Delta-lognormal linear models applied to standardised CPUE abundance indices (1994 to 2003) for orange roughy off Namibia

Anabela Brandão and Doug S. Butterworth

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

March 2004

Abstract

GLM analyses are used to standardise the CPUE data for Namibian orange roughy in a manner that deals with tows that record zero catch of orange roughy. The possibility of there being a “learning” period of lower CPUE for a new vessel when it enters the fishery is taken into account. Further, to allow for areal expansion of the fishery at each aggregation, sub-aggregations are defined and CPUE trends estimated separately for each. Different methods for combining the results for the various sub-aggregations to provide a single index for an aggregation are considered. The standardised CPUE values for 2003 are lower to those for the previous year in all aggregations.

Introduction

In Brandão and Butterworth (2002) commercial CPUE data for orange roughy off Namibia were standardised by applying two Generalised Linear Models (GLMs) that addressed two problems encountered in such analyses for this fishery: i) a considerable number of tows with zero catches and ii) the areal distribution of effort shifting within and even beyond previously defined aggregations (especially notable for the *Johnies* aggregation). These standardised CPUE indices of abundance are then used as an input to a population model to assess the state of the stock (Brandão and Butterworth 2004). In Brandão and Butterworth (2003) these models were updated using the re-entered data for commercial fishing with previously missing records included and with the extra data for 2002. Brandão and Butterworth (2003) found that in all aggregations, the indices obtained from fitting a lognormal model hardly differed from those obtained from fitting a delta-lognormal model assuming binomial errors for the proportion positive. It was thus decided at the Deep Water Fisheries Working Group meeting held in Swakopmund on March 2003, to adopt the delta-lognormal models to standardise the commercial orange roughy CPUE data. In this paper, we present the results of the updated standardised CPUE indices for orange roughy taking an extra year's data into account, using a delta-lognormal model.

The Model

The model applied to the CPUE time series of data for Namibian orange roughy consists of the delta-lognormal which takes into account the presence of tows with zero catch as described by Lo *et al.* (1992) and Stone and Porter (1999).

The delta distribution is often used in instances when there are a considerable number of zero observations, for which zero and non-zero data are consequently treated separately. Final estimates of abundance are obtained from the product of the proportion and the mean of non-zero observations. For the delta-lognormal model, two linear models are fitted to the commercial CPUE data, one to estimate the proportion of tows for which there is a positive catch, and the other to estimate the standardised CPUE for orange roughy for tows that have a positive catch.

Relative abundance indices of orange roughy are then given by:

$$CPUE_y = \sum_{agg} \left(CPUE_{agg,y}^{+ve} Prop_{agg,y}^{non-zero} \right) A_{agg} \qquad (1)$$

where:

$CPUE_{agg,y}^{+ve}$ is the standardised CPUE index for tows which have positive catches for a given sub-aggregation,

$Prop_{agg,y}^{non-zero}$ is the standardised measure of the proportion of tows that have positive catches for a given sub-aggregation, and

A_{agg} is the geographical area for a given sub-aggregation (Table 1).

Standardised indices for the component of positive catches were obtained by fitting a lognormal model that allows for possible differences in abundance trends in orange roughy in the various aggregations, and assumes the possibility that vessels might operate differently in their first year in the fishery, but have the same degree of “effectiveness” in all subsequent years. Brandão and Butterworth (2003) found that only the vessel *Whitby* showed a significant difference in its first year of operation. Therefore only this vessel was differentiated from its first year in the fishery and all subsequent years. The model to estimate the abundance of positive catches is thus given by:

$$\ln(CPUE^{+ve}) = \mu + \alpha_{vessel} + \beta_{year} + \gamma_{month} + \lambda_{agg} + \eta_{year \times agg} + \varepsilon \qquad (2)$$

where:

μ is the intercept,

vessel is a factor with 12 levels associated with each of the vessels that have operated in the fishery:

Bell Ocean II
Conbaroya Cuarto
Dantago
Emanguluko
Harvest Nicola
Hurinis
Petersen
Sea Flower
Southern Aquarius
Whitby (first year)
Whitby (subsequent years)
Will Watch,

year is a factor with 10 levels associated with the “fishing years” 1994–2003 (note: “1996”, for example, refers to the period July 1996 to June 1997),

month is a factor with 12 levels (January– December),

agg is a factor with 12 levels associated with the four aggregations and their sub-aggregations:

Johnies: *Johnies1*
Johnies2
Johnies3
Johnies4

Frankies: *21 Jump Street*
Frankies Flats
Frankies Outer
Three Sisters

Smifton

Rix: *Rix Inner*

Rix Outer

Hotspot,

year×*agg* is the interaction between year and aggregation (this allows for the possibility of different trends for the different sub-aggregations), and

ε is an error term assumed to be normally distributed.

