

# A Gender-Disaggregated Assessment of the South African Hake Resource, Fitting Directly to Age-Length Keys: a New Reference Case

Rebecca A Rademeyer and Doug S Butterworth

November 2009

## SUMMARY

The assessment methodology for the hake resource is refined to take account of recommendations made by the Panel at the December 2008 International Stock Assessment Workshop to move to a gender-disaggregated model, and to fit directly to age data and estimate growth curve parameters directly in the likelihood maximisation process. This approach proves to remove the conflict between fits to catch-at-age and catch-at-length distributions evident in earlier assessments. The current *M. paradoxus* female spawning biomass is estimated to be at 76% of its MSY level, while the corresponding *M. capensis* estimate well above this level.

## INTRODUCTION

This paper is a response in particular to the recommendation of the Panel at the December 2008 International Stock Assessment Workshop regarding the hake resource that “A sex-structured population dynamics model should be fit to the conditional age-at-length data (age-length keys) and length-frequency data (by sex when such data are available). The growth curves (and the variation in length-at-age) should be estimated within the assessment.” The recommendation arose in large part from the need to resolve a conflict between catch-at-age and catch-at-length data, and growth curves input to then current “New Baseline” assessment.

The reason for the gender-disaggregation is that there are very clear gender-specific differences in somatic growth for both *M. paradoxus* and *M. capensis*, in fact more so than between species (see Fig. 4). Routine application of age-length keys to obtain catch-at-age proportions has been conducted without attention to gender-specific differences, but gender-differential growth means that larger sized males are not well represented in the catch. This could confound estimates based on catch-at-age data developed from a gender-aggregated age-length key, which might consequently under-represent the number of older hake present (and therefore affect estimates of natural mortality as well as distort estimates of year-class strength – the comparatively low variability of previous estimates of the latter has been a puzzle and concern, which might be resolved by gender differentiation of the assessments). Furthermore there is a sex-imbalance in certain components of the fishery: for example Geromont *et al.* (1995) estimated a female proportion in the south coast longline catches of 83%.

In the light of these considerations, the assessment of the hake resource has now been refined to incorporate gender- as well as species-differentiation. Thus in this analysis, the genders are modelled separately. As recommended above, the model is also fit directly to age-length keys (ALKs) and length frequencies (as e.g. in Punt *et al.* 2006), rather than to the age frequency information which multiplying the two would provide. There are three reasons for this:

- a) ALKs are not available for all years and surveys or fisheries, so that length distribution data have to be fitted directly in those cases.
- b) The fishery selectivity is essentially length- rather than age-specific; age-specific selectivities as assumed when fitting to age-distribution data will lead to mis-fitting of length distribution data in these circumstances (e.g. the lower tails of the length distributions of younger fish are not present in catches, but an age-specific selectivity requires them to be).

- c) The feature of the data described in b) leads to a bias in the estimation of hake growth curves if estimated directly from hake age data, leading to the lengths at younger ages being positively biased; growth curve parameters need to be estimated within the assessment to correct for this bias.

This paper reports the results from this refined assessment, which is termed the “New Reference Case” in the expectation that it provides a basis from which to develop the Operating Models for the hake resource which will provide the basis for simulation testing of a revised OMP for hake due for adoption in the latter part of 2010.

## DATA and METHODS

Appendix I details the data used in this analysis, while the specifications and equations of the overall model are set out in Appendix II.

## RESULTS

Estimates of management quantities for the new Reference Case are given in Table 1, while Fig. 1 plots the spawning biomass trajectories. The spawning biomass trajectories for *M. paradoxus* show a clear gender difference, with the current male depletion estimated to be at 10% compared to 33% for the females. For *M. capensis*, both male and female spawning biomasses are estimated to be at about 50% of pre-exploitation levels. For *M. capensis* the female spawning biomass is estimated to be well above its MSY level, whereas the corresponding component of the *M. paradoxus* population is estimated at 76% of that level.

The estimated commercial and survey selectivities are shown in Figs 2 and 3 respectively. Apart for *M. paradoxus* on the South Coast, the male and female selectivities at length are assumed to be the same and are then converted to gender specific selectivities at age. Because of selectivity difference between males and females apparent in the South Coast surveys for *M. paradoxus*, gender-specific selectivities are estimated for this species in the South Coast autumn and spring surveys, with the female selectivities scaled down for these two surveys (for *M. paradoxus* only) by a factor estimated in the model fitting procedure. This gender difference is assumed to affect the commercial fleet as well and the female selectivity for the South Coast offshore trawl fleet (the only fleet assumed to catch *M. paradoxus* on the south coast) is therefore also scaled down by a factor estimated in the model fitting. The female selectivity scaling factors estimated for the South Coast spring and autumn surveys and the offshore trawl fleet are 0.13, 0.27 and 0.10 respectively.

All the commercial selectivities show a decrease for large fish. This decrease is estimated for all fleets apart from the offshore trawl and south coast handline fleets for *M. capensis*. For the *M. capensis* offshore trawl fleet, the selectivity slope is fixed to 1/3 of the inshore trawl fleet slope estimated, while for the handline fleet, the selectivity slope is taken as the average of the estimated longline and inshore fleet slopes.

Fig. 4 plots the gender-specific growth curves estimated in the model, as well as the estimated length-at-age distributions. The difference between male and female growth curves is estimated to be more important for *M. paradoxus* than for *M. capensis*. *M. capensis* is estimated to grow larger than *M. paradoxus*.

Fig. 5 plots the estimated stock-recruitment relationships, and the time series of residuals about these relationships for both species. The extent of residual variability indicated for these plots (the output  $\sigma_R$  values) remain low compared to the norm for populations of similar demersal species.

Figs 6 and 7 show the fits the CPUE and survey abundance series. The fits are good for all series.

The fits to the commercial catch-at-length data are shown in Fig. 8. The fits as averaged over the years for which data are available are good for all the data sets. There are however some patterns evident in the bubble plots of residuals which could perhaps be improved by having further periods between which selectivities change.

The fits to the survey gender-aggregated and gender-disaggregated catch-at-age data are shown in Figs 9 and 10 respectively. These fits are also relatively good. In particular, the problem of the lack of fit evident in the corresponding plots for the “New Baseline” assessment (Rademeyer and Butterworth 2009) appears to be resolved.

The fit of the model to the ALKs is shown in Fig. 11. The observed and predicted ALKs are compared aggregated over all ALKs, first summed over ages, which should be exactly equal by construction and then summed over lengths.

The maturity-at-age ogive used in the new Reference Case is shown in Fig. 12. Since resource depletion estimates can be sensitive to the manner in which spawning biomass is defined, Table 2 compares the estimated 2009 depletions for this ogive with those for fish of age 3+ and of age 4+ (definitions used for earlier assessments). Results for the ogive are very similar to those for fish of age 4+.

## REFERENCES

- BENEFIT. 2004. Formal report: BENEFIT/NRF stock assessment workshop, Cape Town, 12-17 January 2004.
- Beverton RJH and Holt SJ. 1957. *On the dynamics of exploited fish populations*. Fisheries Investment Series 2, Vol. 19, U.K. Ministry of Agriculture and Fisheries, London. 533pp.
- Brandão A Rademeyer RA and Butterworth DS. 2004. First attempt to obtain a multiplicative bias calibration factor between the *Africana* with the old and the new gear. Unpublished report, Marine and Coastal Management. WG/11/04/D:H:26. 2pp
- Fairweather T. 2008. Length-weight relationship for *Merluccius capensis* and *M. paradoxus* based on research survey biological data. Unpublished report, Marine and Coastal Management. MCM/2008/SWG-DEM/38. 5pp.
- Fairweather T and Leslie RW. 2008. *Merluccius capensis* and *M. paradoxus* length at 50% maturity based on research survey biological data. Unpublished report, Marine and Coastal Management. MCM/2008/AUG/SWG-DEM/42. 5pp
- Fairweather T. 2009. Updated abundance estimates for *Merluccius capensis* & *M. paradoxus*. Unpublished report, Marine and Coastal Management. MCM/2009/JUN/SWG-DEM/47. 4pp.
- Fairweather T, Butterworth DS, Rademeyer RA and Leslie RW. 2009. Computing proportions at length (and by sex) from catches in a stratum. Unpublished report, Marine and Coastal Management. MCM/2009/OCT/SWG-DEM/75. 2pp
- Gaylard JD and Bergh MO. 2004. A species splitting mechanism for application to the commercial hake catch data 1978 to 2003. Unpublished report, Marine and Coastal Management. WG/09/04/D:H:21. 8pp.
- Glazer JP and Butterworth DS. 2009. Results from sensitivity tests related to the models used to standardize the offshore commercial trawl CPUE data of *M. capensis* and *M. paradoxus* respectively. Unpublished report, Marine and Coastal Management. MCM/2009/NOV/SWG-DEM/95.
- Punt AE, Smith DC, Tuck GN and Methot RD. 2006. Including discard data in fisheries stock assessments: Two case studies from south-eastern Australia. Fisheries Research 79: 239-250.
- Rademeyer RA. 2009. Age-reading error matrices for *Merluccius paradoxus* and *M. capensis*. Unpublished report, Marine and Coastal Management. MCM/2009/OCTOBER/SWG-DEM/73.
- Rademeyer RA and Butterworth DS. 2009. Routine update of the 2008 baseline assessment for the South African hake resource. Unpublished report, Marine and Coastal Management. MCM/2009/NOV/SWG-DEM/92

Table 1: Estimates of management quantities. The SS component of the  $-\ln L$  contribution excludes the  $-\ln \sigma$  term, i.e. this component would be zero if the data matched the model estimates exactly. Note: MSY and related quantities have been calculated assuming a fishing pattern that is the average over the last 5 years.

