# Summary of the most recent South African hake assessments 

R.A. Rademeyer and D.S. Butterworth<br>MARAM (Marine Resource Assessment and Management Group)<br>Department of Mathematics and Applied Mathematics<br>University of Cape Town, Rondebosch 7701, South Africa

April 2006

## Introduction

In the most recent assessments (Rademeyer and Butterworth, 2006) of the South African hake resource, Merluccius paradoxus and M. capensis are treated as two separate stocks, but are assessed simultaneously within a single assessment framework and for the south and west coasts combined. This simultaneous assessment is necessary because much of the data is available in species-aggregated form only. Thus the model is one of two species and two spatial strata (see Fig. 1) with differences in the distributions by age within each stratum handled by allowing for stratum-specific commercial (in principle, though not in this particular implementation) and survey selectivities, rather than explicitly modelling movement. This follows the recommendation from the January 2004 BENEFIT/NRF/BCLME workshop (BENEFIT, 2004), though the further recommendation of that workshop to extend to four spatial strata (two by depth as well as two longshore) has yet to be implemented. The only data available which are explicitly disaggregated by species are those from research surveys that have taken place from 1986 to the present. However the framework does admit implicit disaggregation of data from the commercial fishery as summarised below.

Rather than providing advice based on a single Reference Case, a Reference Set (RS) has been developed. The Reference Set aims to take account of the factors that account for most of the uncertainty regarding the key considerations of resource status and productivity. Although this elaboration is intended primarily to serve the needs of OMP (Operational Management Procedure) development and testing, it has also been used recently for the purposes of "assessment-based" advice. This RS is constructed by including variations around four aspects of the assessment:

1) Natural mortality ( $M$ ): different (age-dependent) upper bounds are implemented ( 2 scenarios).
2) Different assumptions about the species split of the catches pre-1978 (see Fig. 2) (3 scenarios).
3) Steepness parameter ( $h$ ): Different upper bounds are implemented (4 scenarios).
4) Recent stock-recruitment residuals: one scenario with constant variability throughout the period, the other forcing recent recruitment closer to the stock recruitment curve (2 scenarios).

Each scenario is given equal weight, so that the RS consists of a total of 48 equally weighted components.
The reasons for this combination of options are as follows:

1) Best fits to the data yield estimates of natural mortality that are unrealistically large (particularly for older hake) so that these have been constrained not to exceed specified lower values.
2) Plausible assessments do provide some insight on the species split of the pre-1978 catches. If the change from a primarily M. capensis to primarily M. paradoxus fishery occurred much later than indicated in Fig. 2, results reflect an unrealistically large current $M$. capensis biomass, together with
an unrealistically low multiplicative bias for swept-area biomass estimates from the south coast survey.
3) The best fits generally yield estimates of $h$ for both species on the upper bound of 0.95 that is input, so that alternatives for lower $h$ are also included to better allow for the possibility of recruitment overfishing in providing management advice. The four scenarios are as follows: $\mathrm{H} 1: h$ values for both species are estimated in the minimisation process, $\mathrm{H} 2: h$ for $M$. paradoxus is fixed at 0.8 and estimated for M. capensis, H3: $h$ for M. capensis is fixed at 0.7 and estimated for M. paradoxus, H4: h is fixed at 0.8 and 0.7 for M. paradoxus and M. capensis respectively. A higher fixed $h(0.8)$ for $M$. paradoxus than M. capensis ( 0.7 ) has been chosen because of the smaller CV for $h$ for M. paradoxus compared to M. capensis in scenario H1.
4) Given large recent catches of small M. paradoxus, the assessments suggest very high recent recruitment (based on little data) which implies rapid rebuilding. An alternative involving (effectively) greater shrinkage of recent recruitments towards the mean suggested by the stockrecruitment function fitted is considered in case such large recruitment estimates are providing unduly optimistic projection results.

## Data

## Total catches

The South African hake stocks are fished by four fleets: the offshore trawl fleet and the longline fleet operate on both the south and west coasts, while the inshore trawl fleet and the handline fleet operate on the south coast only. Species-disaggregated catch series are external inputs to the model and are shown in Fig. 3. A summary of the assumptions made to disaggregate the catches by species and fleet is given below.

