BEN/DEC04/H/SA/3c~~WG/03/04/D:H:5 – revised~~

Update of the ~~Revised:~~

Proposed Model Framework to Form the Basis for Immediate Future Assessments of the South African Hake Resource

RA Rademeyer and DS Butterworth

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

~~April~~ November 2004

1. Introduction

The final report of the BENEFIT-NRF-BCLME stock assessment workshop held in January 2004 provides recommendations for the future assessment of hake off South Africa (BENEFIT, 2004). The workshop recommended that the approach include the following features:

1) age-length keys for one year should not be applied to the length-frequency data for another year – rather, if length-frequency data are available for a year for which an age-length key is not available, the model should be fitted to the length-frequency data for that year;~~.~~

2) a model which considers both species simultaneously should be developed and its results aggregated to fit to data that cannot be disaggregated between species (e.g. the ICSEAF CPUE series);

3) the initial spatial structure of the model should involve four components (west coast inshore, west coast offshore, south coast inshore, south coast offshore) – the definition of inshore and offshore should be based on biological considerations and data availability;

the initial version of the model should estimate component- and species-specific "selectivity" (which includes both gear selectivity and availability) patterns;

4)

5) the values for the parameter that determines the split among species of the exploitation rate on fully-selected animals should be calculated to mimic the catches by species each year, with a prior placed on the extent to which it may vary over time;

~~the values for the parameter that determines the split among species of the exploitation rate on fully-selected animals should be calculated to mimic the catches by species each year, with a prior placed on the extent to which it may vary over time;~~

6) the longline catches should be split to species, e.g. using observer data to develop a suitable algorithm;

future assessments should be sex-structured with selectivity defined in terms of length (rather than age) because hake are sexually dimorphic – this will require the collection and use of sex-structure data; and

~~7) the longline catches should be split to species, e.g. using observer data to develop a suitable algorithm; and~~

8) allowance should be made for age-determination error when fitting the catch-at-age information.

A framework is outlined below which incorporates these features. Recommendations ~~6~~7) ~~to~~ and 8) are however ~~are~~ not taken into account at this stage as the necessary data are not yet available, and seem unlikely to become so within the next few ~~m~~months.

Formatted: Bullets and Numbering

Formatted: Bullets and Numbering

Formatted: Bullets and Numbering

Formatted: Bullets and Numbering

Formatted: Bullets and Numbering

This paper includes a number of questions, marked **?** in the text, on which the Working Group is asked to provide guidance prior to the initiation of computations along the lines suggested.

2. Data

A strawmanThe structure of the data available is summarised in the tables below to facilitate discussions. This information is divided into four periods, it differentiates between the two species (*M. capensis* and *M. paradoxus*), three four fleets (Small Trawler – e.g. as operate from Mossel Bay, previously refered to as inshore trawler; Standard Trawler, previously refered to as offshore trawler; and Longline/ and handlineHandline) and the surveys (treated effectively as a fifth fleet), and four regions (south coast inshore, south coast offshore, west coast inshore and west coast offshore).

From discussions in the MCM Demersal Working Group, it was suggested that the boundary separating the inshore and offshore regions be at the 300m isobath on the west coast and at the 200m isobath on the south coast. Furthermore, it was decided that the appropriate boundary to use to split the west and south coasts should start from Cape Agulhas (20°E) to 36°S (as the old ICSEAF division), then from MCM block 580, staggered to the south-east by block (20'x20') in order to include the whole of Brown's Bank in the west coast.

? What would a sensible depth be to separate the inshore and offshore regions? 300m or 400m? (Note that 200m would be too shallow; although the Mossel Bay Small Trawlers operate mainly within this depth level, hardly any Standard Trawler operations take place here.)

? What is the most appropriate line to use to split west and south coasts?

Period 1917 – 1947~~?~~

This period represents the beginning of the fishery, when the fishery confined (presumably) its activities to fishing grounds relatively close inshore, around Cape Town and to a lesser extent off the south coast. The bulk of the catch, which did not exceed 50 000 tons p.a. over this period, must have been *M. capensis*.

Note that ╳ in the diagrams that follow indicates a presumption of zero effort and zero catches for the fleet/region/species all concerned over the full period under consideration. Within such cells, data sources are marked ✓/X to indicate where such data are/are not available; positive indication mean for some, but not necessarily <u>all</u> years of the period.

		Inshore region		Offshore region	
	Fleet	*M. capensis*	*M. paradoxus*	*M. capensis*	*M. paradoxus*
West coast	Standard Trawler	Catch ✓ CPUE X CAA X	╳	╳	╳
West coast	Longline	╳	╳	╳	╳
	Small Trawler	╳	╳	╳	╳
South coast	Standard Trawler	Catch X CPUE X CAA X	╳	╳	╳
South coast	Longline	╳	╳	╳	╳
South coast	Handline	╳	╳	╳	╳

<u>General assumptions:</u>

- <u>standard trawlers only operated in the inshore regions during this period</u>
- <u>only *M. capensis* was caught during this period</u>
- <u>hake was landed on the south coast from mid-1930's – catch statistics are not available</u>

Formatted: Bullets and Numbering

~~? Is 1947 a sensible choice to end this period?~~

~~? Is it appropriate to assume that the standard trawlers only operated in the inshore regions during this period?~~

~~? Is it appropriate to assume that only *M. capensis* was caught during this period?~~

~~? Was the catch from the standard trawlers really all from the west coast (as current official statistics indicate)?~~

~~? Were the standard trawlers of this period more like the small trawler type?~~

Period 1948~~?~~ - ~~1966~~1959

During this period, locally based trawlers were presumably starting to move to deeper waters on the west coast, catching both *M. capensis* and *M. paradoxus*.