	Both	SS	
-lnL total	80.5		
CPUE historic	-40.7	3.7	
CPUE GLM	-164.3	55.6	
Survey	-32.7	56.2	
Commercial CAL	-54.9		
Survey CAL (sex-aggr.)	-6.3		
Survey CAL (sex-disaggr.)	23.9	5538.9	
ALK	319.0		
Recruitment penalty	20.7		
Selectivity smoothing penalty	15.6		
	Both	Males	Females
<i>M. paradoxus</i>			
$K^{sp}$	1535	649	886
$h$	0.75		
$B^{sp}_{2009}$	353	62	292
$B^{sp}_{2009}/K^{sp}$	0.23	0.10	0.33
$B^{sp}_{MSY}$	466		
$MSYL^{sp}$	0.30		
$B^{sp}_{2009}/B^{sp}_{MSY}$	0.76		
$MSY$	97		
<i>M. capensis</i>			
$K^{sp}$	784	377	407
$h$	0.98		
$B^{sp}_{2009}$	387	185	202
$B^{sp}_{2009}/K^{sp}$	0.49	0.49	0.50
$B^{sp}_{MSY}$	183		
$MSYL^{sp}$	0.23		
$B^{sp}_{2009}/B^{sp}_{MSY}$	2.11		
$MSY$	73		
2009 species ratio $B^{sp}$	1.10	2.98	0.69

Table 2: Estimated 2009 depletion for the new Reference Case spawning biomass, 3+ biomass and 4+ biomass for *M. paradoxus* and *M. capensis*.

	<i>M. paradoxus</i>	<i>M. capensis</i>
$B^{sp}_{2009}/K^{sp}$	0.23	0.49
$B^{3+}_{2009}/K^{3+}$	0.31	0.57
$B^{4+}_{2009}/K^{4+}$	0.25	0.50

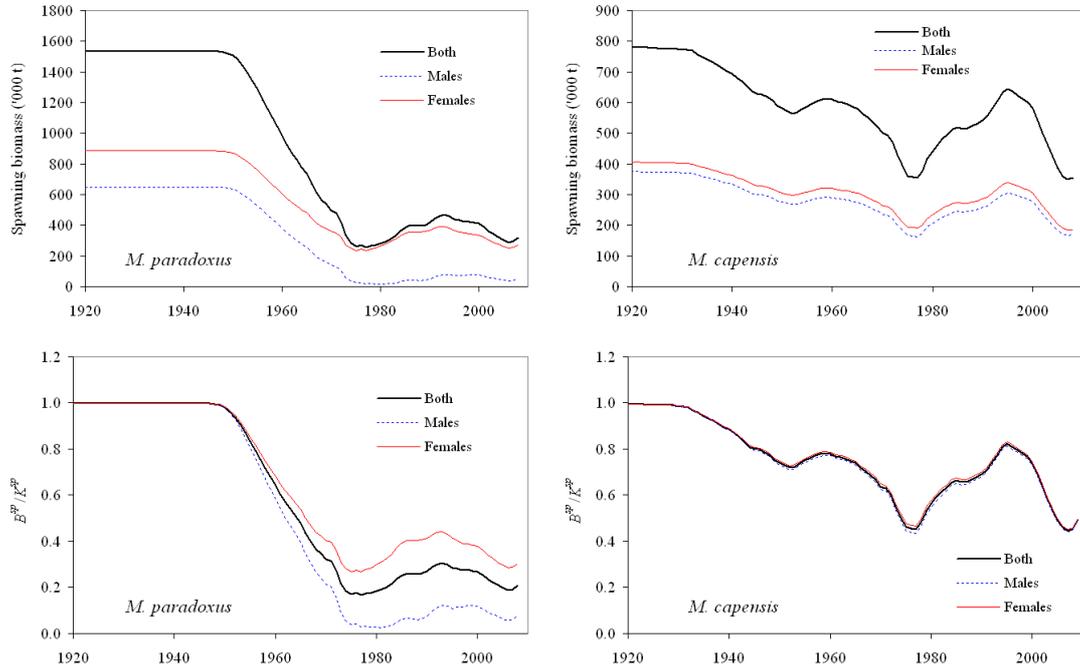


Fig. 1: Estimated spawning biomass trajectories for males and females *M. paradoxus* and *M. capensis*, both in absolute terms and relative to the pre-exploitation level.

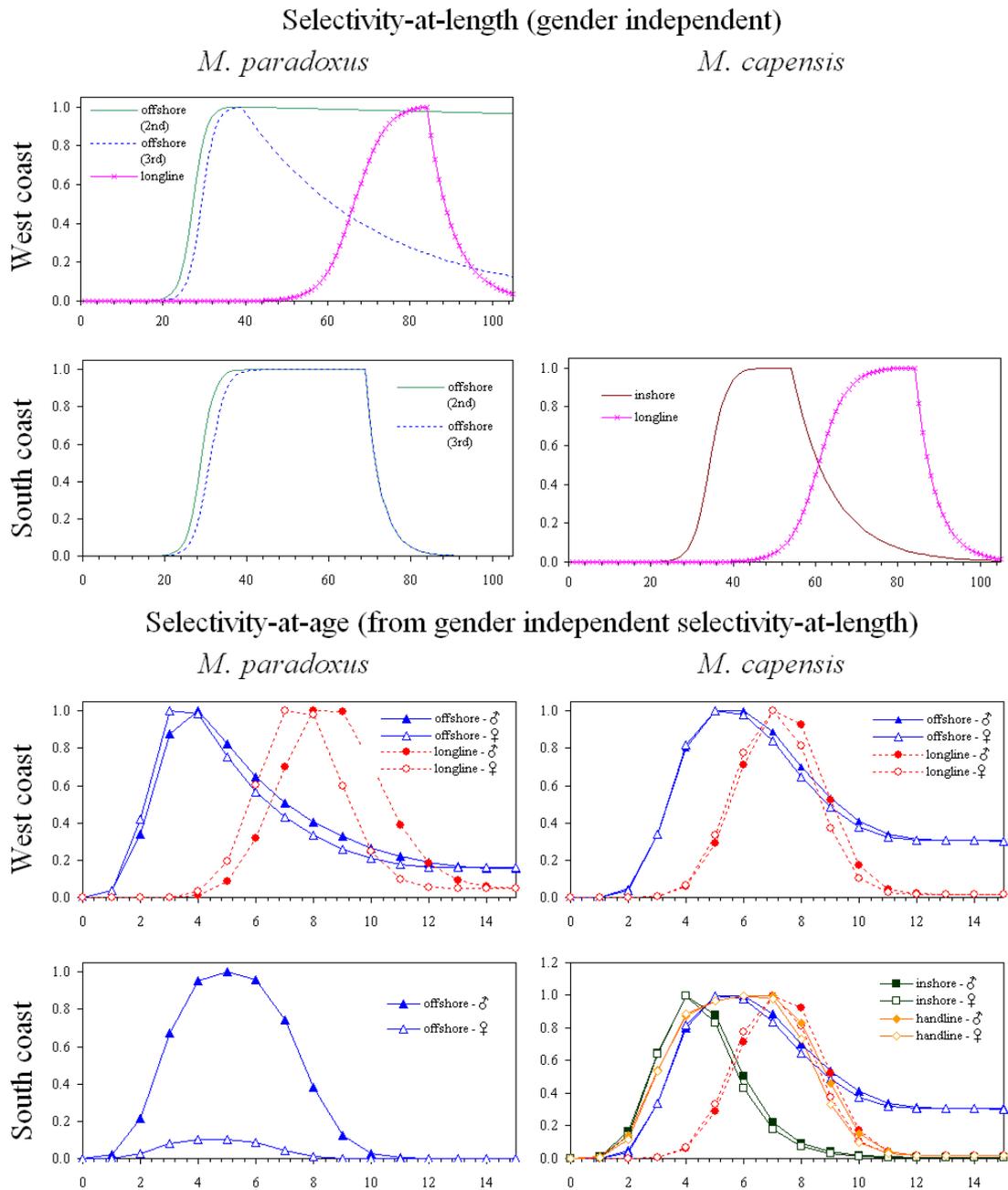


Fig. 2: Commercial gender-independent selectivities-at-length estimated directly in the model-fitting (the other selectivities-at-length (west coast: *M. paradoxus* offshore 1st period, (1917-1976) and *M. capensis* offshore and longline; south coast: *M. capensis* offshore and handline) are based on various assumptions (see text)) and commercial gender-dependent selectivities-at-age that follow from those. Note that because of space constraints, the offshore and longline legends have been omitted for the south coast *M. capensis* selectivity-at-age plots. The legends for these two fleets are as for those for the west coast.