## a) Offshore trawl fleet:

## 1978-2004:

The catches made by the offshore trawl fleet have been split by species by applying the size-based species proportion-by-depth relationships, as estimated from research survey data for the west and south coasts, that have been developed by Gaylard and Bergh (2004). [The possibility that the research survey information provides biased estimates of proportions in the commercial catches has been addressed in separate sensitivity tests.]

## 1917-1977:

Prior to 1978, there is no depth information recorded for the landings so that the proportion of M. capensis caught cannot be estimated using the method above. Previously, the proportional split over the 1917-1977 period has been assumed to equal the average that pertained over the 1978-1982 in dividing the catches for these years. More recently, however, the catch data for the 1917-1977 period have been split by assuming that the proportion of M. capensis caught follows a logistic trend over this period, starting at 1 and then decreasing to stabilise at the 1978-1982 average value. Indeed, trawling was concentrated in inshore areas around Cape Town when the fishery began (i.e. probably catching $M$. capensis exclusively) and progressively moved offshore, so that this seems a more defensible approach. The proportion of M. capensis in the offshore trawl catch in year $y$ on coast $c$ is thus given by:

$$
\begin{equation*}
\text { prop }_{c y}^{\text {off }}=\frac{1-\Delta_{c}}{1+\exp \left[\left(y-P_{1}\right) / P_{2}\right]}+\Delta_{c} \tag{1}
\end{equation*}
$$

where
$\Delta_{c} \quad$ is the average proportion of M. capensis in the offshore catch over the 1978-1982 period for coast $c$ ( $24 \%$ and $60 \%$ on the west and south coasts respectively for scenarios C3a-c), and
$P_{1}, P_{2}$ are parameters of the logistic function: $P_{1}$ is the year in which the proportion of $M$. capensis in the catch is half-way between $100 \%$ and $\Delta_{c}$; while $P_{2}$ defines how rapidly this change in proportion occurs.

The following scenarios have been included in the Reference Set:

C3a: $P_{1}=1950$ and $P_{2}=1.5$;
C3b: $P_{1}=1940$ and $P_{2}=1.5$;
C3c: $P_{1}=1957$ and $P_{2}=1.5$.
The proportion of $M$. capensis consequently assumed for the offshore trawl catches for scenarios C3a-c is shown in Fig. 2 for the west and for the south coasts.
b) Inshore trawl fleet:

The inshore trawl fleet operates on the south coast only. Catches made by this fleet are assumed to consist of M. capensis only, as it operates in relatively shallow water.

Because fleet-disaggregated catch data are not available prior to 1974 , the assumption has been made that the annual catch of the inshore trawl fleet from 1960 to 1973 increased linearly from 1000 t to 5000t, and that the balance of the total catch recorded was taken by the offshore trawl fleet.
c) Longline fleet:

Longline catches on the west coast are assumed to consist of $30 \%$ M. capensis for the whole period that longlining has been in operation (from 1983), while on the south coast catches by this fleet are assumed to consist of $M$. capensis exclusively.

## d) Handline fleet:

The handline fleet operates on the south coast only. As for the inshore fleet, catches made by this fleet are assumed to consist of M. capensis only.

The overall catch in 2004 is taken to be the TAC for that year, with the same proportion of each species as caught by each fleet in 2003 assumed.

## Abundance indices

Historic (ICSEAF) (1955-1977 on the west coast and 1969-1977 on the south coast) and GLM-standardised (post-1978) CPUE data are available for the hake offshore trawl fishery. The historic CPUE series cannot be disaggregated by species, as there are no effort-by-depth data available for this pre-1978 period, and so is fitted to an appropriate combination of M. capensis and M. paradoxus exploitable biomass (see Appendix A).The GLM-standardized CPUE series are species-specific indices (and based also on the new Gaylard and Bergh estimated species-proportion vs. depth relationships).

