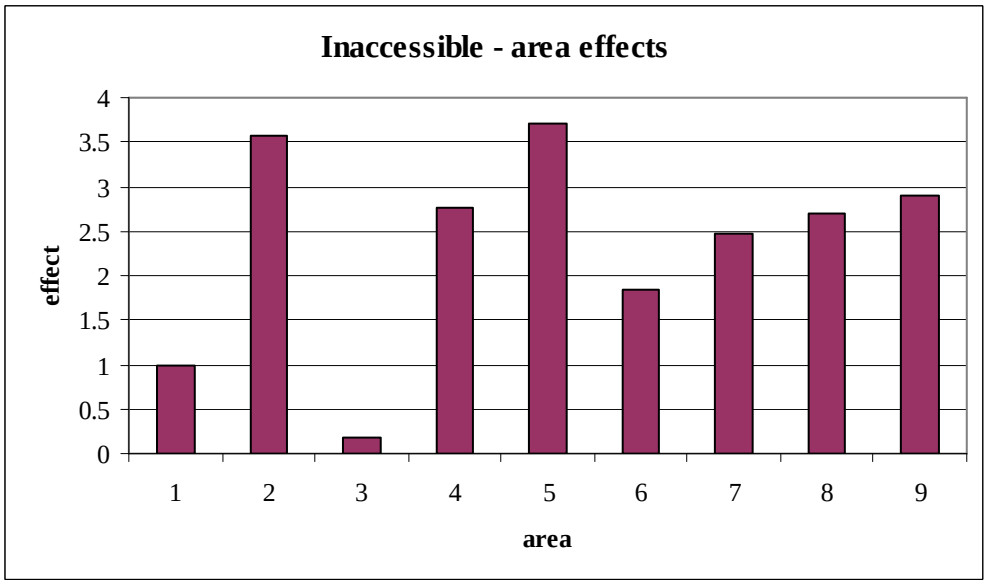

Figure 3b: GLMM area effects for **Nightingale** Island.

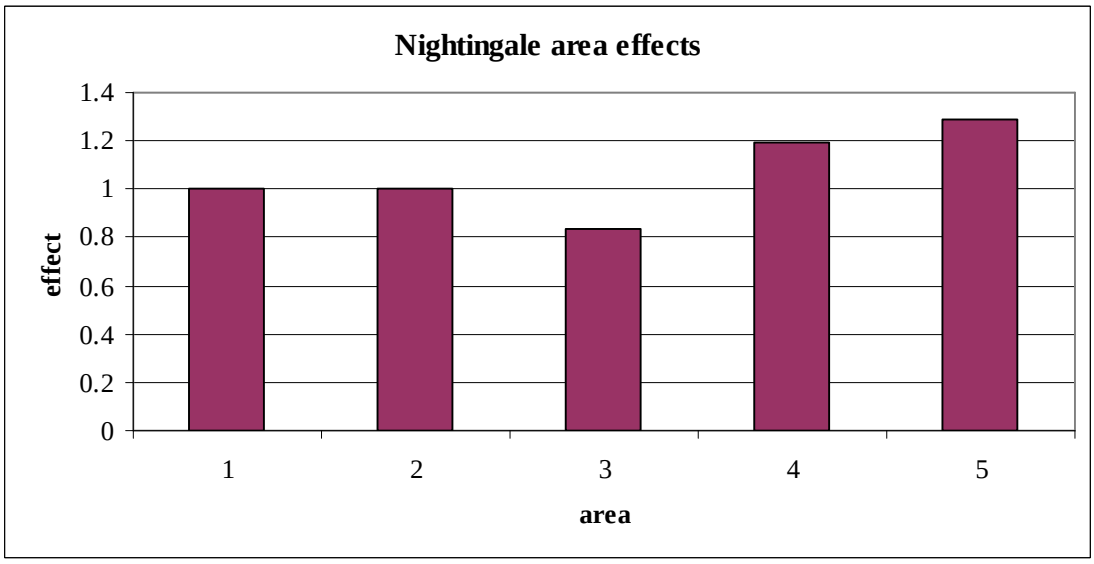

Figure 3c: GLMM area effects for **Gough** Island.