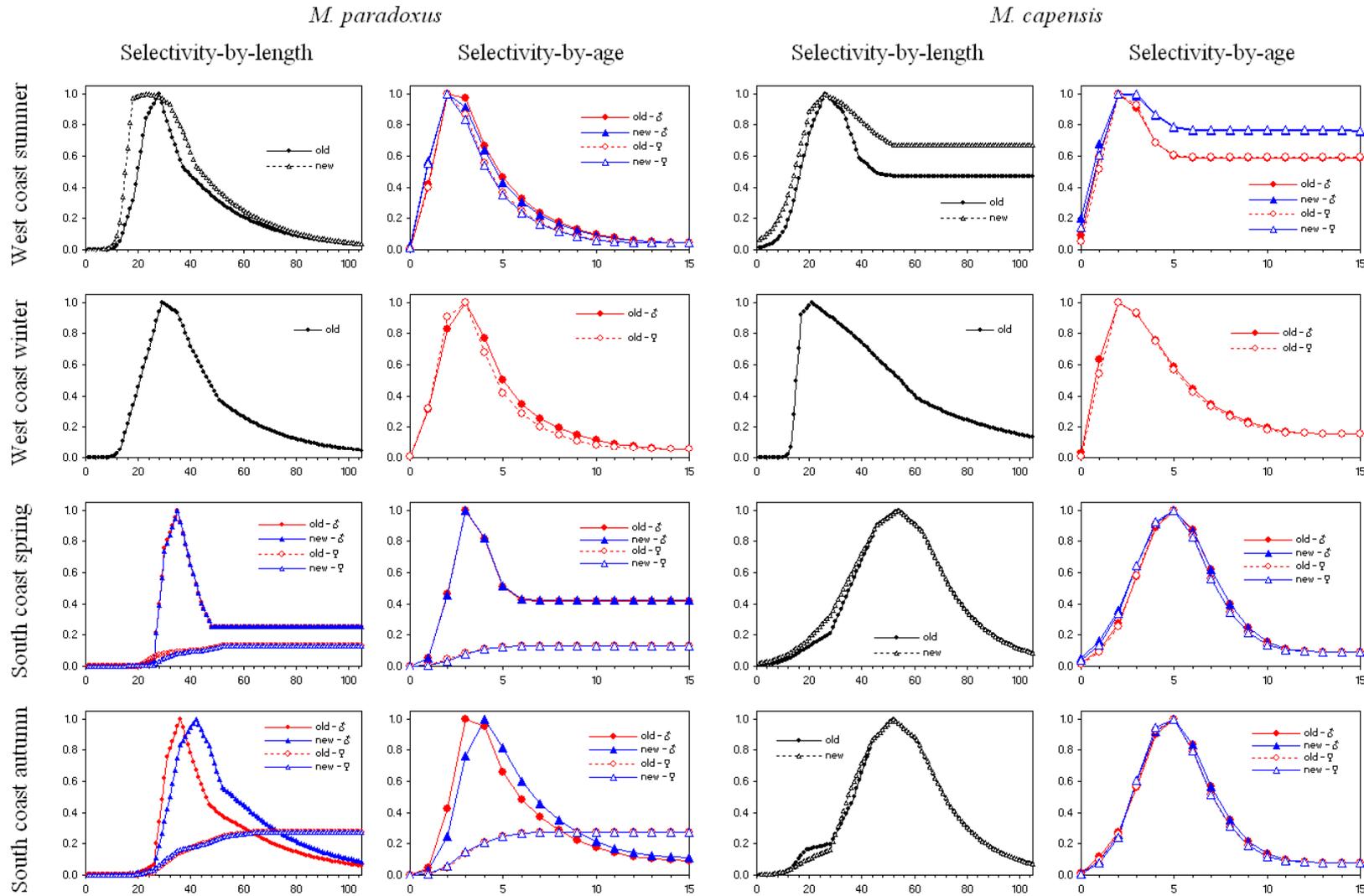


Fig. 3: Survey selectivity-at-length (estimated) for the 'old' and 'new' Africana gear and derived gender-dependent survey selectivity-at-age for each of the four surveys for *M. paradoxus* and *M. capensis*. Note that for the south coast spring and autumn surveys, gender-specific selectivities have been estimated for *M. paradoxus*. For all the other combinations of species and surveys, the selectivities-at-length are gender independent.

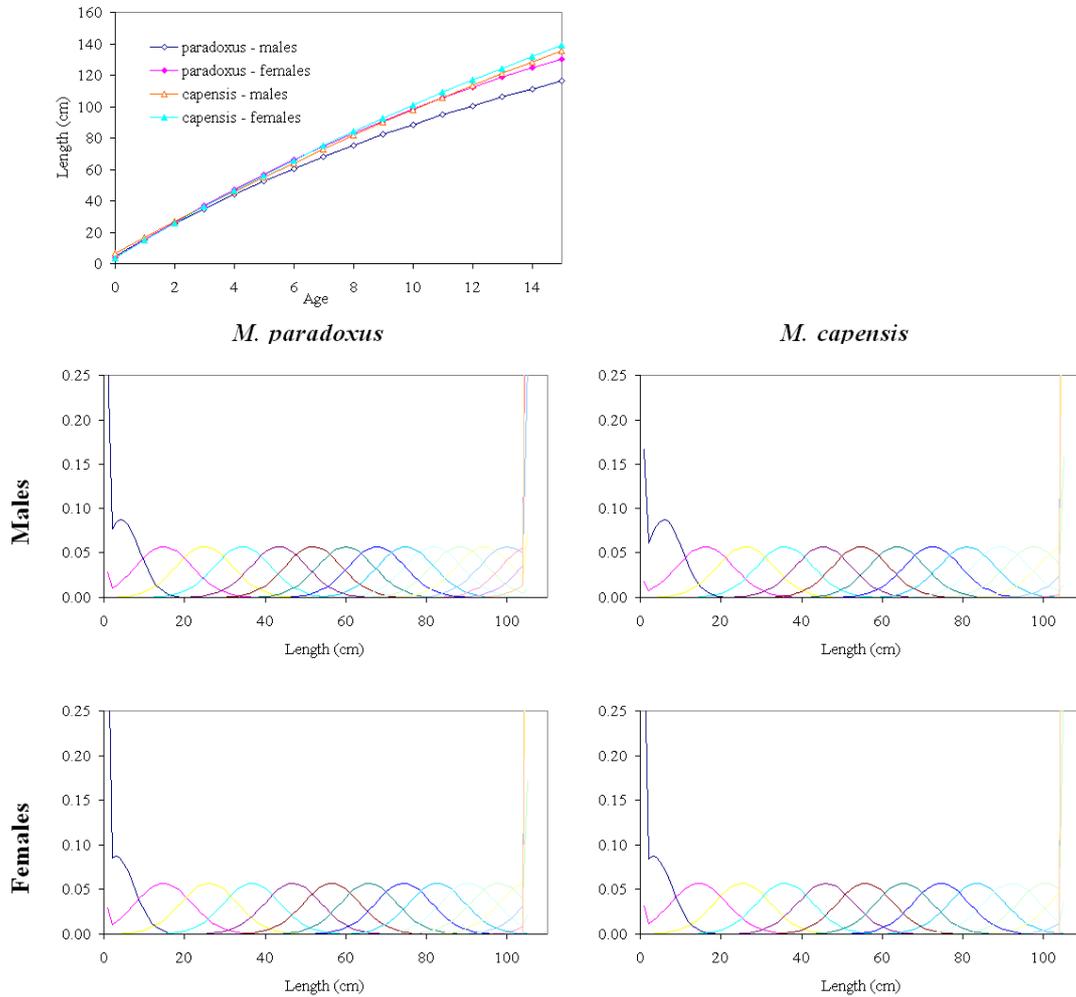


Fig. 4: Estimated length-at-age relationship and resulting length-at-age distributions for males and females *M. paradoxus* and *M. capensis*. In the lower plots the distributions, starting from the left, correspond to ages 0, 1, 2, ... The 0-year old group has a distribution overlapping zero, which is accumulated into a minus-group of 1 cm length; though such a model for this length distribution is clearly unrealistic, this hardly matters as in implementation the lowest minus-group considered for length is 10 cm, so that implied structure below that length is ignored.

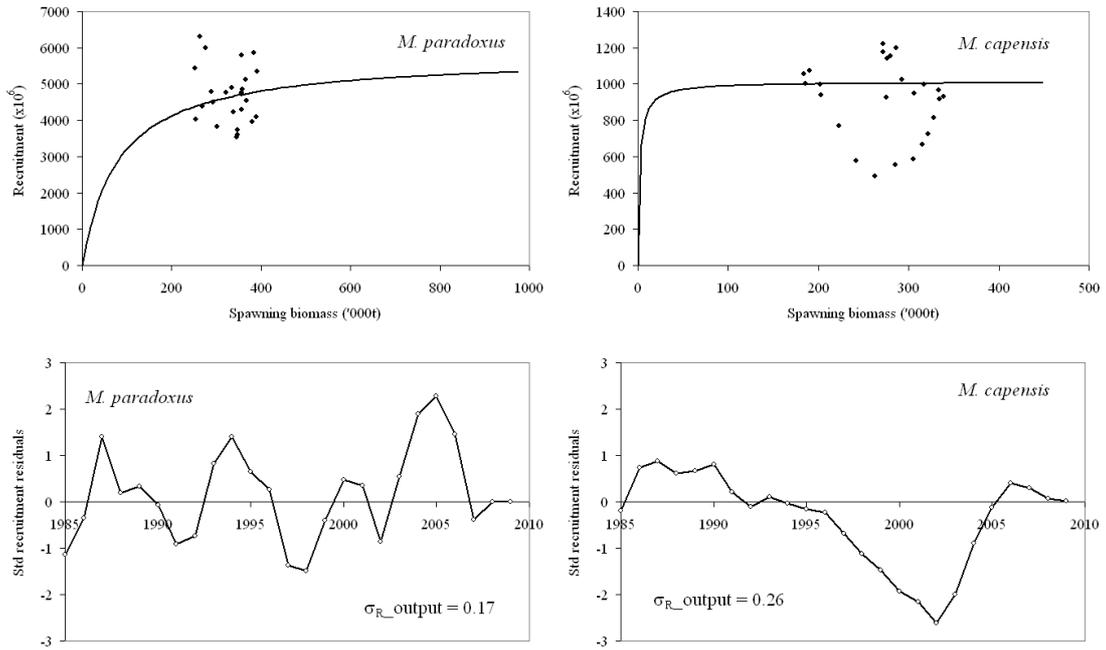


Fig.5: Estimated stock-recruitment relationships and time series of standardised stock-recruitment residuals. The output  $\sigma_R$  values are also given.

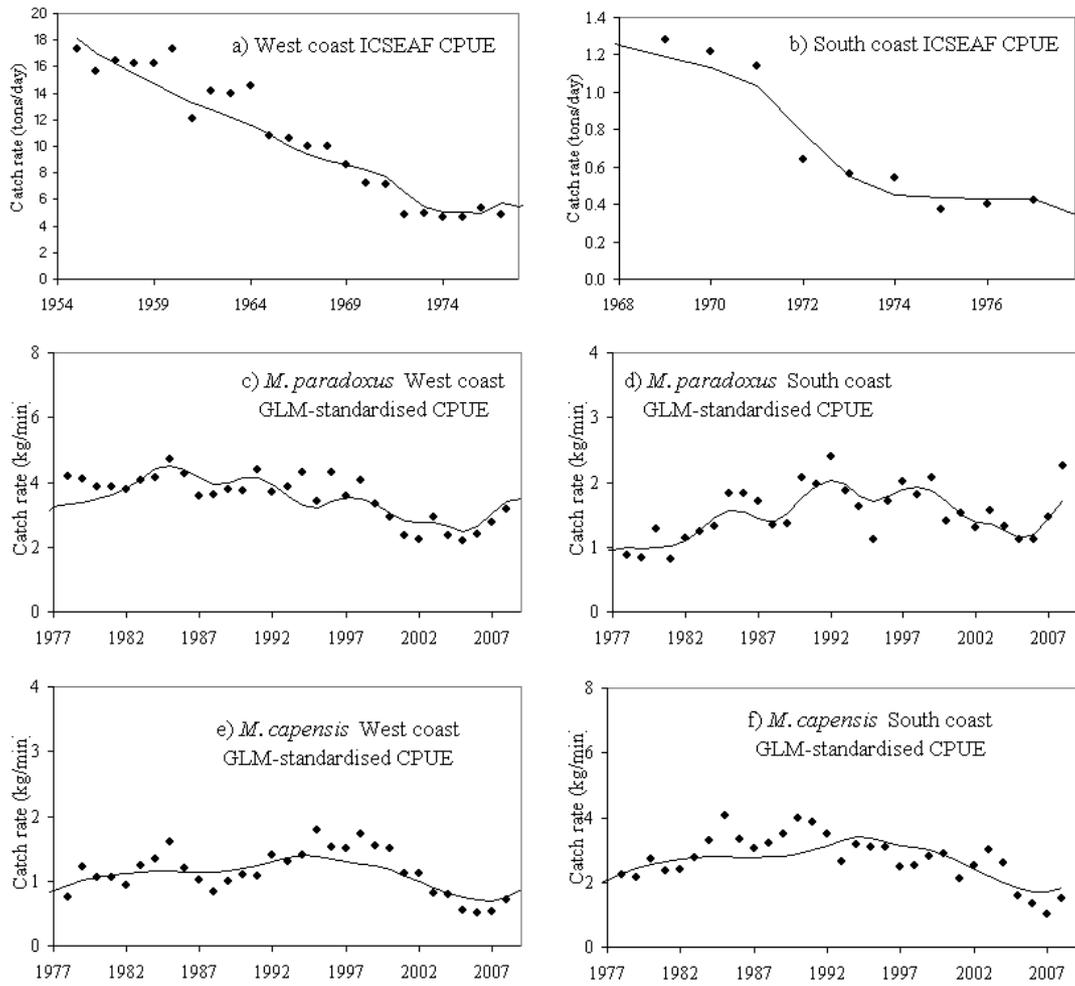


Fig.6: Fit of the model to the CPUE data.