Survey biomass estimates for the west (summer and winter) and south (spring and autumn) coasts are available for M. paradoxus and M. capensis separately. The winter and spring surveys have been occasional, but the summer west coast and autumn south coast surveys virtually annual since 1988 and 1985 respectively (except for four and six missing years respectively). However for 2000-2001, the summer surveys were carried out by the Nansen instead of the Africana, so that appropriate adjustments to allow for different survey selectivity functions for these two vessels are made in the assessment.

## Catches-at-age

Survey catch-at-age data are available for M. paradoxus and M. capensis separately.
Commercial catches-at-age are available for the offshore (1975-1996) (both coasts combined) and the inshore (1989-2000) and longline (1994-1997, 2000) (south coast only) fleets. They cannot be split by species on an age-basis, but for this assessment this is not a problem for the south coast inshore and longline fleets as their catches are assumed to consist of M. capensis only in the region in question. For the offshore trawl catches, the model fits available data to predicted catches-at-age for the two species together.

Annual age-length keys are used throughout, but are averages over sex and species.

## Key Assumptions

The model used in the South African hake assessment is an Age-Structured Production Model (ASPM). It now includes a new method introduced to model the historic (pre-1978) ICSEAF CPUE series which are available in terms of species-aggregated catches only - see Appendix A.

A summary of the specifications for the Reference Set assessments is given below.
a) Plus-group:

Age 15 is used as the plus-group for both species. Though recent catches reflect few fish of more than 7 years of age, this is necessary to take proper account of the greater weights-at-age of the older hake present during the initial years of the fishery.

## b) Natural mortality:

$M_{a}$ is taken to be age-dependent $\left(M_{a}\right)$ - see eqn 2 below. Upper bounds of 0.5 and 0.3 (scenario M1) and 1.0 and 0.5 (scenario M4) for ages 2 and 5 respectively are implemented. As there are not enough data to inform on the natural mortalities at ages above 5 for $M$. paradoxus and above 7 for M. capensis, the natural mortalities estimated for age 5 for M. paradoxus and age 7 for $M$. capensis are assumed to apply to older ages as well.
$M_{s a}=\left\{\begin{array}{ccc}M_{s 2} & \text { for } & a \leq 1 \\ \alpha_{s}^{M}+\frac{\beta_{s}^{M}}{a+1} & \text { for } & a \geq 2\end{array}\right.$
where $s$ signifies the species, and $\mathrm{M}_{\mathrm{s} 0}$ and $\mathrm{M}_{\mathrm{s} 1}$ are set equal to $\mathrm{M}_{\mathrm{s} 2}$ as there are no data (hake of ages younger than 2 are rare in catch and survey data) which would allow independent estimation of $\mathrm{M}_{\mathrm{s} 0}$ and $\mathrm{M}_{\mathrm{s} 1}$.

## c) Commercial selectivity-at-age:

The commercial selectivities take the form of logistic curves. Selectivity at low ages in the commercial catches has changed over time, likely due to the phasing out of net liners. To take account of this change, periods of fixed and changing selectivity have been assumed for the offshore fleet.
The offshore trawl selectivity for M. paradoxus is assumed to decrease exponentially from age 3 (with a slope parameter estimated in the model fitting procedure), while for M. capensis the offshore selectivity is assumed flat for older ages as these are assumed to be fully available to this fleet. This selectivity applies for the whole region (i.e. for west and south coasts combined). For the inshore fleet on the south coast, the selectivity is allowed to decrease exponentially from age 5, as this fleet does not fully select older fish because the distribution of hake extends deeper than its area of operation. Because the longline fishery targets principally older fish, the selectivity for that fleet on the south coast is also assumed to be flat for older ages. There are no catch-at-age data available to estimate a selectivity vector for the handline fleet, so the assumption is made that the selectivity for this fleet (assumed to catch M. capensis only) is intermediate between the inshore trawl and longline selectivities.
In all cases, the exponential decrease (or constancy) is assumed to continue from age 5 for M. paradoxus and age 7 for $M$. capensis to age $15+$.

## d) Survey selectivity-at-age

Survey selectivities are estimated directly for each age. A separate selectivity is assumed for each species, season and coast combination for which data are available.