In the case of the orange roughy tow data, the proportion of tows with a positive catch is either "0" or "1" for an individual tow, and therefore a model for the proportion positive assuming binomially distributed errors is considered, given by:

$$Prop^{non-zero} = \mu + \alpha_{vessel} + \beta_{year} + \gamma_{month} + \lambda_{agg} + \eta_{year \times agg} + \varsigma \quad (3)$$

where

ζ is an error term assumed to be binomially distributed.

Standardised measures of the abundance of orange roughy in positive tows for a given (sub)-aggregation are estimated by calculating:

$$C\hat{P}UE^{+ve}_{agg,y} = \exp\left[\hat{\mu} + \hat{\beta}_{year} + \hat{\lambda}_{agg} + \hat{\eta}_{year \times agg}\right]\psi^{+ve}_{y} \quad (4)$$

where in this application we are standardising on the vessel *Southern Aquarius* and on the month of *August*, and where

ψ^{+ve}_{v} is a correction factor for bias (Lo *et al.* 1992), given by:

$$\psi^{+ve}_{y} = g_m\left[\frac{m+1}{2m}\left(\hat{\xi}^2 - \hat{\xi}^2_{\hat{\theta}}\right)\right] \quad (5)$$

where

$\hat{\xi}^2$ is the residual variance,

m is the degrees of freedom for the estimate of residual variance,

$\hat{\theta}$ is given by $\hat{\mu} + \hat{\beta}_{year} + \hat{\lambda}_{agg} + \hat{\eta}_{year \times agg}$,

$\hat{\xi}^2_{\theta}$ is the variance of $\hat{\theta}$, and

$$g_m(t) = \sum_{p=0}^{\infty} \left[\frac{m^p (m+2p)}{m(m+2)\cdots(m+2p)} \left(\frac{m}{m+1} \right)^p \frac{t^p}{p!} \right]$$

where t is the argument of the function.

Standardised measures of the proportion of positive catches of orange roughy is given by:

$$P\hat{r}op_{agg,y}^{non-zero} = \frac{\exp\left[\hat{\mu} + \hat{\beta}_{year} + \hat{\lambda}_{agg} + \hat{\eta}_{year \times agg}\right]}{1 + \exp\left[\hat{\mu} + \hat{\beta}_{year} + \hat{\lambda}_{agg} + \hat{\eta}_{year \times agg}\right]}. \quad (6)$$

Model Implementation

To take into account movement of orange roughy within a known aggregation, the analyses in Brandão and Butterworth (2002) took into consideration not only tows that lie within the inner strata of an aggregation, but also tows that take place in the outer strata of the aggregation. The levels of the factor for aggregations in the GLMs are then given as the various sub-aggregations. The definition of aggregations and their sub-aggregations given by Brandão and Butterworth (2002) are used in this paper.

Commercial tow information inside the known aggregations of orange roughy in Namibia for the fishing years (July–June) 1994 to 2003, as provided by A. Staby has been used. The year 2003 is incomplete as this fishing year ends only in June 2004. Data until the end of December were available. A total of 17 529 data points was available for the analyses. Of these, 14 688 data points consisted of a non-zero catch. Bottom distances were calculated from the GPS positions for each tow. For tows that did not have haul positions, but did have bottom time information, bottom distances were calculated by the following regression relationship:

Bottom distance [km] = *bottom time* [h] * 5.6082+0.1259.

