SOME INSIGHTS INTO THE ESTIMABILITY OF NON-LINEAR DEPENDENCE OF CPUE ON ABUNDANCE IN ORANGE ROUGHY FISHERIES, BASED UPON A SIMPLE AGE-AGGREGATED PRODUCTION MODEL

D S Butterworth and A Brandão

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

ABSTRACT

A simple Schaefer-like production model is used in simulations to assess the potential benefits or otherwise of attempting to estimate the extent of non-linearity in a CPUE-abundance relationship (reflected by the parameter β). The resource situation considered is similar to that analysed by Hicks (2005), though this paper evaluates only estimation based upon CPUE data alone. Simulations are conducted for three pre-exploitation values of orange roughy abundance, which correspond to a resource which at present is either still declining, is approximately stable, or is increasing slightly. CPUE data are generated for five different values of the β parameter, and estimators which either fix this value, or try to estimate it from the fit of the population model to the CPUE data, are considered. An initial impression of the results obtained for estimates of the initial biomass, current depletion and current replacement yield for the simulated resource is that, in terms of root -mean-square errors, estimating β can achieve smallish gains in some circumstances, but leads to much larger losses in others.

INTRODUCTION

A matter which has an important impact on the results of assessments of some New Zealand orange roughy resources is whether such analyses should internally attempt to estimate the extent of possible non-linearity in the relationship between CPUE and abundance, rather than assume this relationship to be one of linear proportionality as is customary.

Hicks (2005) has addressed this question using a relatively complex age-structured model of the underlying resource dynamics. Here the intention is to complement Hicks' results by using a simple age-aggregated production model (AAPM) for the resource dynamics, to see whether this can contribute further insight.

METHODS

The intention here is, as far as possible, to mimic the situation considered by Hicks (2005), which was based on the New Zealand Mid-East Coast orange roughy fishery, to facilitate the contrasting of results. Accordingly, the simulations carried out here utilise the catches in that fishery from 1982 to 2003, and assume an annual CPUE index with a CV of 0.28 to be available over the period from 1984 to 2003 with the exception of 1989.

The deterministic AAPM used for the underlying resource dynamics is a simple "distortion" of the Schaefer model that yields *MSY* at a depletion (*MSYL* or B_{msy}) of $0.3B_0$ (as conventionally assumed for orange roughy resources) rather than $0.5B_0$ (see Appendix). The *MSYR* (= MSY/B_{msy}) parameter of the model is set at 0.04 and is known exactly to the estimator – this is roughly equivalent to tests of age-structured model estimators for orange roughy which assume natural mortality M and steepness h to be known without error. Three underlying scenarios for current resource status are investigated, spanning a range of current resource depletion (B_{2004}/B_0) of some 10-35% within which current estimates (depending on assumptions) for the Mid-East Coast orange roughy resource lie (P. Mace, pers. commn). The values of B_0 selected for these scenarios are 130 000, 140 000 and 160 000 thousand tons, and correspond respectively to current depletions of 10, 20 and 33% of the resource, or equivalently to instances where the resource over the most recent five years has been declining quite rapidly (5.5% p.a), declining slightly (1% p.a.) and increasing slightly (0.2% p.a.).

The possible non-linearity in the relationship between CPUE and abundance (B) is modelled as:

$$CPUE = qB^{\beta}e^{\varepsilon} \text{ where } \varepsilon \sim N(0, \sigma^2) \tag{1}$$

where in a slight extension of Hicks (2005), values of β of 0.625, 1, 1.6, 2 and 4 are considered. As in Hicks (2005) $\sigma = 0.28$. The combination of three values for B_0 and five values for β leads to 15 scenarios (see Table 1). For each of these scenarios, 100 sets of CPUE values over the period from 1982[1] to 2003 (excluding 1989) are generated to provide a basis to contrast the performance of different estimators.

The two AAPM-based estimators considered both assume exact knowledge of the form of the surplus production function and of the *MSYR* parameter. The first assumes linear proportionality between CPUE and B (i.e. that $\beta = 1$), and estimates B_0, σ and q. The second attempts estimation of β as well as these other three parameters.