An exponential decrease in selectivity is assumed from age 5 for M. paradoxus with the slope parameter fixed at 0.5 , and from age 7 for $M$. capensis with the slope parameter fixed at 1.0 (these values have been roughly computed from previous assessments) - note that "slope" in this context means:

$$
S_{a+1}=S_{a} e^{-s l o p e}
$$

## e) Stock-recruitment residuals

The residuals are estimated from year 1985 to 2004 (the lack of age data before that time precludes estimation of earlier values in a MLE context). The variability level (standard deviation of the logged residuals) $\sigma_{R}$ is fixed at 0.25 . This is low compared to other similar stocks worldwide, but the age data in this instance do not suggest high levels of variability.

## f) Age-at-maturity

For both species, $100 \%$ of fish of age 4 and above are assumed to be mature and to contribute fully to the spawning biomass and hence recruitment, and none below this.

## Key Results

The overall average and range of estimates of management quantities for the Reference Set are shown in Table 1. Fig. 4 plots the corresponding biomass trajectories, focusing on the median, maximum and minimum values across the scenarios for each year. Fig. 5 shows the survey and commercial fishing selectivities. In these Table and Figures, results are also shown for one specific case of the RS ("Case 21 ": higher natural mortality (M4), C3c historic species split option, best estimates (high) of steepness (H1) and "unrestricted" estimates of recent recruitment (SR1)). This case was chosen for presentation because it has the highest likelihood.

The current status of the M. paradoxus resource in terms of spawning biomass is estimated to be low, at around $10 \%$ of the pristine level, while M. capensis is estimated to be in a relatively good state, above the estimated MSYL.

Figs 6 and 7 show the fits of Case 21 to the CPUE and survey indices respectively, while Figs 8 and 9 show the fit of this model to the commercial and survey catch-at-age data. Fig. 10 shows the stock-recruitment residuals and stock-recruitment curve estimated for each species.

## Key Concerns

The aspects of greatest concern in these assessments are:
i) the high natural mortality, particularly at large age, which unconstrained fits to the model suggest;
ii) the low levels of recruitment variability estimated ( $\sigma_{R}$-output of 0.26 and 0.15 for M. paradoxus and $M$. capensis respectively for case 21 ).

Both of these results are biologically questionable, particularly in the light of comparisons with stocks of similar species elsewhere in the world. They are consequences of the catch-at-age data input, and may reflect systematic bias in ageing, or biases introduced by aggregating over species and sex in developing age-length keys to apply to catch-at-length distributions.

## References

BENEFIT 2004. Formal report: BENEFIT/NRF/BCLME Stock Assessment Workshop 2004. 12-17 January 2004. University of Cape Town. South Africa.

Gaylard JD and Bergh MO. 2004. A species splitting mechanism for application to the commercial hake catch data 1978 to 2003. Unpublished report, MCM, South Africa. WG/09/04/D:H:21.
Glazer JP. Incorporating the bias in the species splitting algorithm applied to the offshore hake catches. Unpublished report, MCM, South Africa. WG/03/05/D:H:13.

Rademeyer RA and Butterworth DS. 2006. Detailed methodology and results for the final reference set of the South African Merluccius paradoxus and M. capensis resources for use in OMP testing. Unpublished report, MCM, South Africa. WG/02/06/D:H:5.

Table 1: Average and range (in parenthesis) of estimates of management quantities of the M. paradoxus and M. capensis coast-combined resources over the 48 scenarios of the Reference Set. Results are also shown for case 21 of the Reference Set (see text for details). $M S Y$ and associated quantities are given in relation to the selectivity for the offshore fleet. The multiplicative bias estimate $(q)$ is shown for $M$. capensis for the sweptarea estimates from south coast autumn survey only, as this has proved important in excluding certain historic catch species split scenarios which have been viewed as unrealistic because they lead to implausibly low estimates for this $q$.



Fig 1: Demarcation of the "old" and "new" boundaries separating the west and south coasts in the hake fishery assessments (from Glazer, 2005). The new boundary applies to the analyses summarised in this paper.

C3 variants


Fig. 2: Assumed proportion of M. capensis in the offshore catches for the west coast and south coast for the catch variants C3a, b and c (see text for details).


Fig. 3: Catch series assumed in the most recent South African hake assessments.