GLM Results and Discussion

The lognormal model applied to tows with a positive catch (equation (2)) accounts for 45.8% of the total variation of orange roughy positive CPUE. Table 2 shows the parameter estimates obtained for the factor vessel for the component of positive catches and for the proportion of positive tows. Tables 3 to 6 show the index of abundance provided by the delta-lognormal model assuming binomial errors for the proportion positive for each aggregation. Observations are not available for all years in all of the sub-aggregations. Two of the three methods of combining the standardised CPUE indices from each individual sub-aggregation to obtain a standardised CPUE index for each aggregation of Brandão and Butterworth (2002) were used to deal with such empty cells. The first method, referred to as the "zero" method, assumes that empty cells mean that there was no orange roughy in those areas for those years. The second method referred to as the "proportional" method, assumes that although no observations were made, there was orange roughy present. It is further assumed that the amount present is in the same proportion relative to the previous year to that observed in the other constituent sub-aggregation of that aggregation for that year. The overall standardised index for each aggregation is obtained by summing the standardised CPUE for each sub-aggregation multiplied by its associated geographical area (equation(1)).

Figures 1 to 4 show the index of abundance provided by the delta-lognormal model assuming binomial errors for the proportion positive for each aggregation. For each aggregation (except *Hotspot* for which there are no empty cells) a comparison is provided of the indices of abundance of orange roughy obtained by fitting the delta-lognormal model to the CPUE data for the two methods of combining the individual indices of the sub-aggregations. All aggregations show differences between the two methods of combining individual indices (Tables 3 to 5 and Figs. 1 to 3). Differences are most marked in the first few years of the series (mostly for pre–1997).

The standardised CPUE values for 2003 are lower to those for the previous year in all aggregations.

Acknowledgements

A number of people have willingly provided data for this study. Assistance from Arved Staby, in particular, of NatMIRC is gratefully acknowledged, as is funding from the Namibian Deepwater Fishing Industry and the Namibian Ministry of Fisheries and Marine resources.

References

Brandão, A. and Butterworth, D.S. 2002. Standardised CPUE abundance indices for orange roughy off Namibia based on lognormal and delta-lognormal models. Namibian Ministry of Fisheries and Marine Resources document: DWFWG/WkShop/Feb02/doc1.

Brandão, A. and Butterworth, D.S. 2003. Application of lognormal and delta-lognormal linear models to standardised CPUE abundance indices (1994 to 2002) for orange roughy off Namibia. Namibian Ministry of Fisheries and Marine Resources document: DWFWG/WkShop/Mar03/doc2.

Brandão, A. and Butterworth, D.S. 2004. Updated stock assessment of Namibian orange roughy populations under the assumption of intermittent aggregation. Namibian Ministry of Fisheries and Marine Resources document: DWFWG/WkShop/Mar04/doc2.

Lo, N.C., Jacobson, L.D., and Squire, J.L. 1992. Indices of relative abundance from fish spotter data based on delta-lognormal models. Can. J. Fish. Aqaut. Sci. 49: 2515–2526.

Stone, H.H. and Porter, J.M. (1999) Standardized CPUE indices for Canadian bluefin tuna fisheries based on commercial catch rates. International Commission for the Conservation of Atlantic Tunas, collective Volume of Scientific Papers: XLIX (2): 206–221.

Table 1. Geographical area for each sub-aggregation of orange roughy off Namibia.

Aggregation	Sub-aggregation	Area (km^2)
Johnies	***Johnies1***	82.8
	Johnies2	457.2
	Johnies3	198.2
	Johnies4	587.1
Frankies	***21 Jump Street***	39.2
	Frankies Flats	17.8
	Frankies Outer	1 255.0
	Three Sisters	39.6
	Smifton	15.8
Rix	***Rix Inner***	99.4
	Rix Outer	685.6
Hotspot	***Hotspot Inner***	97.3
	Hotspot Outer*	89.0

* Too few tows fall within the *Hotspot Outer* sub-aggregation and therefore these tows are omitted from the GLM analyses.

Table 2. Parameter estimates for the vessel factor when the lognormal model applied to tows with a positive catch (equation (2)) and the model for the proportion positive (equation (3)) are fitted.