		Inshore region		**Offshore region**	
	Fleet	*M. capensis*	*M. paradoxus*	*M. capensis*	*M. paradoxus*
West coast	Standard Trawler	Catch not split by spp or region CPUE not split by spp or region CAA X			
	Longline				
South coast	Small Trawler				
	Standard Trawler	Catch X CPUE X CAA X			
	Longline				
	Handline				

Assumptions:

- standard trawlers moved to deeper waters on the west coast during this period
- both *M. capensis* and *M. paradoxus* were caught on the west coast
- standard trawlers on the south coast stayed in the inshore region and catch *M. capensis* only during this period

Formatted: Bullets and Numbering

~~? When did the south coast catches start? During this period?~~

~~? When did the small trawlers start appreciable fishing for hake (from Mossel Bay)?~~

~~? Would standard trawlers on the south coast have taken any *M. paradoxus*? (The answer likely depends on how the line dividing the west/south coasts is drawn.)~~

Period ~~1967~~ 1960 - 1976

The first ~~'official'~~ ~~c~~catches ~~on the south coast~~by small trawlers on the south coast. Standard trawlers on the south coast moving to deeper water and starting to catch *M. paradoxus*. This is also the period over which the foreign fleets operated.

		Inshore region		Offshore region	
	Fleet	*M. capensis*	*M. paradoxus*	*M. capensis*	*M. paradoxus*
West coast	Standard Trawler	Catch not split by spp or region CPUE not split by spp or region CAA X			
West coast	Longline				
South coast	Small Trawler	Catch ✓ CPUE X CAA X			
South coast	Standard Trawler	Catch not split by spp or region CPUE not split by spp or region CAA X			
South coast	Longline				
South coast	Handline				

Assumptions:

- small trawlers fish only in the inshore region on the south coast
- small trawlers only catch *M. capensis*
- standard trawlers on the south coast move to deeper water and start catching *M. paradoxus*

Formatted

Formatted: Bullets and Numbering

~~? The catches before 1974 are not split between the small trawler and the standard trawler fleets. How is this best done?~~

~~? Catches by the small trawler fleet are assumed to consist of *M. capensis* only. Is this reasonable?~~

Period 1978 - present

From 1978 onwards, depth information has been recorded for tows. This information allows the catch and effort to be split between the inshore and offshore regions. It also allows the catch to be species-disaggregated, using a proportion-by-depth (and by size) relationship. Survey data are also available on a species and depth basis.

		Inshore region		Offshore region	
	Fleet	*M. capensis*	*M. paradoxus*	*M. capensis*	*M. paradoxus*
West coast	Standard Trawler	Catch ✓ CPUE ✓	Catch ✓ CPUE ✓	Catch ✓ CPUE ✓	Catch ✓ CPUE ✓
West coast	Standard Trawler	CAA not split by spp or region			
West coast	Longline	Catch ✓ CPUE X CAA ?	Catch ✓ CPUE X CAA ?		
West coast	Survey	biomass ✓ CAA ✓	biomass ✓ CAA ✓	biomass ✓ CAA ✓	biomass ✓ CAA ✓
South coast	Small Trawler	Catch ✓ CPUE X CAA ✓			
South coast	Standard Trawler	Catch ✓ CPUE ✓	Catch ✓ CPUE ✓	Catch ✓ CPUE ✓	Catch ✓ CPUE ✓
South coast	Standard Trawler	CAA X ?			
South coast	Longline	Catch ✓ CPUE X CAA ✓			
South coast	Handline	Catch ✓ CPUE X CAA ✓			
South coast	Survey	biomass ✓ CAA ✓	biomass ✓ CAA ✓	biomass ✓ CAA ✓	biomass ✓ CAA ✓

Assumptions:

- small trawlers fish only in the inshore region on the south coast
- small trawlers only catch *M. capensis*
- longline fleet operates only in the inshore regions on both coasts
- longline fleet on the west coast catches both *M. capensis* and *M. paradoxus*
- longline fleet on the south coast catches *M. capensis* only
- handline fleet operates in the inshore region only
- handline fleet catches *M. capensis* only

Formatted: Bullets and Numbering

Formatted

? Which species/depth disaggregation algorithm should be used? With or without the size information from industry classes incorporated?

? Similarly, which algorithm should be used to split the catch to obtain the CPUE series by species?

? Is it appropriate to assume that the longline/handline fleet operates in the inshore regions only and catches *M. capensis* only (at least as an initial approximation)?

? What is the status of data on longline/handline catch from the west coast and CPUE for both coasts?

? For the standard trawler fleet, splitting the catch by region should be almost exact, but the split by species is only approximate. In the model therefore, it is proposed that the species-aggregated catch by region would be assumed to be known almost exactly, but allowance would be made for errors in the fit to the species-disaggregated catches. Is this reasonable?

? Commercial catch-at-age data should be updated (to include data up to 2003). By when might this be expected to be completed (or should models rather fit to catch-at-length data for the currently missing years)?

? Should an attempt be made to use CPUE data from the south coast small trawler fleet in Mossel Bay?