Further details of the underlying model and the estimators may be found in the Appendix.

[1] At the stage of checking of the final results for this paper, it was realised that CPUE data had inadvertently been generated from 1982, rather than from 1984 as intended to duplicate Hicks (2005). This should not, however, have a major impact on the results.

RESULTS AND DISCUSSION

Results of the simulations in terms of distributions of estimates (reflected by medians and 90% probability intervals) compared to true underlying values are shown in Fig. 1 for three quantities of management interest: B_0, B_{2003}/B_0 and the current replacement yield RY_{2003} (i.e. the catch that would maintain the biomass at its 2003 level). For the case where $\beta = 4$, the results may not be entirely reliable as there were indications that the estimation minimisation had not converged on all occasions. This also occurred for lower values of β, but with much less frequency.

As expected, estimating β increases estimation variance. The question is whether there is sufficient compensatory decrease in bias to justify such estimation. Note that although estimators which assume $\beta = 1$ when this happens to be the true value give unbiased results for the three quantities shown in Fig. 1, bias is not totally removed when β is estimated (even when the true β is 1). Because of possible confounding as a result of convergence difficulties, results for $\beta = 4$ are not considered in the comments below.

Patterns amongst the results differ depending on the true value for B_0, though the more fundamental distinction is likely whether the biomass trend is still downwards, is almost stable, or is slightly increasing. As far as bias is concerned, for the three quantities considered:

- B_0: marginal improvement for the downward trend case; otherwise little to choose.
- B_{2003}/B_0: definite improvement for the almost stable case; otherwise minimal.
- RY_{2003}: comments as for B_{2003}/B_0.

To assist gauge whether these decreases in bias do offset the increases in variance, Table 2 lists the ratio of the root-mean-square errors (RMSE's) for the estimator estimating β to that for the estimator fixing $\beta = 1$. A general impression from this Table is that while estimating β is advantageous in some circumstances, the benefit obtained in terms of the RMSE's are not very large; on the other hand, there are a number of other situations where estimating β leads to a fairly substantial deterioration in overall estimation performance.

The calculations reported have considered the case where only CPUE information (being a relative measure of abundance) is available. Hydroacoustic surveys (for example) have the potential to provide information on abundance B in absolute terms. A possible further step for computations such as those reported here is to consider the impact of the availability of such information on the performance of different estimators, but taking account of the fact that such information will have associated variance and may also be biased.

REFERENCE

Hicks, A C. 2005. Estimating a parameter for non-linear CPUE in orange roughy fisheries. Unpublished report dated 23 March 2005, held by Ministry of Fisheries, Wellington.

Table 1. Description of the different scenarios and estimators for the combinations considered in this paper.

Combination	Scenario	Estimator specification
1	B_0 = 130 000, β = 0.625	Set β = 1
2	B_0 = 130 000, β = 1.0	Set β = 1
3	B_0 = 130 000, β = 1.6	Set β = 1
4	B_0 = 130 000, β = 2.0	Set β = 1
5	B_0 = 130 000, β = 4.0	Set β = 1
6	B_0 = 130 000, β = 0.625	Estimate β
7	B_0 = 130 000, β = 1.0	Estimate β
8	B_0 = 130 000, β = 1.6	Estimate β
9	B_0 = 130 000, β = 2.0	Estimate β
10	B_0 = 130 000, β = 4.0	Estimate β
11	B_0 = 140 000, β = 0.625	Set β = 1
12	B_0 = 140 000, β = 1.0	Set β = 1
13	B_0 = 140 000, β = 1.6	Set β = 1
14	B_0 = 140 000, β = 2.0	Set β = 1
15	B_0 = 140 000, β = 4.0	Set β = 1
16	B_0 = 140 000, β = 0.625	Estimate β
17	B_0 = 140 000, β = 1.0	Estimate β
18	B_0 = 140 000, β = 1.6	Estimate β
19	B_0 = 140 000, β = 2.0	Estimate β
20	B_0 = 140 000, β = 4.0	Estimate β
21	B_0 = 160 000, β = 0.625	Set β = 1
22	B_0 = 160 000, β = 1.0	Set β = 1
23	B_0 = 160 000, β = 1.6	Set β = 1
24	B_0 = 160 000, β = 2.0	Set β = 1
25	B_0 = 160 000, β = 4.0	Set β = 1
26	B_0 = 160 000, β = 0.625	Estimate β
27	B_0 = 160 000, β = 1.0	Estimate β
28	B_0 = 160 000, β = 1.6	Estimate β
29	B_0 = 160 000, β = 2.0	Estimate β
30	B_0 = 160 000, β = 4.0	Estimate β