Vessel	Vessel factor = $e^{\alpha_{vessel}}$ (positive catches)	Vessel factor = $e^{\alpha_{vessel}}$ (proportion positive)
Bell Ocean II	0.560	0.221
Conbaroya Cuarto	0.316	1.212
Dantago	0.318	0.775
Emanguluko	0.453	1.138
Harvest Nicola	0.221	0.503
Hurinis	0.322	0.627
Petersen	0.434	4.287
Sea Flower	0.502	2324
Southern Aquarius	**1.000**	**1.000**
Whitby (first year)	1.035	1.017
Whitby (subsequent years)	0.506	210
Will Watch	0.997	2022

Table 3. Standardised CPUE series (each normalised to their mean over the years considered) for the *Johnies* aggregation obtained by fitting the delta-lognormal model assuming binomial errors for the proportion positive to the observed CPUE data for Namibian orange roughy. Two methods (“zero” and “proportional” of dealing with years in which no observations were made in the sub-aggregations are considered.

Year	**Standardised indices**	
	“Zero” method	**“Proportional” method**
1994	5.348	7.045
1995	0.771	1.016
1996	1.089	1.435
1997	1.465	0.264
1998	0.536	0.097
1999	0.232	0.042
2000	0.198	0.036
2001	0.112	0.020
2002	0.141	0.026
2003	0.107	0.019

Table 4. Standardised CPUE series (each normalised to their mean over the years considered) for the *Frankies* aggregation obtained by fitting the delta-lognormal model assuming binomial errors for the proportion positive to the observed CPUE data for Namibian orange roughy. Two methods ("zero" and "proportional" of dealing with years in which no observations were made in the sub-aggregations are considered.

Year	**Standardised indices**	
	"Zero" method	**"Proportional" method**
1995	1.309	6.785
1996	4.007	1.271
1997	1.246	0.395
1998	0.594	0.188
1999	0.266	0.090
2000		0.043
2001	0.412	0.155
2002	0.141	0.062
2003	0.026	0.012

Table 5. Standardised CPUE series (each normalised to their mean over the years considered) for the *Rix* aggregation obtained by fitting the delta-lognormal model assuming binomial errors for the proportion positive to the observed CPUE data for Namibian orange roughy. Two methods ("zero" and "proportional" of dealing with years in which no observations were made in the sub-aggregations are considered.

Year	**Standardised indices**	
	"Zero" method	**"Proportional" method**
1995	0.558	2.006
1996	0.665	2.392
1997	4.404	2.606
1998	1.862	1.102
1999	0.369	0.218
2000	0.383	0.227
2001	0.274	0.162
2002	0.287	0.170
2003	0.197	0.117

Table 6. Standardised CPUE series (each normalised to their mean over the years considered) for the *Hotspot* aggregation obtained by fitting the delta-lognormal model assuming binomial errors for the proportion positive to the observed CPUE data for Namibian orange roughy. There are no zero- cells for *Hotspot*, so the “zero” and “proportional” options need not be considered.

Year	Standardised indices
1994	5.347
1995	2.246
1996	0.800
1997	0.289
1998	0.459
1999	0.240
2000	0.093
2001	0.155
2002	0.327
2003	0.044