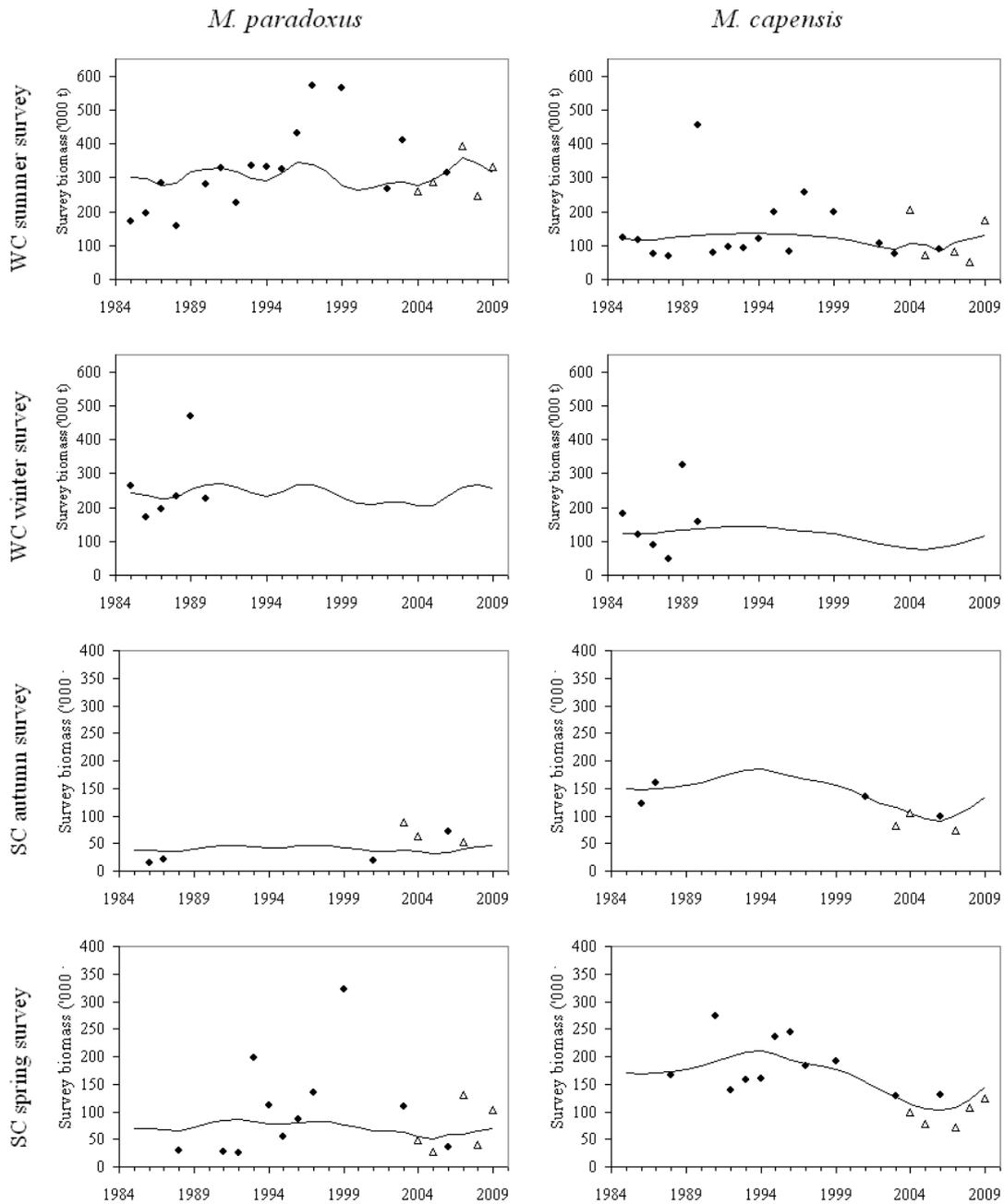


Fig.7: Fit of the model to the survey abundance indices. The triangles represent surveys that have been conducted with the new gear on the *Africana*. These are rescaled by the calibration factor for the species concerned to make them comparable to the other survey results.

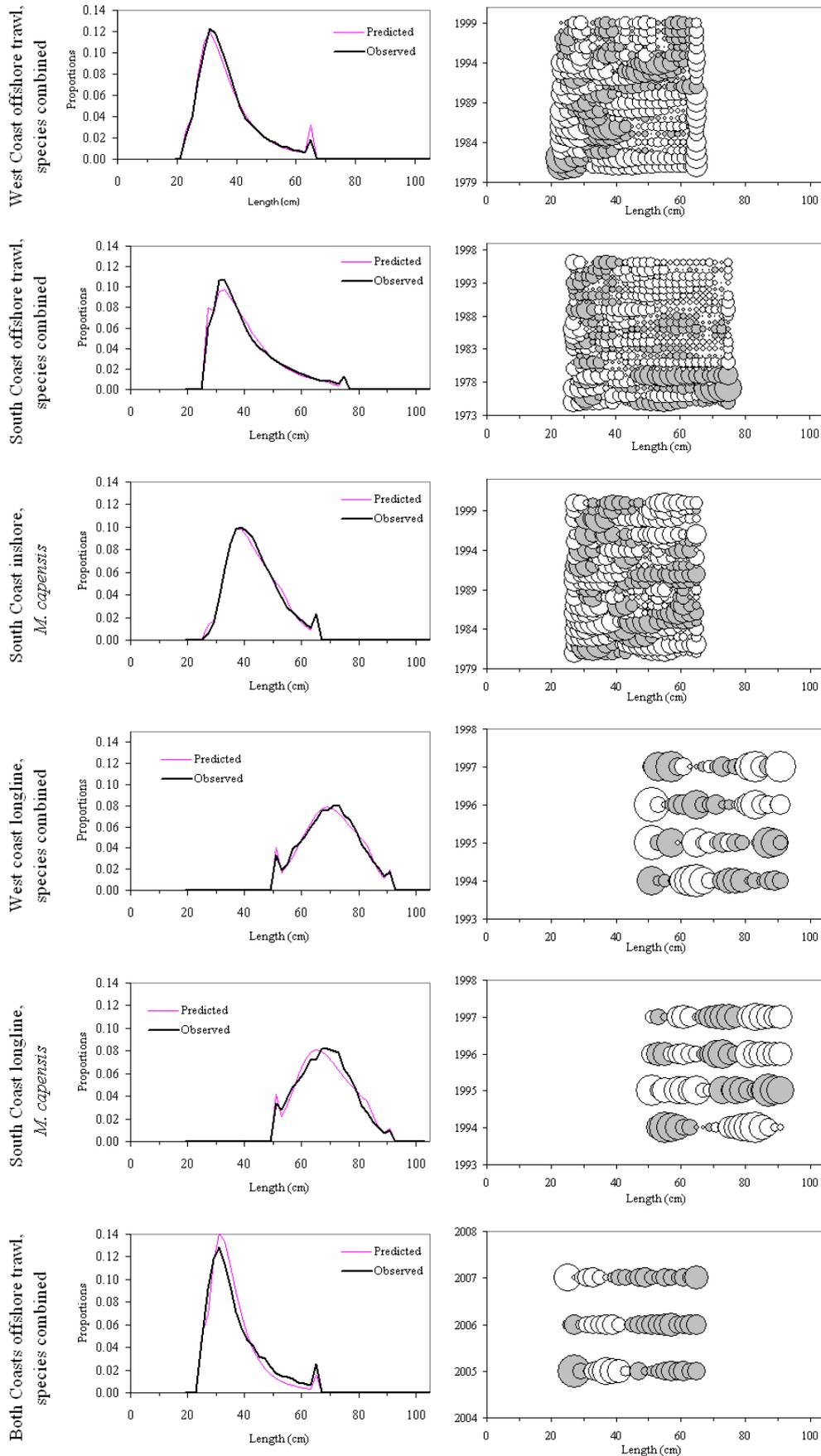


Fig.8: Fit of the model to the commercial proportion-at-length data, aggregated over years for which data are available for the plots on the left.