Summary of data which still needs to be prepared

Catches:

a) Pre-1978, redistribute catches by coast/region once the queries above have been addressed.

b)Post-1978, splitting the catches from the standard trawler fleet by species, coast and region.

c)b) Catches for the longline(/handline) fleet on the west coast.

Formatted: Bullets and Numbering

CPUE:

a) GLM-standardised CPUE series for the standard trawl fleet for each combination of species (*M. capensis* and *M. paradoxus*) and region (south coast inshore, south coast offshore, west coast inshore and west coast offshore), from 1978 to the present.

Commercial catches-at-age:

a) Update the west coast, species-combined, catch-at-age data from the standard trawler fleet to include data up to 2003.

b) Update the south coast catch-at-age data for the small trawler fleet (assumed to consist of *M. capensis* only) to include data up to 2003.

c)Update the south coast catch-at-age data for the longline fleet (also assumed to consist of *M. capensis* only) to include data up to 2003 (why are such data not available in 1998 and 1999 when the 2000 data are?).

Formatted: Bullets and Numbering

d)Are there data available for the west coast longline fleet?

e)c) Are there catch-at-age data available for the standard trawlers on the south coast?

Survey biomass estimates:

a) For each coast, the biomass estimates for each species need to be split between the inshore and offshore regions.

Survey catches-at-age:

a) For each coast, the catch-at-age estimates for each species need to be split between the inshore and offshore regions.

Other recommendations from the BENEFIT workshop that need to be discussed in due course

B.1 The catch by the handline sector and its species-, sex- and size-structure should be monitored.

B.2 The observer data should be used to test the validity of the algorithms for splitting the past commercial trawl catches among species.

B.3 The algorithm used to split the historical trawl catches to species should take the fish size as well as depth of capture into account.

B.6 The observer programme for South Africa needs to provide regular and reliable information on the species-split of the hake catch.

B.7 The spatial and temporal trends in the catch and effort data for the longline fishery should be analyzed.

B.9 Industry should be consulted to develop alternative hypotheses regarding the levels and spatial distribution of the historical catches.

3. Methods

The model proposed for future assessments of the southern African hake stocks is an ASPM similar to those used for "standard" assessments (Rademeyer, 2003, for example), but involves a single spatially disaggregated formulation, in which both species are assessed separately. The initial spatial structure considered will include four regions: all combination of west coast/south coast and inshore/offshore.

The model equations and the general specifications of the model proposed are described below, followed by details of the contributions to the log-likelihood function from the different data considered. Quasi-Newton minimisation is to be used to minimise the total negative log-likelihood function (implemented using AD Model Builder™, Otter Research, Ltd.).

3.1 Population Dynamics

3.1.1 Numbers-at-age

The resource dynamics of the southern African hake stocks are modelled by the following set of population dynamics equations:

$$N_{s,y+1,0} = R_{s,y+1} \tag{1}$$

$$N_{s,y+1,a+1} = \left(N_{sya}\, e^{-M_{sa}/2} - \sum_{rf} C_{srfya} \right) e^{-M_{sa}/2} \qquad \text{for } 0 \le a \le m_s - 2 \tag{2}$$

$$N_{s,y+1,m} = \left(N_{sy,m-1}\, e^{-M_{s,m-1}/2} - \sum_{rf} C_{srfy,m-1} \right) e^{-M_{s,m-1}/2} + \left(N_{sym} e^{-M_{sm}/2} - \sum_{rf} C_{sfym} \right) e^{-M_{sm}/2} \tag{3}$$

where

N_{sya} is the number of fish of species *s* and age *a* at the start of year *y*,

R_{sy} is the recruitment (number of 0-year-old fish) of species *s* at the start of year *y*,

M_{sa} denotes the natural mortality rate on fish of species *s* and age *a*,

C_{srfya} is the number of fish of species *s* and age *a* caught in year *y* by fleet *f* in region *r*, and

m_s is the maximum age considered (taken to be a plus-group) for species *s*.

These equations simply state that for a closed population, i.e. with no immigration or emigration, the only sources of loss are natural mortality (predation, disease, etc.) and fishing mortality (catch). They reflect Pope's approximation (Pope, 19841972) (the catches are assumed to be taken as a pulse in the middle of the year) rather than the more customary Baranov catch equations (Baranov, 1918) (where catches are incorporated in the form of a continuous fishing mortality). As long as mortality rates are not too high, the differences between the Baranov and Pope formulation will be minimal. Tests showed this approximation to be adequate for the hake stocks (Punt, University of Washington, pers. commn).

3.1.2 Recruitment

Next year's recruitment depends upon the reproductive output of this year's fish. The number of recruits of each species (i.e. new zero-year old fish) at the start of year *y* in region *r* is assumed to be related to the spawning stock size (i.e., the biomass of mature fish) in that region by a stock-recruitment relationship. Traditionally, the Beverton-Holt function (Beverton and Holt, 1957) has been used for southern African hake assessments.