Table 2. Ratio of RMSE (estimating β) to RMSE (fix $\beta = 1$) for the 15 scenarios considered.

Scenario		RMSE ratio		
B_0	β	B_0	B_{2003}/B_0	RY_{2003}
130 000	0.625	3.543	2.718	2.312
130 000	1.0	1.652	0.981	0.991
130 000	1.6	0.846	0.935	0.922
130 000	2.0	0.948	0.988	0.984
130 000	4.0	0.490	1.222	1.251
140 000	0.625	5.209	2.566	2.319
140 000	1.0	4.128	1.029	0.996
140 000	1.6	1.422	0.807	0.828
140 000	2.0	0.969	0.840	0.840
140 000	4.0	3.349	0.585	0.382
160 000	0.625	7.510	1.243	4.114
160 000	1.0	12.64	2.051	2.080
160 000	1.6	3.040	2.051	0.818
160 000	2.0	1.072	1.204	0.747
160 000	4.0	0.969	0.974	1.022

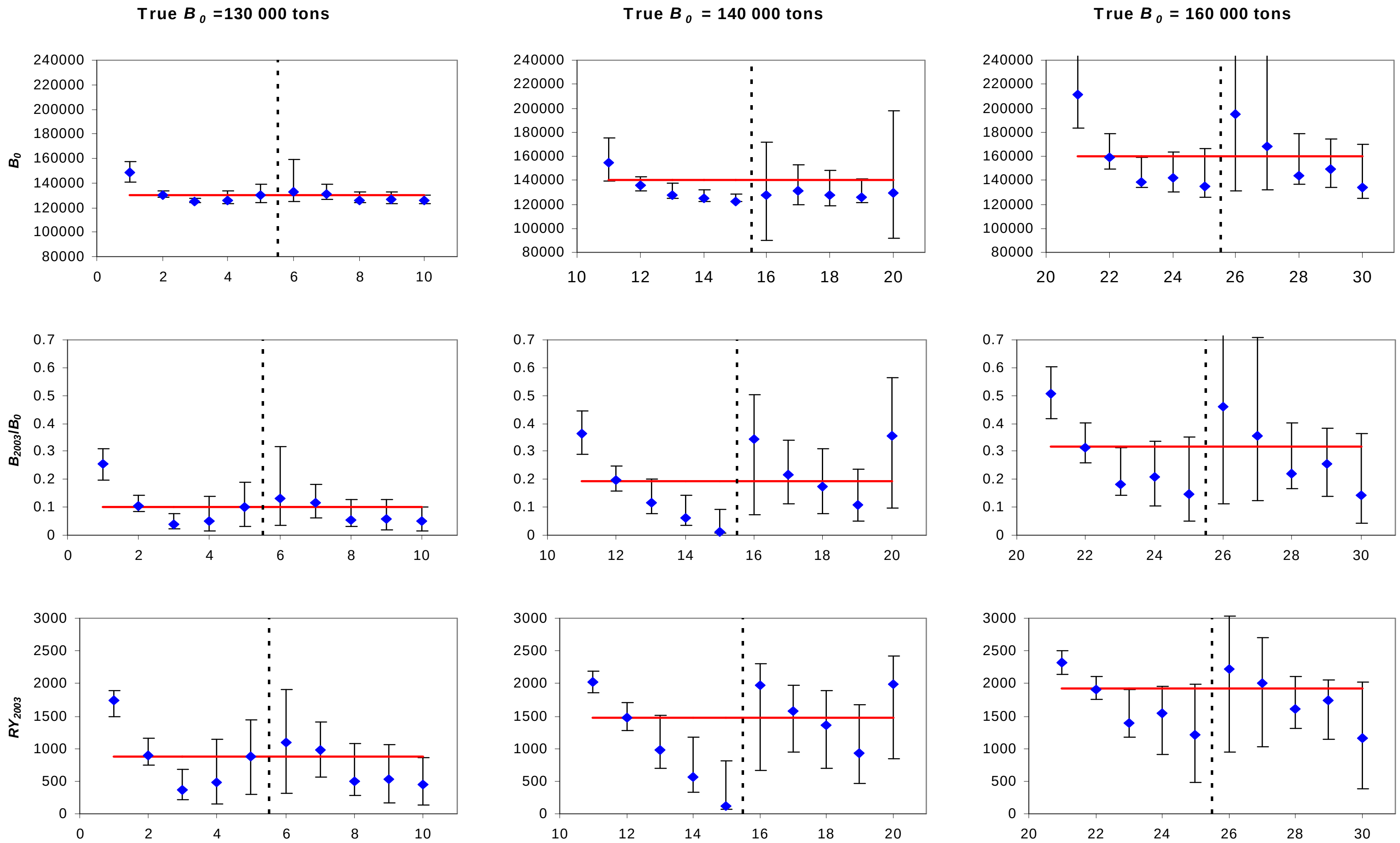

Figure 1. Results of the simulations in terms of distributions of estimates (reflected by medians and 90% probability intervals) compared to true underlying values for three quantities of management interest: B_0, B_{2003}/B_0 and RY_{2003}. The horizontal axis denotes the different combinations of scenarios and estimators considered (see Table 1), with the first five corresponding to combinations for which the parameter β is set to 1, and the next five to ones for which β is estimated. Note that within each group, the true value of β changes left to right from 0.625, to 1, to 1.6, to 2.0, and to 4.0.

APPENDIX

Underlying Population Model

The dynamics of the resource is modelled by an age-aggregated production model (AAPM) as follows:

$$B_{t+1} = B_t + f(B) - C_t \qquad \text{(A.1)}$$

$$CPUE_t = qB_t^{\beta} e^{\varepsilon_t} \qquad \text{(A.2)}$$