Fig. 4: Trajectories of resource abundance for the Reference Set. Resource abundance is expressed in terms of a) spawning biomass, b) spawning biomass as a proportion of its pre-exploitation level, c) exploitable biomass (to which offshore trawl CPUE is assumed to be proportional) and d) biomass of fish of age 2 and above. The median is indicated by a thick line while the shaded area represents the full uncertainty encompassed by the Reference Set (minimum and maximum for each year). The dashed lines plot the biomasses for case 21 of the Reference Set (see text for details).

Survey selectivities


Commercial selectivities


Inshore





Fig. 5: Estimated survey and commercial fishing selectivities for the Reference Set. The median is indicated by a thick line while the shaded area represents the full uncertainty encompassed by the Reference Set (minimum and maximum for each age). The dashed lines plot the selectivities for case 21 of the Reference Set (see text for details).


Fig. 6: Fits to the CPUE abundance indices of Case 21 of the Reference Set.


Fig. 7: Fits to the survey abundance indices of Case 21 of the Reference Set.




Fig. 8: Case 21 model fits to commercial catch-at-age data (averaged over all the years for which data are available).


Fig. 9: Case 21 model fits to survey catch-at-age data (averaged over all the years for which data are available).


Fig. 10: Stock-recruit residuals (in relation to $\ln$ recruitment) and estimated stock-recruitment curves (together with recruitments as estimated from 1985, before which values from the curves are taken to apply) for M. paradoxus and M. capensis for Case 21 of the Reference Set.

## Appendix A - Fitting to species combined CPUE

In cases where the CPUE series are based upon species-aggregated catches (as available pre-1978), the corresponding model estimate is derived by assuming two types of fishing zones: z1) an "M. capensis only zone", corresponding to the shallow water and z2) a "mixed zone" (see Fig. A1 below).

| Zone 1 (z1): | Zone 2 (z2): |
| :---: | :---: |
| M. capensis only | Mixed zone |
| M. capensis: | M. capensis: |
| biomass $\left(B_{C z 1}\right)$, catch $\left(C_{C z 1}\right)$ | biomass $\left(B_{C z}\right)$, catch $\left(C_{C z 2}\right)$ |
|  | M. paradoxus: |
| biomass $\left(B_{P}\right)$, catch $\left(C_{P}\right)$ |  |
| Effort in zone $1\left(E_{z 1}\right)$ | Effort in zone $1\left(E_{z 2}\right)$ |

Fig. A1: Diagrammatic representation of the two theoretical fishing zones.

The total catch of hake of both species $(B S)$ by fleet $f$ in year $y\left(C_{B S, f y}\right)$ can be written as
$C_{B S, f y}=C_{C z 1, f y}+C_{C z 2, f y}+C_{P, f y}$, where
$C_{C z 1, f y}$ is the M. capensis catch by fleet $f$ in year $y$ in the $M$. capensis only zone,
$C_{C z 2, f y}$ is the M. capensis catch by fleet $f$ in year $y$ in the mixed zone, and
$C_{P, f y}$ is the M. paradoxus catch by fleet $f$ in year $y$ in the mixed zone.
Catch rate is assumed to be proportional to exploitable biomass. Furthermore, let $\gamma$ be the proportion of the M. capensis exploitable biomass in the mixed zone $\left(\gamma=B_{C z 2, f y}^{e x} / B_{C, f y}^{e x}\right)$ (assumed to be constant throughout the period) and $s_{f y}$ be the proportion of the effort of fleet $f$ in the mixed zone in year $y\left(s_{f y}=E_{f y}^{z 2} / E_{f y}\right)$, so that:

$$
\begin{align*}
& C_{C z 1, f y}=q_{C z 1}^{i} B_{C z 1, f y}^{e x} E_{f y}^{z 1}=q_{C z 1}^{i}(1-\gamma) B_{C, f y}^{e x}\left(1-s_{f y}\right) E_{f y}  \tag{A1}\\
& C_{C z 2, f y}=q_{C z 2}^{i} B_{C z 2, f y}^{e x} E_{f y}^{z 2}=q_{C z 2}^{i} \gamma B_{C, f y}^{e x} s_{f y} E_{f y} \quad \text { and }  \tag{A2}\\
& C_{P, f y}=q_{P}^{i} B_{P, f y}^{e x} E_{f y}^{z 2}=q_{P}^{i} B_{P, f y}^{e x} s_{f y} E_{f y} \tag{A3}
\end{align*}
$$