The Beverton-Holt stock-recruitment relationship, allowing for annual fluctuations, is written as:

$$R_{sy} = \frac{\alpha_s B_{sy}^{sp}}{\beta_s + B_{sy}^{sp}} e^{(\varsigma_{sy} - \sigma_R^2/2)} \qquad (4)$$

where

α_s and β_s are spawning biomass-recruitment relationship parameters for species *s*, α being the maximum number of recruits produced, and β the spawning stock needed to produce a recruitment equal to $\alpha/2$, in the deterministic case;

ς_{sy} reflects fluctuation about the expected recruitment for species *s* in year *y*, which is assumed to be normally distributed with standard deviation σ_R σ_R (whose value is input in the applications considered here); these residuals are treated as estimable parameters in the model fitting process. Estimating the stock-recruitment residuals is made possible by the availability of catch-at-age data, which give some indication of the age-structure of the population. The $-\sigma_R^2/2$ term is to correct for bias given the skewness of the log-normal distribution; it ensures that, on average, recruitments will be as indicated by the deterministic component of the stock-recruitment relationship;

Formatted

B_{sy}^{sp} is the spawning biomass of fish of species *s* at the start of year *y*, computed as:

$$B_{sy}^{sp} = \sum_{a=1}^{m} f_{sa} w_{sa} N_{sya} \qquad (5)$$

where

w_{sa} is the begin-year mass of fish of species *s* and age *a*, and

f_{sa} is the proportion of fish of species *s* and age *a* that are mature.

In order to work with estimable parameters that are more biologically meaningful, the stock-recruitment relationship is re-parameterised in terms of the pre-exploitation equilibrium spawning ("virgin") biomass, K_s^{sp}, and the "steepness", h_s, of the stock-recruitment relationship, which is the proportion of the virgin recruitment (R_{1s}) that is realised at a spawning biomass level of 20% of the virgin spawning biomass:

$$\alpha_s = \frac{4h_s R_{s1}}{5h_s - 1} \qquad (6)$$

and

$$\beta_s = \frac{K_s^{sp}(1-h_s)}{5h_s - 1} \qquad (7)$$

where

$$R_{s1} = K_s^{sp} / \left[\sum_{a=1}^{m-1} f_{sa} w_{sa} \exp\left(-\sum_{a'=0}^{a-1} M_{sa'} \right) + f_{sm} w_{sm} \frac{\exp\left(-\sum_{a'=0}^{m-1} M_{sa'} \right)}{1-\exp(-M_{sm})} \right] \qquad (8)$$

In the fitting procedure, both h_s and K_s^{sp} are estimated. The steepness parameter is important, as the overall potential yield of a resource estimated by an ASPM depends primarily on the steepness of the stock-recruitment curve and on the natural mortality rate.

3.1.3 Total catch and catches-at-age

The fleet-disaggregated catch by mass for species *s*, in year *y* and region *r* is given by:

$$C_{srfy} = \sum_{a=0}^{m} w_{s,a+1/2}\, C_{srfya} = \sum_{a=0}^{m} w_{s,a+1/2}\, N_{sya}\, e^{-M_{sa}/2}\, S_{srfya}\, F_{srfy} \qquad (9)$$

where

$w_{s,a+1/2}$ denotes the mid-year mass of fish of species *s* and age *a*, which is assumed to be the same for each fleet (as there are no data available to discriminate between fleets),

C_{srfya} is the catch-at-age, i.e. the number of fish of species *s* and age *a*, caught in year *y* and region *r* by fleet *f*,

S_{srfya} is the commercial selectivity (i.e. vulnerability to fishing gear, which may depend not only on the gear itself, but also on distribution patterns of the fish by age compared to the areal distribution of fishing effort) of species *s* at age *a* for year *y*, region *r* and fleet *f*; when $S_{srfya} = 1$, the age-class *a* is said to be fully selected, and

F_{srfy} is the fished proportion of a fully selected age class of species *s*, for fleet *f* in region *r*.

Rather than having the annual catches as fixed input parameters (as it is in the 'standard' assessments), the fishing proportions of *M. capensis* in the west coast inshore region for each fleet ($F_{cap,wc-in,fy}$) are treated as estimable parameters in the fitting procedure. To obtain the fishing proportions of the other combinations of species and regions, it is assumed that the relative selectivities are related by a series of parameters, θ_{srfy}, so that:

$$F_{srfy} = \theta_{srfy} F_{cap,wc-in,fy} \qquad (10)$$

$$F_{cap,wc-off,fy} = \theta_{fy}^{2} F_{cap,wc-in,fy} \qquad (11)$$

$$F_{cap,sc-in,fy} = \theta_{fy}^{3} F_{cap,wc-in,fy} \qquad (12)$$

$$F_{par,wc-off,fy} = \theta_{fy}^{4} F_{cap,wc-off,fy} \qquad (13)$$

$$F_{par,sc-in,fy} = \theta_{fy}^{5} F_{cap,sc-in,fy} \qquad (14)$$

$$F_{cap,sc-off,fy} = \theta_{fy}^{6} F_{cap,sc-in,fy} \qquad (15)$$

$$F_{par,sc-off,fy} = \theta^7_{fy} F_{cap,sc-off,fy}$$ (16)

where θ_{srfy} is modelled by a random walk (once fishing has commenced in that region; before that θ=0):

$$\theta_{srf,y+1} = \theta_{srfy} e^{\zeta_{srfy}}$$ (11) (17)

with

ζ_{srfy} is $N\left(0,(\sigma_{srfy})^2\right)$

For years in which data are available to estimate the fishing mortality for a particular combination of species and region directly and with reasonable precision, σ_{srfy} can be fixed at a very large value, so that the likelihood contribution from equation (1711) plays little role.