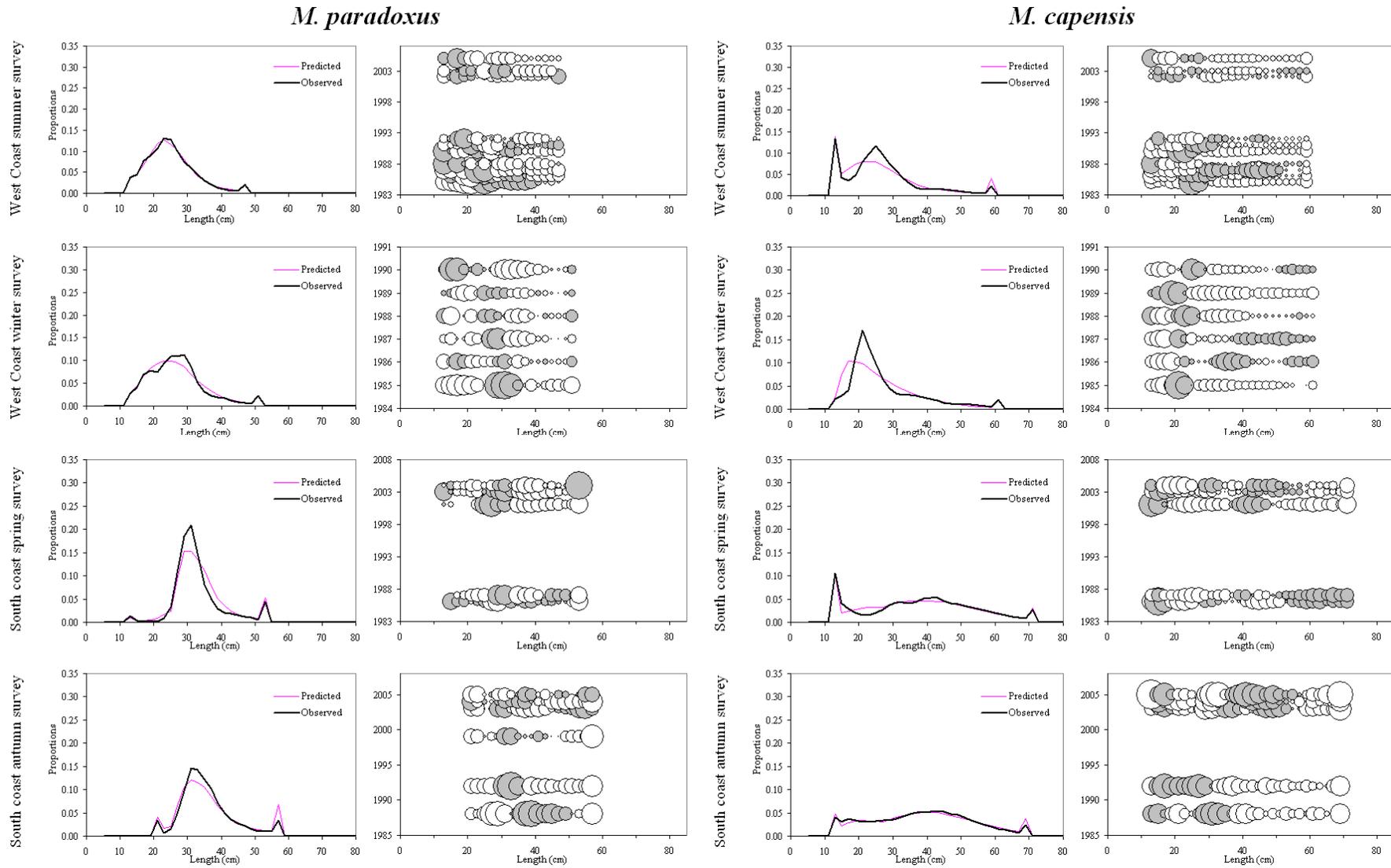


Fig.9: Fit of the model to the survey gender-aggregated surveys proportion-at-length data (in some plots, aggregated over years for which data are available).

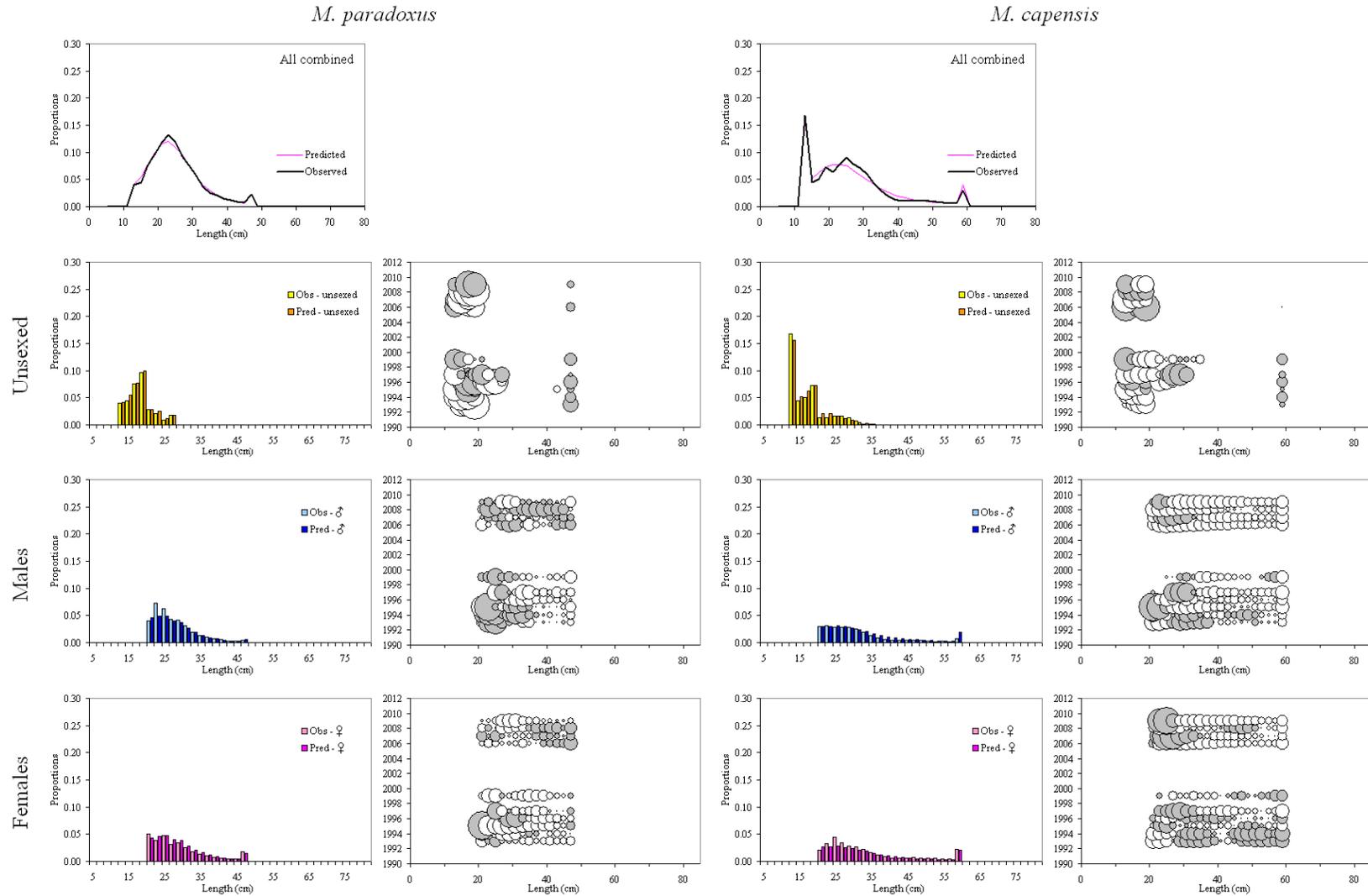


Fig. 10a: Fit of the model to the west coast summer survey gender- disaggregated proportion-at-length data (in some plots, aggregated over years for which data are available).

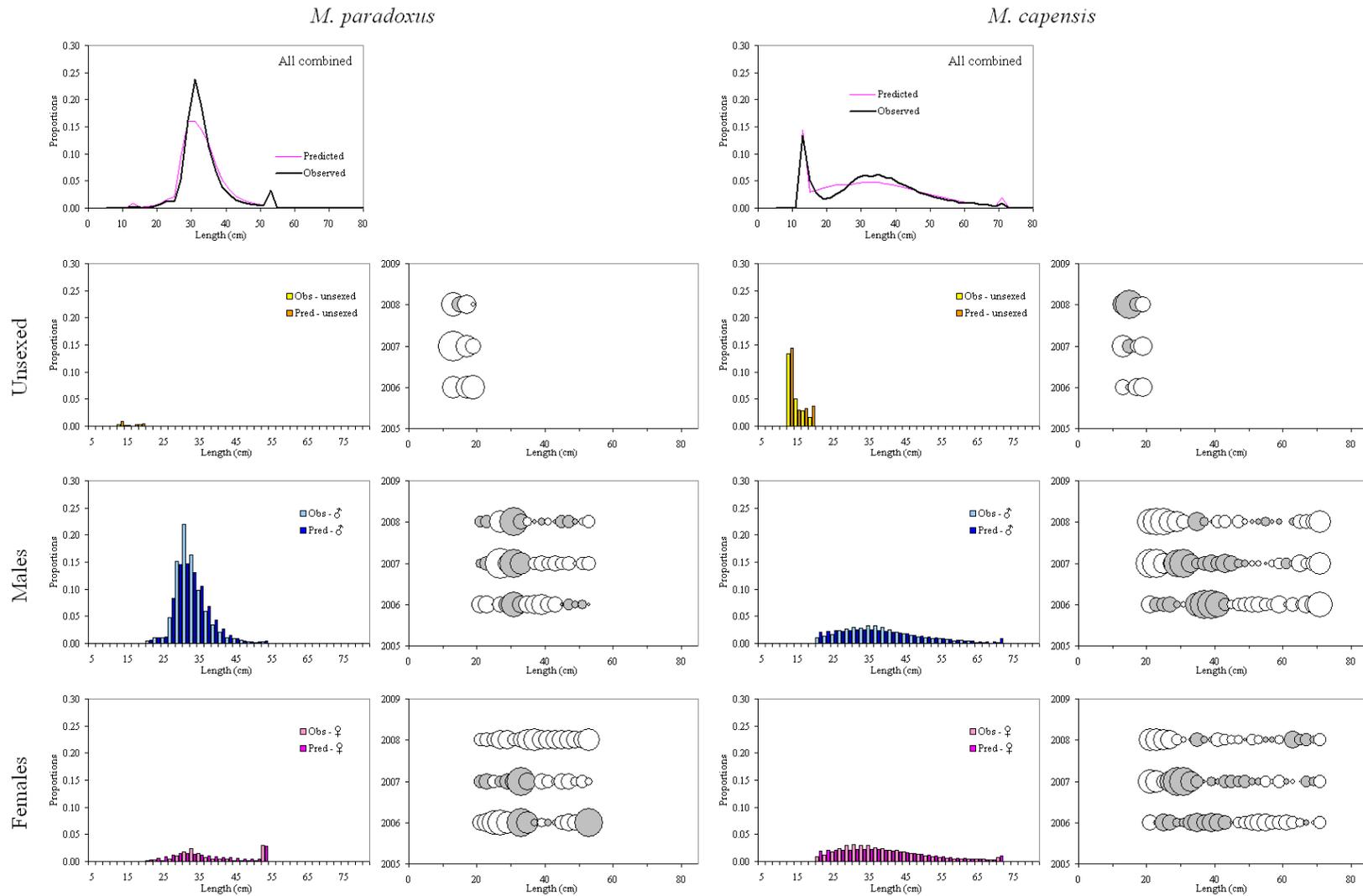


Fig. 10b: Fit of the model to the south coast spring survey gender- disaggregated proportion-at-length data (in some plots, aggregated over years for which data are available).