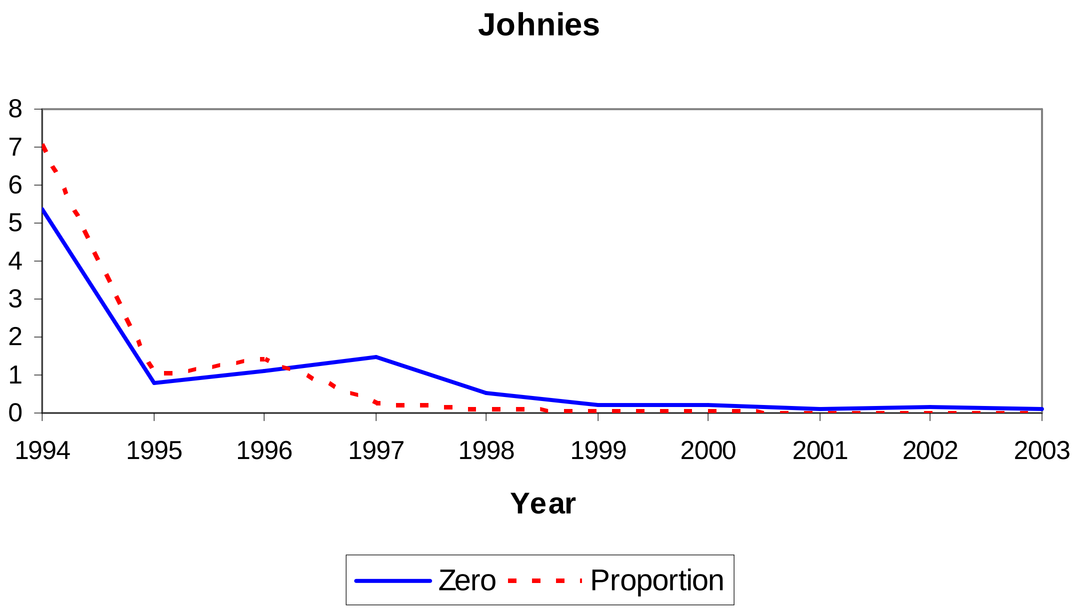

Figure 1. Index of abundance for the *Johnies* aggregation (normalised to its mean over the ten year period) for Namibian orange roughy obtained from fitting the delta-lognormal model assuming binomial errors for the proportion positive. Results are shown for the two methods of dealing with empty cells when combining the indices from sub-aggregations.

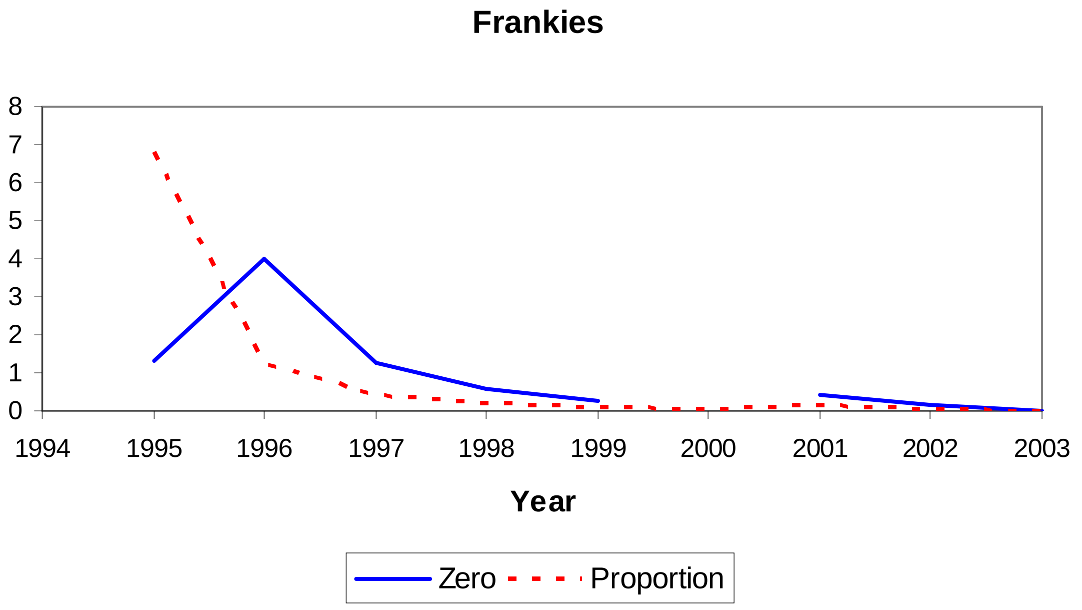

Figure 2. Index of abundance for the *Frankies* aggregation (normalised to its mean over the ten year period) for Namibian orange roughy obtained from fitting the delta-lognormal model assuming binomial errors for the proportion positive. Results are shown for the two methods of dealing with empty cells when combining the indices from sub-aggregations.