where:

- B_t is the biomass at the start of year t,
- $f(B)$ is the surplus production function,
- C_t is the catch made in year t (see Table A.1),
- $CPUE_t$ is the (simulated) CPUE in year t,
- q is the constant in the relationship between CPUE and biomass, and
- ε_t is the log of the observation error for the CPUE in year t, which is assumed to be normally distributed with constant variance: N(0, σ^2).

The annual surplus production for the Schaefer (logistic) model was adjusted to have B_{MSY}/B_0 to be at 30% (where B_0 is the pre-exploitation biomass) instead of at 50%. To ensure derivative continuity, the production function was accordingly given by:

$$f(B_t) = \begin{cases} rB_t\left(1 - B_t/0.6B_0\right) & \text{for } B_t \le 0.3B_0 \\ 0.1225r\left(1 - B_t/B_0\right)(1 + 2.5B_t) & \text{for } B_t > 0.3B_0 \end{cases} \qquad \text{(A.3)}$$

where:

- r is the intrinsic growth rate parameter, set here to be 0.08, so that $MSYR$ = 0.04, and
- B_{1982} is set equal to B_0.

This adjustment was preferred to use of the Pella-Tomlinson form because that has an unrealistic infinite slope at the origin for B_{MSY}/B_0 values that are as low as 30%.

Equations (A.1)–(A.3) are used in simulating CPUE abundance indices, where the constant in the CPUE series is taken to be 1. The values considered for β in equation (A.2) were 0.625, 1.0, 1.6, 2, and 4.

Fig. A.1 shows the form of the surplus production function, and Fig. A.2 plots the biomass trajectories for the three values considered for B_0.