The model estimate of the mid-year exploitable ("available") component of biomass for each species, fleet and region is calculated by converting the numbers-at-age into mid-year mass-at-age (using the mid-year individual weights) and applying natural and fishing mortality for half the year:

$$B^{ex}_{srfy} = \sum_{a=0}^{m_s} w_{s,a+1/2} S_{srfya} N_{sya} e^{-M_{sa}/2}(1 - \sum_f S_{srfya} F_{srfy} / 2)$$ (1812)

Total exploitable biomass for a fleet and region is taken as the sum of the *M. capensis* and *M. paradoxus* exploitable biomass for that fleet and region, i.e.:

$$B^{ex}_{BS,rfy} = B^{ex}_{cap,rfy} + B^{ex}_{par,rfy}$$ (19)

where the subscript '*BS*' stands for 'both species'.

The model estimate of the survey biomass at the start of the year (summer) for each species and region is given by:

$$B^{surv}_{sry} = \sum_{a=0}^{m_s} w_{sa} S^{surv}_{sra} N_{sya}$$ (2013)

and in mid-year (winter):

$$B^{surv}_{sry} = \sum_{a-0}^{m_s} w_{s,a+1/2} S^{surv}_{sra} N_{sya}\ e^{-M_{sa}/2}(1 - \sum_f S_{srfya} F_{srfy} / 2)$$ (2114)

where

S^{surv}_{sra} is the survey selectivity for age *a* for species *s* in region *r*, and

$w_{s,a+1/2}$ is the mid-year weight of fish of species *s* and age *a* at the start of the year.

It is assumed that the resource is at the deterministic equilibrium that corresponds to an absence of harvesting at the start of the initial year considered, i.e., $B^{sp}_{sy0} = K^{sp}_s$.

3.2 The likelihood function

The model is fitted to CPUE and survey abundance indices, catch information and commercial and survey catch-at-age data, as well as to the stock-recruitment curve to estimate model parameters. Contributions by each of these to the negative of the log-likelihood (- $\ell\text{n}L$) are as follows.

3.2.1 CPUE relative abundance data

The likelihood is calculated assuming that the observed abundance index is log-normally distributed about its expected value:

$$I_y^i = \hat{I}_y^i \exp\left(\varepsilon_y^i\right) \quad \text{or} \quad \varepsilon_y^i = \ell n\left(I_y^i\right) - \ell n\left(\hat{I}_y^i\right) \qquad (2215)$$

where

I_y^i is the abundance index for year y and series i (referring which corresponds to a combination of species, fleet and region – for example,)the ICSEAF CPUE series on the west coast is for the offshore fleet, for both species combined and for both the inshore and offshore regions combined; another example of a CPUE series would be the GLM-standardised CPUE series for the offshore fleet, for the offshore region of the south coast, for *M. paradoxus* only),

$\hat{I}_y^i = \hat{q}^i \hat{B}_{srfy}^{ex}$ is the corresponding model estimate, where $\hat{B}_{srfy}^{ex}$ is the model estimate of exploitable resource biomass, given by equation 1812,

$\hat{q}^i$ is the constant of proportionality for abundance series i, and

ε_y^i from $N\left(0, \left(\sigma_y^i\right)^2\right)$.

In cases where the CPUE series are based upon species-aggregated and region-aggregated (offshore and inshore for the west or south coasts) catches (as available pre-1978), the corresponding model estimate is derived as follow:

Combining equations 9 and 10, we get:

$$C_{srfy} = \theta_{srfy} F_{cap,wc-in,fy} \sum_{a=0}^{m} w_{s,a+1/2} N_{sya} e^{-M_a/2} S_{srfya} \qquad (16)$$

The total catch (C_{cfy}^{TOT}) (*M. capensis* and *M. paradoxus* combined) by fleet *f* in year *y* off coast *c* (inshore and offshore regions combined), can then be written as:

$$C_{cfy}^{TOT} = \sum_{s=1}^{2}\sum_{r=1}^{2} C_{srfy} = F_{cap,wc-in,fy} \sum_{s=1}^{2}\sum_{r=1}^{2}\left[\theta_{srfy} \sum_{a=0}^{m} w_{s,a+1/2} N_{sya} e^{-M_a/2} S_{srfya}\right] \qquad (17)$$

The effort of fleet *f* in region *r* fishing species *s*, in year *y* (E_{srfy}) is proportional to the corresponding fishing mortality:

$$E_{srfy} = \frac{1}{q_{srfy}} F_{srfy} \qquad (18)$$

Calibrating from when the disaggregation is known, we have:

$$q_{srf} = \lambda_{srf}\, q_{cap,wc-in,f} \qquad (19)$$

where λ_{srf} and $q_{cap,wc-in,f}$ are estimable parameters

so that the total effort (E_{cfy}^{TOT}) (to catch *M. capensis* and *M. paradoxus*) by fleet *f* in year *y* off coast *c* (inshore and offshore regions combined) can be written as:

$$E_{cfy}^{TOT} = \sum_{s=1}^{2}\sum_{r=1}^{2} \frac{1}{q_{srf}} F_{srfy} = F_{cap,wc-in,fy} \frac{1}{q_{cap,wc-in,f}} \sum_{s=1}^{2}\sum_{r=1}^{2} \frac{\theta_{srfy}}{\lambda_{srf}} \qquad (20)$$

The model estimate of the species- and region-aggregated abundance index for year *y* and series *i* is then written as:

Formatted
Formatted
Formatted
Formatted
Formatted
Formatted
Formatted
Formatted
Formatted
Formatted
Formatted
Formatted
Formatted
Formatted
Formatted
Formatted

$$\hat{I}_y^i = \left(C/E\right)^{TOT}_{cfy} = q_{cap,wc-in,f} \frac{\sum_{s=1}^{2}\sum_{r=1}^{2}\left[\theta_{srfy}\sum_{a=0}^{m} w_{s,a+1/2} N_{sya} e^{-M_a/2} S_{srfya}\right]}{\sum_{s=1}^{2}\sum_{r=1}^{2}\left[\theta_{srfy}/\lambda_{srf}\right]} \qquad (21)$$

To correct for possible negative bias in estimates of variance $\left(\sigma_y^i\right)$ and to avoid according unrealistically high precision (and so giving inappropriately high weight) to the CPUE data, a lower bound ($\left(\sigma_A^i\right)^2$) on each CPUE series is input to the assessment model.

The contribution of the CPUE data to the negative of the log-likelihood function (after removal of constants) is then given by:

$$-\ell n L^{CPUE} = \sum_i \sum_y \left\{ \ell n \sqrt{\left(\sigma_y^i\right)^2 + \left(\sigma_A^i\right)^2} + \left(\varepsilon_y^i\right)^2 / \left[2\left(\left(\sigma_y^i\right)^2 + \left(\sigma_A^i\right)^2\right)\right]\right\} \qquad (2322)$$

where

σ_y^i is the (minimum, when $\sigma_A^i = 0$) standard deviation of the residuals for the logarithms of index *i* in year *y*,

σ_A^i is the square root of the additional variance for abundance series *i*, which is an input value; alternatively, this can be used to as a means of specifying an effective lower bound for σ_y^i.

Homoscedasticity of residuals is usually assumed, so that $\sigma_y^i = \sigma^i$ is estimated in the fitting procedure by its maximum likelihood value:

$$\hat{\sigma}^i = \sqrt{1/n_i \sum_y \left(\ell n(I_y^i) - \ell n(\hat{I}_y^i)\right)^2 - \left(\sigma_A^i\right)^2} \qquad (2423)$$

where n_i n_i is the number of data points for abundance index *i*.

The catchability coefficient q^i for abundance index *i* is estimated by its maximum likelihood value, which in the more general case of heteroscedastic residuals, is given by:

$$\ln \hat{q}^i = \frac{\sum_y \left[1 \Big/ \left\{\left(\sigma_y^i\right)^2 + \left(\sigma_A^i\right)^2\right\}\right]\left(\ln I_y^i - \ln \hat{B}_{srfy}^{ex}\right)}{\sum_y \left[1 \Big/ \left\{\left(\sigma_y^i\right)^2 + \left(\sigma_A^i\right)^2\right\}\right]} \qquad (2524)$$

3.2.2 Survey abundance data

Data from the research surveys are treated as relative abundance indices in a similar manner to the CPUE series above, with survey selectivity function S_{sra}^{surv} replacing the commercial selectivity S_{srfya} (see equations 20 13 and 21 14 above). Account is also taken of the begin- or mid-year nature of the survey.

An estimate of sampling variance is available for most surveys and the associated σ_y^i is generally taken to be given by the corresponding survey CV. However, these estimates likely fail to include all sources of variability, and unrealistically high precision (low variance and hence high weight) could hence be accorded to these indices. The contribution of the survey data to the negative log-likelihood is of the same form as that of the CPUE abundance data (see equation 2322). The procedure adopted takes into account an additional variance in the same manner as for the CPUE abundance indices, but instead of being input, the additional variance $\left(\sigma_A\right)^2$ is treated as another estimable

Formatted

Formatted

parameter in the minimisation process. This procedure is carried out enforcing the constraint that $(\sigma_A)^2>0$, i.e. the overall variance cannot be less than its externally input component.

3.2.3 Commercial Catches

The contribution of the annual catch estimates to the negative log-likelihood ($-\ell nL^{catch}$) is calculated by assuming that the observed catch (C_{srfy}) is log-normally distributed about its expected value ($\hat{C}_{srfy}$):

$$C_{srfy} = \hat{C}_{srfy}\exp(\varepsilon_{srfy}) \quad \text{or} \quad \varepsilon_{srfy} = \ell n(C_{srfy}) - \ell n(\hat{C}_{srfy}) \qquad (2625)$$

where

ε_{srfy} from $N(0,(\sigma_{catch})^2)$, with $\sigma_{catch}=0.1$ an input parameter.

$-\ell nL^{catch}$ is then given (after removal of constants) by:

$$-\ell nL^{catch} = \sum_s\sum_r\sum_f\sum_y \ell n\sigma_{catch} + \frac{(\varepsilon_{srfy})^2}{2(\sigma_{catch})^2} \qquad (2726)$$

Generally species-disaggregated catches by region are relatively poorly estimated, and only for years from 1978 onwards, whereas species-combined catches are relatively well known. To allow for this in a simple way, we assume simply that:

$$\varepsilon_{BS,rfy} = \ell n\left(\sum_s C_{srfy}\right) - \ell n\left(\sum_s \hat{C}_{srfy}\right) \qquad (2827)$$

where

$\varepsilon_{BS,rfy}$ from $N(0,(\sigma_{catchBS})^2)$, with $\sigma_{catchBS}=0.01$ also an input parameter.