where
$q_{s}^{i} \quad$ is the constant of proportionality for abundance series $i$ and species $s$,
$E_{f y}=E_{f y}^{z 1}+E_{f y}^{z 2}$ is the total effort of fleet $f$, corresponding to combined-species CPUE series $i$ which consists of the effort in the $M$. capensis only zone $\left(E_{f y}^{z 1}\right)$ and the effort in the mixed zone $\left(E_{f y}^{z 2}\right)$.

It follows that:

$$
\begin{align*}
& C_{C, f y}=B_{C, f y}^{e x} E_{f y}\left[q_{C z 1}^{i}(1-\gamma)\left(1-s_{f y}\right)+q_{C z 2}^{i} s_{f y}\right]  \tag{A4}\\
& C_{P, f y}=B_{P, f y}^{e x} E_{f y} q_{P}^{i} s_{f y} \tag{A5}
\end{align*}
$$

By solving equations A4 and A5, it becomes evident that:
$s_{f y}=\frac{q_{C z 1}^{i}(1-\gamma)}{\left\{\frac{C_{C f y} B_{P f y}^{e x} q_{P}^{i}}{B_{C f y}^{e x} C_{P f y}}-q_{C_{z 2}}^{i} \gamma+q_{C z 1}^{i}(1-\gamma)\right\}}$
so that:
$\hat{I}_{y}^{i}=\frac{C_{f y}}{E_{f y}}=\frac{C_{f y} B_{P f}^{e x} q_{P}^{i} S_{f y}}{C_{P f y}}$

To correct for possible negative bias in estimates of standard deviation $\left(\sigma_{y}^{i}\right)$ and to avoid according unrealistically high precision (and so giving inappropriately high weight) to the CPUE data, lower bounds ( $\sigma_{A}^{i}$ ) on the standard deviations of the residuals for the logarithm of the CPUE series have been enforced for all such series considered in the population model fit; for the historic ICSEAF CPUE series (separate west coast and south coast series) $\sigma_{A}^{i}$ is set to 0 , as already $\sigma^{\text {ICSEAF }} \geq 0.25$.

The contribution of the CPUE data to the negative of the log-likelihood function (after removal of constants) is then given by:
$-\ln L^{\text {CPUE }}=\sum_{i} \sum_{y}\left\{\ln \sqrt{\left(\sigma_{y}^{i}\right)^{2}+\left(\sigma_{A}^{i}\right)^{2}}+\left(\varepsilon_{y}^{i}\right)^{2} /\left[2\left(\left(\sigma_{y}^{i}\right)^{2}+\left(\sigma_{A}^{i}\right)^{2}\right)\right]\right\}$
where
$\sigma_{y}^{i}$ is the (minimum, when $\sigma_{A}^{i}=0$ ) standard deviation of the residuals for the logarithms of index $i$ in year $y$,
$\sigma_{A}^{i}$ 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 $\sigma_{y}^{i}$.

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

$$
\begin{equation*}
\hat{\sigma}^{i}=\sqrt{1 / n_{i} \sum_{y}\left(\ln \left(I_{y}^{i}\right)-\ln \left(\hat{I}_{y}^{i}\right)\right)^{2}-\left(\sigma_{A}^{i}\right)^{2}} \tag{A9}
\end{equation*}
$$

where $n_{i}$ is the number of data points for abundance index $i$.
In the case of the species-combined CPUE, $q_{C Z 1}^{i}, q_{C Z 2}^{i}, q_{P}^{i}$ and $\gamma$ are directly estimated in the fitting procedure.
Two species-aggregated CPUE indices are available: the ICSEAF west coast and the ICSEAF south coast series. For consistency, $q$ 's for each species (and zone) are forced to be in the same proportion:
$q_{s}^{S C}=r q_{s}^{W C}$