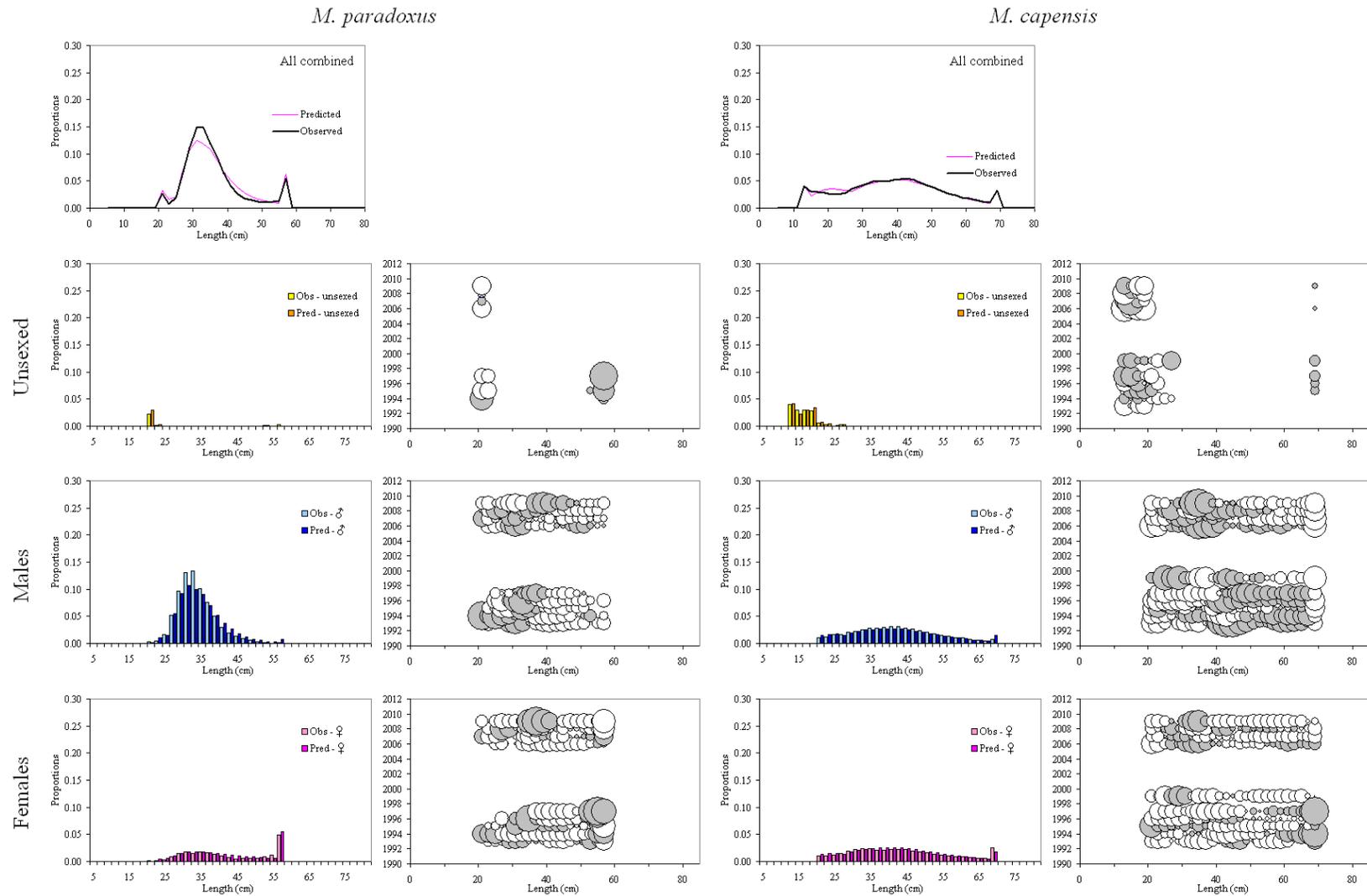


Fig. 10c: Fit of the model to the south coast autumn survey gender- disaggregated proportion-at-length data (in some plots, aggregated over years for which data are available).

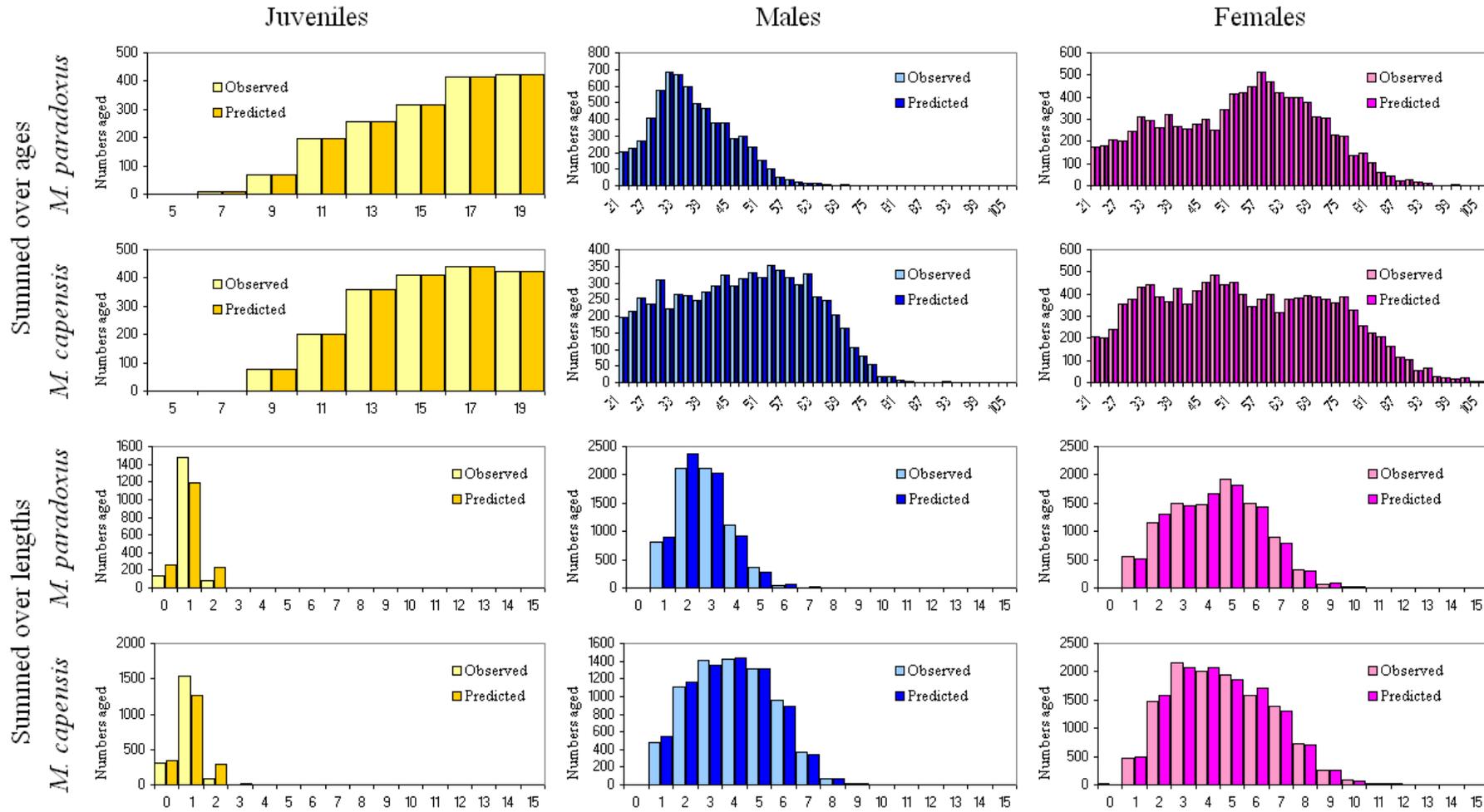


Fig. 11: Fit of the model to the ALKs, aggregated over all ALKs. The first two rows are summed over ages (so that model estimate and observation should be exactly equal by construction) and second two rows summed over lengths.

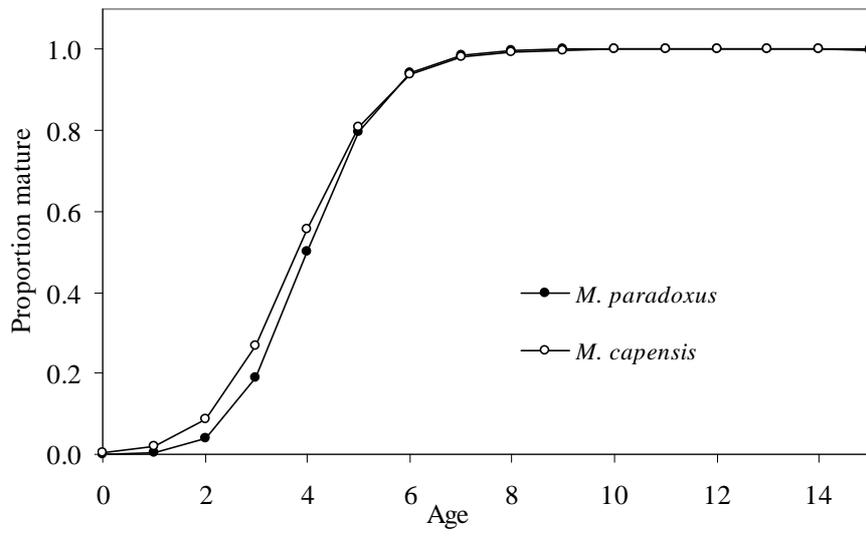


Fig. 12: Maturity-at-age used in the new Reference Case for *M. paradoxus* and *M. capensis*

## APPENDIX I – The Data Utilized

### I.1 Annual catches

The species-split of the catches is carried out external to the model. A summary of the assumptions made to disaggregate the catches by species for the New Reference Case assessment is given below. The reported or assumed catches by fleet and species are given in Table App.I.1 and plotted in Fig. App.I.1.