A further term is then added to the $-\ell nL^{catch}$:

$$-\ell nL^{catch} \to -\ell nL^{catch} + \sum_r\sum_f\sum_y \ell n\sigma_{catchBS} + \frac{(\varepsilon_{BSrfy})^2}{2(\sigma_{catchBS})^2} \qquad (2928)$$

3.2.4 Commercial catches-at-age

Catches-at-age cannot be disaggregated by species or by inshore and offshore region, the model is therefore fitted to the catches-at-age for both species and both inshore and offshore regions combined. The contribution of the catch-at-age data to the negative of the log-likelihood function when assuming an "adjusted" lognormal error distribution is given by:

$$-\ell n\, L^{age} = \sum_i\sum_y\sum_a \left[\ell n\left(\sigma^i_{com}/\sqrt{p_{iya}}\right) + p_{iya}\left(\ell n\, p_{iya} - \ell n\,\hat{p}_{iya}\right)^2 / 2\left(\sigma^i_{com}\right)^2\right] \qquad (3029)$$

where

the subscript '*i*' refers to a particular series of catch-at-age data which reflect a specific combination of fleet and coast (region???)region or coast.

$p_{iya} = \frac{C_{BS,ins+off,fya}}{\sum_{a'} C_{BS,ins+off,fya'}}$ is the observed proportion of fish (*M. capensis* and *M. paradoxus* combined) caught by fleet *f* in year *y* and region *r* that are of age *a*,

$\hat{p}_{iya} = \frac{\hat{C}_{BS,ins+off,fya}}{\sum_{a'} \hat{C}_{BS,ins+off,fya'}} = \frac{\sum_s \hat{C}_{s,ins,fya} + \hat{C}_{s,off,fya}}{\sum_{a'} \sum_s \left(\hat{C}_{s,ins,fya'} + \hat{C}_{s,off,fya'}\right)}$ is the model-predicted proportion of fish caught by fleet *f* in year *y* that are of age *a*, where:

$$\hat{C}_{srfya} = N_{sya}\, e^{-M_{sa}/2}\, S_{srfya}\, F_{srfya} \qquad (3130)$$

and

σ^i_{com} is the standard deviation associated with the catch-at-age data, which is estimated in the fitting procedure by:

$$\hat{\sigma}^i_{com} = \sqrt{\sum_y \sum_a p^i_{y,a} \left(\ln p^i_{y,a} - \ln \hat{p}^i_{y,a}\right)^2 / \sum_y \sum_a 1} \qquad (3231)$$

The log-normal error distribution underlying equation 30 29 is chosen on the grounds that (assuming no ageing error) variability is likely dominated by a combination of interannual variation in the distribution of fishing effort, and fluctuations (partly as a consequence of such variations) in selectivity-at-age, which suggests that the assumption of a constant coefficient of variation is appropriate. However, for ages poorly represented in the sample, sampling variability considerations must at some stage start to dominate the variance. To take this into account in a simple manner, motivated by multinomial distribution properties, Punt (pers. commn) advocates weighting by the expected observed proportions (as in equation 3029) so that undue importance is not attached to data based upon a few samples only.

Commercial catches-at-age are incorporated in the likelihood function using equation 30, for which the summation over age *a* is taken from age a_{minus} (considered as a minus group) to a_{plus} (a plus group). The ages for the minus- and plus-groups are chosen so that typically a few percent, but no more, of the fish sampled fall into these two groups.

3.2.5 Survey catches-at-age

The survey catches-at-age are incorporated into the negative of the log-likelihood in an analogous manner to the commercial catches-at-age, assuming an adjusted log-normal error distribution (equation 3029). In this case however, the data is available disaggregated by species and for the inshore and offshore regions separately.

$p^{surv}_{srya} = C^{surv}_{srya} / \sum_{a'} C^{surv}_{srya'}$ is the observed proportion of fish of species *s* and age *a* from survey *surv* in year *y* and region *r*,

$\hat{p}^{surv}_{srya}$ is the expected proportion of fish of species *s* and age *a* in year *y* and region *r* in the survey *surv*, given by:

$$\hat{p}^{surv}_{srya} = \frac{S^{surv}_{sra} N_{sya}}{\sum_{a'=0}^{m_s} S^{surv}_{sra'} N_{sya'}} \qquad (3332)$$

for begin-year (summer) surveys, or

$$\hat{p}^{surv}_{srya} = \frac{S^{surv}_{sra} N_{sya} \exp(-M_{sa}/2)\left(1 - \sum_f S_{srfya} F_{srfy}/2\right)}{\sum_{a'=0}^{m_s} S^{surv}_{sra'} N_{sya'} \exp(-M_{sa'}/2)\left(1 - \sum_f S_{srfya} F_{srfy}/2\right)} \qquad (3433)$$

for mid-year (winter) surveys.

3.2.6 Stock-recruitment function residuals

The stock-recruitment residuals are assumed to be log-normally distributed and serially correlated. Thus, the contribution of the recruitment residuals to the negative of the log-likelihood function is given by:

$$-\ell n L^{SR} = \sum_s \sum_{y=y1+1}^{y2} \left[\ell n \sigma_R + \left(\frac{\varsigma_{sy} - \rho \varsigma_{s,y-1}}{\sqrt{1-\rho^2}} \right)^2 \Big/ 2\sigma_R^2 \right] \qquad (3534)$$

where

$\varsigma_{sy} = \rho\varsigma_{s,y-1} + \sqrt{1-\rho^2}\,\varepsilon_{sy}$ is the recruitment residual for species *s*, and year *y*, which is estimated for year *y1* to *y2* (see equation 4),

ε_{sy} from $N\left(0,(\sigma_R)^2\right)$

σ_R is the standard deviation of the log-residuals, which is input, and

ρ is the serial correlation coefficient, which is input.