Estimators considered

The model parameters, q, B_0, and σ, and also β for situations where this is estimated rather than fixed at 1, are estimated by minimising the negative log-likelihood function:

$$-\ln L = \sum_{t=1982, t\neq 1989}^{2003} \left[\ln\sigma + \frac{1}{2\sigma^2}\left(\ln CPUE_t - \ln(q) - \beta \ln(B_t)\right)^2 \right] \qquad \text{(A.4)}$$

where

q is the constant in the CPUE-abundance relationship, whose maximum likelihood estimate is given by:

$$\ln\hat{q} = \frac{1}{n}\sum_{t}^{CPUE}\left(\ln(CPUE_t) - \hat{\beta}\ln\hat{B}_t\right), \text{ and}$$

σ is the standard deviation of the CPUE series, whose maximum likelihood estimate is given by:

$$\hat{\sigma} = \sqrt{\frac{1}{n}\sum_{t}^{CPUE}\left(\ln(CPUE_t) - \ln\hat{q} - \hat{\beta}\ln(\hat{B}_t)\right)^2} \; .$$

Table A.1. Yearly catches of orange roughy (in tons) considered in this paper.

Year	Catches
1982	700
1983	4000
1984	9000
1985	10000
1986	10000
1987	10000
1988	12000
1989	11000
1990	12000
1991	11000
1992	11000
1993	9500
1994	7000
1995	6000
1996	1900
1997	2200
1998	2300
1999	2300
2000	2600
2001	1800
2002	1500
2003	900

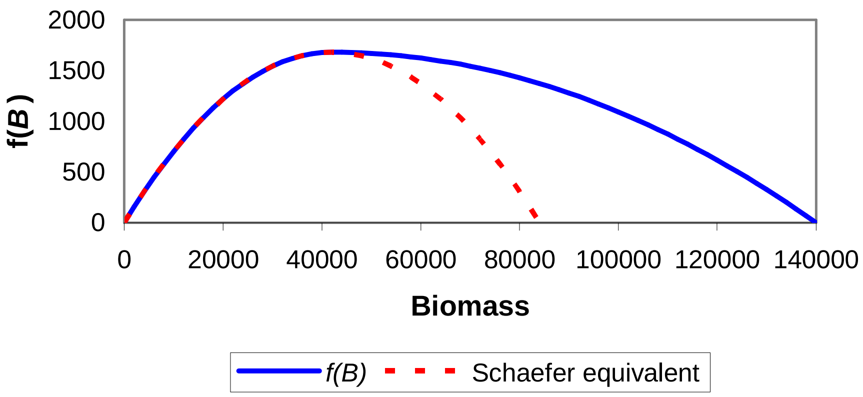

Figure A.1. Production function when B_0 is 140 000 tons, showing also the corresponding form for a Schaefer function.

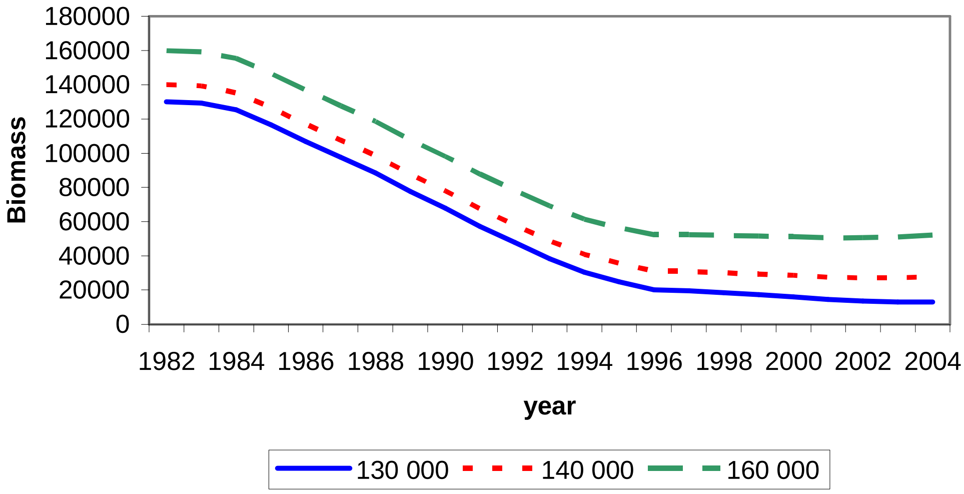

Figure A.2. Deterministic biomass trajectories that correspond to the three values chosen for B_0.