#### Offshore trawl fleet

From 1978 onwards, the catches made by the offshore trawl fleet have been split by species by applying the size-based species proportion-by-depth relationships for the west and south coasts which were developed by Gaylard and Bergh (2004) from research survey data.

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. The catch data for the 1917-1977 period are split by assuming that the proportion of *M. capensis* caught follows a logistic function over this period, starting at 1 and then decreasing to stabilise at the 1978-1982 average value. As trawling was concentrated in inshore areas around Cape Town and to the east when the fishery began (i.e. probably catching *M. capensis* exclusively) and progressively moved offshore, this seems a more defensible approach. To reflect a change from a *M. capensis* only fishery to the species ratio in the catch in 1978, the changing proportion with year  $y$  of *M. capensis* in the offshore trawl catch on coast  $c$  is modelled by:

$$prop_{cy}^{prop} = \frac{1 - \Delta_c}{1 + \exp[(y - P_1)/P_2]} + \Delta_c \quad (\text{App.I.1})$$

where

$\Delta_c$  is the average proportion of *M. capensis* in the offshore catch over the 1978-1982 period for coast  $c$  (24% and 60% for the west and south coasts respectively), 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 mid-way between 100% and  $\Delta_c$ , while  $P_2$  determines how rapidly this change in proportion occurs.

The New Reference Case assessment assumes:  $P_1=1950$  and  $P_2=1.5$ .

#### Inshore trawl and handline fleets

Catches made by these fleets are assumed to consist of *M. capensis* only, as they operate in relatively shallow water on the south coast.

#### Longline fleet

Longline catches on the west coast are assumed to consist of 30% *M. capensis* for the whole period, while on the south coast, catches by this fleet are assumed to consist of *M. capensis* exclusively (Andrew Penney, PISCES, pers. commn).

The total catch in 2009 is assumed equal to the TAC for that year (118 500 t); it is split between the different fleets and species assuming the same proportions as in 2008.

### I.2 Abundance indices

Six CPUE time-series are available for assessing the status of the resource (Table App.I.2): a CPUE series for each of the south and west coasts developed by the International Commission for South East Atlantic Fisheries (ICSEAF, 1989) and a GLM-standardised CPUE series for each coast, for each of *M. paradoxus* and *M. capensis* (Table App.I.2) from the offshore trawl fleet (Glazer and Butterworth, 2009). The two historical CPUE series cannot be disaggregated by species, as there are no effort-by-

depth data available for this pre-1978 period. The GLM standardized CPUE indices are species-specific (the catch data being based on the Gaylard and Bergh (2004) algorithm).

Research surveys have been conducted on board the FRS *Africana* from 1986 in spring and/or autumn on the south coast and from 1985 in summer and/or winter on the west coast, and provide fully species-specific information. Since 2003, new fishing gear has occasionally been used on the *Africana*, for which a calibration factor is available. Survey biomass estimates and their estimated (sampling) standard errors are listed in Tables App.I.3-4 (Fairweather, 2009). Only surveys extending to the deepest depth (500m) normally included in the survey design are considered for reasons of comparability.

### I.3 Length frequencies

Survey length frequencies are available disaggregated by species and in some years disaggregated by gender (Table App.I.5) (Fairweather, 2009).

Sex-aggregated proportions-at-length for each survey stratum ( $p_{yl}^{surv,i}$ ) are provided in 1cm length classes (Fairweather, 2009). In some instances, the proportions of males and females for a particular survey stratum and length class are available ( $q_{yl}^{g,surv,i}$ , where  $\sum_g q_{yl}^{g,surv,i} = 1$ ). These are converted to survey specific (i.e. aggregated over all strata for a particular cruise) proportions-at-lengths for males ( $g=1$ ), females ( $g=2$ ) and unsexed ( $g=0$ ) (with  $\sum_{g=0}^2 p_{yl}^{g,surv,i} = 1$ ) as follows:

The proportions-at-length are grouped into 2cm length classes.

- a. For all length classes < 21 cm, the proportions-at-length are assumed to be unsexed;
- b. For length classes > 20 cm:

- If there is no sex-information for either of the two 1 cm length classes to group (i.e.  $\sum_g q_{yl}^{g,surv,i} = 0$  and  $\sum_g q_{y,l+1}^{g,surv,i} = 0$ ), then the proportion for the resulting 2 cm length class is assumed to be unsexed:

$$p_{yL}^{g,surv,i} = \begin{cases} p_{yl}^{surv,i} + p_{y,l+1}^{surv,i} & \text{for } g = 0 \\ 0 & \text{for } g = 1/2 \end{cases} \quad (\text{App.I.1})$$

- If there is sex-information for one of the two 1 cm length classes to group (i.e. if  $\sum_g q_{yl}^{g,surv,i} = 1$  or  $\sum_g q_{y,l+1}^{g,surv,i} = 1$ ), then the sex-information from the one length class is used for both:

$$p_{yL}^{g,surv,i} = \begin{cases} 0 & \text{for } g = 0 \\ q_{yl}^{g,surv,i} (p_{yl}^{surv,i} + p_{y,l+1}^{surv,i}) & \text{for } g = 1/2 \end{cases} \quad (\text{App.I.2})$$

- If there is sex-information for both of the two 1 cm length classes to group (i.e. if  $\sum_g q_{yl}^{g,surv,i} = 1$  and  $\sum_g q_{y,l+1}^{g,surv,i} = 1$ ), then the sex-information is used directly:

$$p_{yL}^{g,surv,i} = \begin{cases} 0 & \text{for } g = 0 \\ q_{yl}^{g,surv,i} p_{yl}^{surv,i} + q_{y,l+1}^{g,surv,i} p_{y,l+1}^{surv,i} & \text{for } g = 1/2 \end{cases} \quad (\text{App.I.3})$$

c. The strata proportions-at-length are weighted by the estimated total number in the strata to obtain a survey specific gender-disaggregated proportions-at-length ( $p_{yl}^{g,surv}$ ). The estimated total number in each stratum is calculated as:

$$N_y^{surv,i} = B_y^{surv,i} / \bar{W}_y^{surv,i} \quad (\text{App.I.4})$$

where

$B_y^{surv,i}$  is the survey biomass estimate for stratum  $i$  in survey  $surv$ , and

$\bar{W}_y^{surv,i}$  is the mean weight of fish for stratum  $i$  in survey  $surv$ , with

$$\bar{W}_y^{surv,i} = \sum_l (p_{yl}^{surv,i} \alpha l^\beta) \quad (\text{App.I.5})$$

d. For each 2 cm length class, if the unsexed proportion is less than 20% of the total proportion in that length class, the sexed proportion is used to split the unsexed proportion into males and females.

Figs. App.I.2-3 plot the survey length frequencies available.

Length frequency information from the commercial catch is not available by species, the reason being that it is often based on cleaned (headed and gutted) fish, which cannot be easily identified by species. As a result it is not possible to disaggregate the commercial length frequencies by species. Commercial catches-at-length for the offshore and for the inshore and longline fleets are shown in Tables App.II.10-12. The south coast inshore and longline fleet catches are assumed to consist of *M. capensis* only.

#### I.4 Age-Length Keys

Table App.I.6 lists the age-length keys available. Data from animals with frills on gills (FOG) have been discarded (<3% of the total). All aged animals less or equal to 20cm in length are assumed to be juveniles, i.e. of unknown gender. The few unsexed data from animals greater than 20cm have been discarded (<1% of the total).

The 'unknown reader' is in fact a combination of data from two or three readers. When the data from the two or three readers are available directly, these are used rather than the aggregated data.

Table App.I.1: Species-disaggregated catches (in thousand tons) of South African hake from the south and west coasts (see text for details).