In the interest of simplicity, equation 35 omits a term in $\varsigma_{s,y1}$ for the case when serial correlation is assumed ($\rho \neq 0$), which is generally of little quantitative consequence to values estimated (Cryer, 1986).

3.2.7 Smoothness in changes over time in focus of fishing effort by region and species

The smoothness overtime desired of the θ_{fy}^k parameters (see equations 10-17) is effected by the addition of a further contribution to the negative log-likelihood function:

$$-\ell nL^{smooth} = \sum_{k=1}^{7}\sum_f\sum_y \ell n\tilde{\sigma}_{fy}^k + \frac{\left(\zeta_{fy}^k\right)^2}{2\left(\tilde{\sigma}_{fy}^k\right)^2}$$

(3635)

where the $\tilde{\sigma}_{fy}^k$ values will be input.

3.3 Model parameters

3.3.1 Estimable parameters

In addition to the species- and region-specific virgin spawning biomass $\left(K_s^{sp}\right)$ and the species-specific "steepness" of the stock-recruitment relationship (h_s), the following parameters are also estimated in some of the model fits undertaken.

Natural mortality:

Natural mortality (M_{sa}) is assumed either to be independent of age or age-specific, and input (fixed) or estimated using the following functional form in the latter case:

$$M_{sa} = \begin{cases} M_{s2} & \text{for} \quad a \le 1 \\ \alpha_s^M + \dfrac{\beta_s^M}{a+1} & \text{for} \quad a \ge 2 \end{cases} \qquad (3736)$$

M_{s0} and M_{s1} are set equal to M_{s2} ($=\alpha_s^M + \beta_s^M/3$) as there are no data (hake of ages younger than 2 are rare in catch and survey data) which would allow independent estimation of M_{s0} and M_{s1}.

Fishing selectivity-at-age:

The fishing selectivity-at-age for each species, region and fleet, S_{srfa}, is either estimated directly:

$$S_{srfa} = \begin{cases} \text{estimated separately} & \text{for } a \le a_{est} \\ =1 & \text{for } a > a_{est} \end{cases} \qquad (3837)$$

or in terms of a logistic curve given by:

$$S_{srfa} = \begin{cases} 0 & \text{for } a = 0 \\ \left[1+\exp\left(-\left(a-a_{srf}^c\right)/\delta_{srf}^c\right)\right]^{-1} & \text{for } a \ge 1 \end{cases} \qquad (3938)$$

where

a^c_{srf} years is the age-at-50% selectivity,

δ^c_{srf} year^{-1} defines the steepness of the ascending limb of the selectivity curve.

The selectivity is sometimes modified to include a decrease in selectivity at older ages, as follows:

$$S_{srfa} \rightarrow S_{srfa} \exp\left(-s_{srfa}\left(a - a_{slope}\right)\right) \qquad \text{for } a > a_{slope}, \qquad (4039)$$

where

s_{srfa} measures the rate of decrease in selectivity with age for fish older than a_{slope} for the fleet concerned, and is referred to as the "selectivity slope" in this thesis.

Time dependence may be incorporated into these specification, so that $S_{srfa} \rightarrow S_{srfya}$.

3.3.2 Input parameters

Age-at-maturity:

The proportion of fish of species *s* age *a* that are mature is approximated by

$$f_{sa} = \begin{cases} 0 & \text{for } a < a_s^{mat} \\ 1 & \text{for } a \geq a_s^{mat} \end{cases} \qquad (4140)$$

where $a_s^{mat} = 4$ for the *M. capensis* and *M. paradoxus* stocks (Punt and Leslie, 1991).

Weight-at-age:

The weight-at-age (begin and mid-year) for each species is calculated from the combination of the von Bertalanffy growth equation and the mass-at-length function.

4. References

Baranov F.T. 1918. On the question of the dynamics of the fishing industry. *Nauch. Issled. Ikhtiol. Inst. Izv*. **1**: 81-128 (In Russian).

BENEFIT. 2004. Formal report: BENEFIT/NRF Stock Assessment Workshop. University of Cape Town. 12-17 January 2004. BENEFIT Programme, Swakopmund. 48pp.

Beverton R.J.H. and Holt S.J. 1957. *On the dynamics of exploited fish populations.* Fisheries Investment Series 2, Vol. 19, U.K. Ministry of Agriculture and Fisheries, London, 533pp.

Cryer J.D. 1986. *Time series analysis*. Wadsworth Publishing, Boston, Massachusetts. 286pp.

Pope J.G. 19841972. The performance of short cut methods for catch forecasts. *ICES Document CM 1984/D:3*. 12ppAn investigation of the accuracy of virtual population analysis using cohort analysis. *Res. Bull. Int. Commn. Northwest Atlant. Fish.* **9**: 65-74.

Punt A.E. and Leslie R.W. 1991. Estimates of some biological parameters for the Cape hakes off the South African west coast. *S. Afr. J. mar. Sci*. **10**: 271-284.

Rademeyer R.A. 2003. Assessment of and Management Procedures for the Hake Stocks off Southern Africa. MSc thesis, University of Cape Town.

Formatted

Formatted