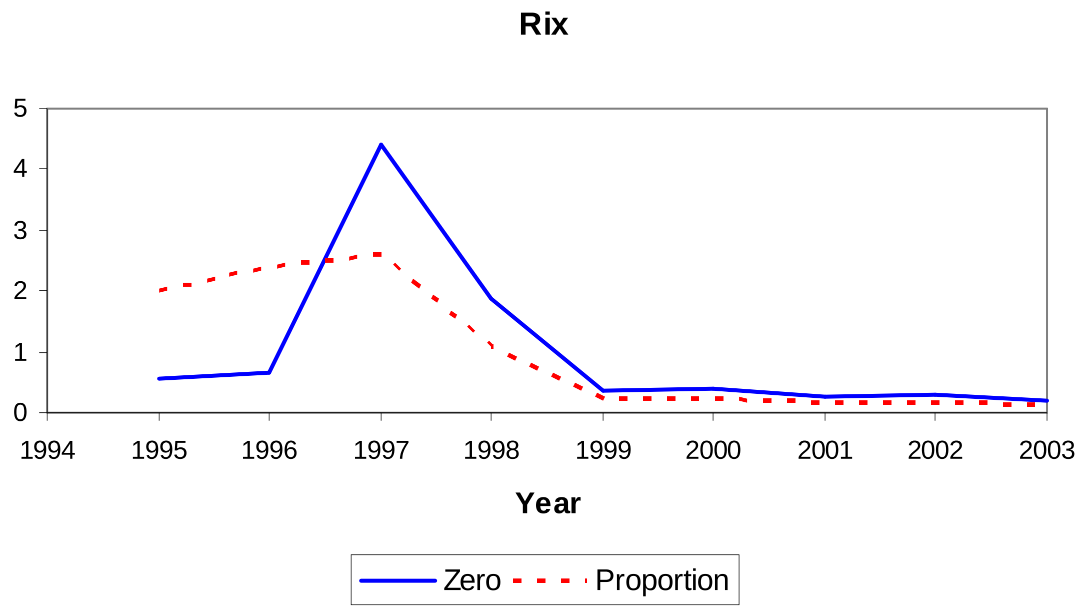

Figure 3. Index of abundance for the *Rix* aggregation (normalised to its mean over the ten year period) for Namibian orange roughy obtained from fitting the delta-lognormal model assuming binomial errors for the proportion positive. Results are shown for the two methods of dealing with empty cells when combining the indices from sub-aggregations.

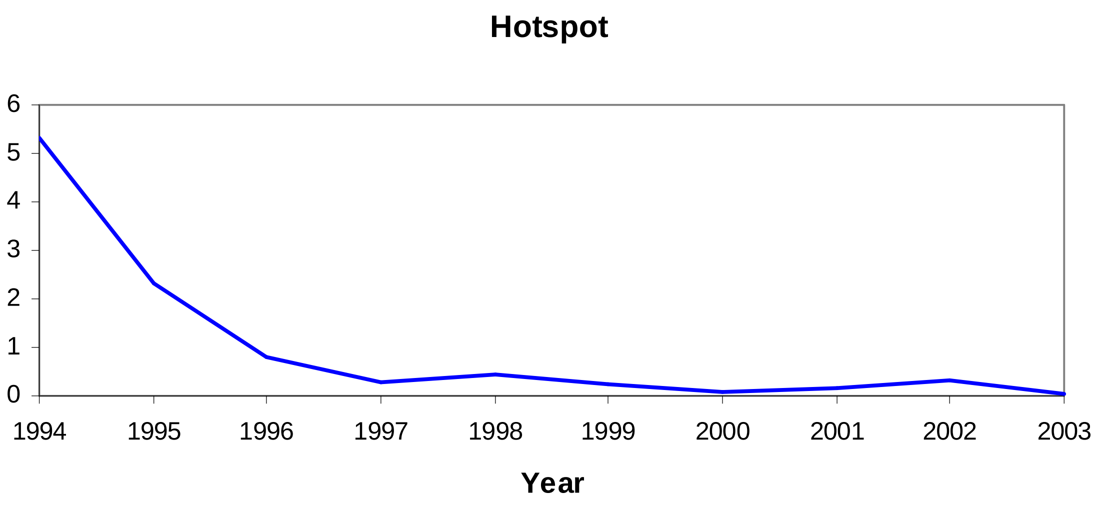

Figure 4. Index of abundance for the *Hotspot* aggregation (normalised to its mean over the nine year period) for Namibian orange roughy obtained from fitting the delta-lognormal model assuming binomial errors for the proportion positive.