	<i>M. paradoxus</i>			<i>M. capensis</i>					
	Offshore		Longline	Offshore		Inshore	Longline		Handline
	West coast	South coast	West coast	West coast	South coast	South coast	West coast	South coast	South coast
1917				1.000					
1918				1.100					
1919				1.900					
1920				0.000					
1921				1.300					
1922				1.000					
1923				2.500					
1924				1.500					
1925				1.900					
1926				1.400					
1927				0.800					
1928				2.600					
1929				3.800					
1930				4.400					
1931				2.800					
1932				14.300					
1933				11.100					
1934				13.800					
1935	0.001			14.999					
1936	0.001			17.699					
1937	0.003			20.197					
1938	0.005			21.095					
1939	0.010			19.990					
1940	0.028			28.572					
1941	0.057			30.543					
1942	0.126			34.374					
1943	0.268			37.632					
1944	0.465			33.635					
1945	0.763			28.437					
1946	1.991			38.409					
1947	3.743			37.657					
1948	9.304			49.496					
1949	14.770			42.630					
1950	27.306			44.694					
1951	44.856			44.644					
1952	53.304			35.496					
1953	62.466			31.034					
1954	74.752			30.648					
1955	84.517			30.883					
1956	88.043			30.157					
1957	94.982			31.418					
1958	98.660			32.040					
1959	110.468			35.532					
1960	121.131			38.769		1.000			
1961	112.716			35.984		1.308			
1962	111.918			35.682		1.615			
1963	128.545			40.955		1.923			
1964	123.095			39.205		2.231			
1965	153.970			49.030		2.538			
1966	147.905			47.095		2.846			
1967	134.026	5.661		42.674	8.525	3.154			
1968	108.921	11.136		34.679	16.772	3.462			
1969	125.229	15.136		39.871	22.795	3.769			
1970	108.087	9.466		34.413	14.257	4.077			
1971	153.218	12.017		48.782	18.098	4.385			
1972	185.025	18.633		58.908	28.062	4.692			
1973	119.679	28.873		38.103	43.483	5.000			
1974	93.296	36.254		29.704	54.599	10.056			
1975	67.975	26.920		21.642	40.543	6.372			
1976	109.144	20.722		34.750	31.208	5.740			
1977	77.616	14.753		24.712	22.219	3.500			
1978	104.101	4.018		23.411	3.568	4.931			
1979	95.374	2.759		38.149	4.161	6.093			
1980	100.779	2.950		32.749	3.510	9.121			
1981	91.624	1.302		29.306	4.184	9.400			
1982	84.990	4.240		28.359	7.118	8.089			
1983	71.202	6.124	0.161	23.231	6.392	7.672	0.069		
1984	81.804	4.843	0.256	29.451	6.092	9.035	0.110	0.016	
1985	91.096	10.442	0.817	33.974	9.574	9.203	0.350	0.292	0.065
1986	103.494	10.214	0.965	30.311	5.751	8.724	0.413	0.302	0.084
1987	94.742	9.269	2.500	22.783	6.415	8.607	1.071	0.353	0.096
1988	83.041	7.430	3.628	23.541	7.086	8.417	1.555	0.331	0.071
1989	82.592	7.640	0.203	24.667	11.272	10.038	0.087	0.032	0.137
1990	76.754	12.028	0.270	26.882	10.764	10.012	0.116		0.348
1991	86.172	14.004		20.278	8.070	8.206		3.000	1.270
1992	81.812	20.023		21.475	6.439	9.252		1.500	1.099
1993	101.982	11.086		16.083	3.175	8.870			0.278
1994	104.347	7.601	1.130	19.520	3.451	9.569	0.484	0.626	0.449
1995	93.332	4.552	0.670	27.564	2.598	10.630	0.287	0.650	0.756
1996	109.831	9.759	1.676	19.344	3.530	11.062	0.718	1.828	1.515
1997	99.881	11.904	1.806	17.273	3.932	8.834	0.774	1.872	1.404
1998	110.856	10.796	0.647	16.797	3.357	8.283	0.277	1.471	1.738
1999	87.509	12.435	1.963	16.252	2.911	8.595	0.841	4.144	2.749
2000	95.035	8.592	3.456	22.613	4.991	10.906	1.481	2.077	5.500
2001	100.614	9.539	2.793	17.262	6.338	11.836	1.197	1.688	7.300
2002	87.862	14.199	4.772	17.488	3.965	9.581	2.045	3.945	3.500
2003	97.044	18.904	4.668	10.292	4.169	9.883	2.000	4.878	3.000
2004	89.494	27.668	3.758	11.251	4.588	10.004	1.611	4.429	1.600
2005	91.148	22.594	4.172	7.136	3.634	7.881	1.788	4.559	0.700
2006	88.571	16.555	3.592	9.525	3.474	5.524	1.539	4.032	0.400
2007	98.212	12.573	3.151	13.427	1.902	6.350	1.350	3.834	0.400
2008	91.957	13.788	2.170	9.962	1.165	5.496	0.930	2.740	0.231
2009	84.906	12.731	2.004	9.205	1.078	5.075	0.859	2.530	0.213

Table App.I.2: South and west coast historic (ICSEAF, 1989) and GLM standardized CPUE data (GLM3 of Glazer and Butterworth, 2009) for *M. paradoxus* and *M. capensis*. The historic CPUE series are for *M. capensis* and *M. paradoxus* combined.

Year	ICSEAF CPUE (t hr <sup>-1</sup> )		Year	GLM CPUE (kg min <sup>-1</sup> )			
	Species-aggregated			<i>M. paradoxus</i>		<i>M. capensis</i>	
	West Coast	South Coast		West Coast	South Coast	West Coast	South Coast
1955	17.31		1978	4.18	0.88	0.76	2.23
1956	15.64		1979	4.11	0.83	1.22	2.17
1957	16.47		1980	3.85	1.28	1.06	2.71
1958	16.26		1981	3.85	0.81	1.06	2.34
1959	16.26		1982	3.79	1.13	0.94	2.42
1960	17.31		1983	4.05	1.25	1.25	2.77
1961	12.09		1984	4.14	1.32	1.34	3.28
1962	14.18		1985	4.70	1.84	1.60	4.08
1963	13.97		1986	4.26	1.83	1.20	3.34
1964	14.60		1987	3.56	1.70	1.01	3.05
1965	10.84		1988	3.60	1.34	0.84	3.21
1966	10.63		1989	3.79	1.36	1.00	3.49
1967	10.01		1990	3.75	2.07	1.10	4.00
1968	10.01		1991	4.40	1.97	1.08	3.87
1969	8.62	1.28	1992	3.71	2.39	1.39	3.50
1970	7.23	1.22	1993	3.86	1.87	1.29	2.65
1971	7.09	1.14	1994	4.30	1.63	1.41	3.19
1972	4.90	0.64	1995	3.41	1.12	1.78	3.10
1973	4.97	0.56	1996	4.31	1.71	1.52	3.08
1974	4.65	0.54	1997	3.57	2.00	1.51	2.48
1975	4.66	0.37	1998	4.07	1.80	1.73	2.52
1976	5.35	0.40	1999	3.32	2.07	1.54	2.79
1977	4.84	0.42	2000	2.91	1.40	1.50	2.88
			2001	2.34	1.52	1.11	2.12
			2002	2.24	1.31	1.12	2.52
			2003	2.92	1.56	0.81	3.00
			2004	2.34	1.33	0.80	2.61
			2005	2.18	1.12	0.56	1.59
			2006	2.41	1.12	0.51	1.36
			2007	2.74	1.46	0.54	1.01
			2008	3.16	2.25	0.70	1.51

Table App.I.3: Survey abundance estimates and associated standard errors in thousand tons for *M. paradoxus* for the depth range 0-500m for the south coast and for the west coast. Values in bold are for the surveys conducted by the *Africana* with the new gear.

Year	West coast				South coast			
	Summer		Winter		Spring (Sept)		Autumn (Apr/May)	
	Biomass	(s.e.)	Biomass	(s.e.)	Biomass	(s.e.)	Biomass	(s.e.)
1985	169.959	(36.680)	264.839	(52.949)	-	-	-	-
1986	196.111	(36.358)	172.477	(24.122)	13.758	(3.554)	-	-
1987	284.805	(53.101)	195.482	(44.415)	21.554	(4.605)	-	-
1988	158.758	(27.383)	233.041	(64.003)	-	-	30.316	(11.104)
1989	-	-	468.780	(124.830)	-	-	-	-
1990	282.174	(78.945)	226.862	(46.007)	-	-	-	-
1991	327.020	(82.180)	-	-	-	-	26.638	(10.460)
1992	226.687	(32.990)	-	-	-	-	24.304	(15.195)
1993	334.151	(50.234)	-	-	-	-	198.849	(98.452)
1994	330.270	(58.319)	-	-	-	-	111.469	(34.627)
1995	324.554	(80.357)	-	-	-	-	55.068	(22.380)
1996	430.908	(80.604)	-	-	-	-	85.546	(25.484)
1997	569.957	(108.200)	-	-	-	-	135.192	(51.031)
1998	-	-	-	-	-	-	-	-
1999	562.859	(116.302)	-	-	-	-	321.478	(113.557)
2000	-	-	-	-	-	-	-	-
2001	-	-	-	-	19.929	(9.956)	-	-
2002	267.487	(35.068)	-	-	-	-	-	-
2003	411.177	(69.431)	-	-	<b>88.442</b>	(36.051)	108.857	(37.528)
2004	<b>259.527</b>	(56.021)	-	-	<b>63.900</b>	(17.894)	<b>48.898</b>	(20.343)
2005	<b>286.416</b>	(39.849)	-	-	-	-	<b>26.605</b>	(7.952)
2006	315.310	(49.490)	-	-	72.415	(15.500)	34.799	(8.325)
2007	<b>392.812</b>	(70.043)	-	-	<b>52.287</b>	(19.231)	<b>129.646</b>	(60.661)
2008	<b>246.542</b>	(51.973)	-	-	<b>24.816</b>	(8.775)	<b>39.505</b>	(11.408)
2009	<b>330.235</b>	(28.526)	-	-	-	-	<b>102.834</b>	(28.670)

Table App.I.4: Survey abundance estimates and associated standard errors in thousand tons for *M. capensis* for the depth range 0-500m for the south coast and for the west coast. Values in bold are for the surveys conducted by the *Africana* with the new gear.

Year	West coast				South coast			
	Summer		Winter		Spring (Sept)		Autumn (Apr/May)	
	Biomass	(s.e.)	Biomass	(s.e.)	Biomass	(s.e.)	Biomass	(s.e.)
1985	124.647	(22.707)	181.487	(27.476)	-	-	-	-
1986	117.810	(23.636)	119.587	(18.489)	121.197	(16.625)	-	-
1987	75.693	(10.241)	87.391	(11.198)	159.088	(17.233)	-	-
1988	66.725	(10.765)	47.120	(9.568)	-	-	165.939	(21.871)
1989	-	-	323.833	(67.295)	-	-	-	-
1990	455.798	(135.237)	157.800	(23.561)	-	-	-	-
1991	77.357	(14.995)	-	-	-	-	274.298	(44.395)
1992	95.407	(11.744)	-	-	-	-	138.085	(15.357)
1993	92.598	(14.589)	-	-	-	-	158.340	(13.733)
1994	121.257	(35.951)	-	-	-	-	160.555	(23.701)
1995	199.142	(26.812)	-	-	-	-	236.025	(31.840)
1996	83.337	(9.285)	-	-	-	-	244.410	(25.107)
1997	257.293	(46.056)	-	-	-	-	183.087	(18.906)
1998	-	-	-	-	-	-	-	-
1999	198.716	(32.467)	-	-	-	-	191.203	(14.952)
2000	-	-	-	-	-	-	-	-
2001	-	-	-	-	133.793	(20.858)	-	-
2002	106.253	(15.813)	-	-	-	-	-	-
2003	75.960	(13.314)	-	-	<b>82.928</b>	(9.010)	128.450	(20.062)
2004	<b>205.939</b>	(33.216)	-	-	<b>106.119</b>	(15.596)	<b>99.902</b>	(12.027)
2005	<b>70.983</b>	(13.845)	-	-	-	-	<b>76.932</b>	(5.965)
2006	88.420	(22.851)	-	-	99.867	(9.803)	130.900	(14.816)
2007	<b>82.270</b>	(11.441)	-	-	<b>74.615</b>	(7.383)	<b>70.940</b>	(5.615)
2008	<b>50.877</b>	(5.355)	-	-	<b>94.232</b>	(11.456)	<b>108.195</b>	(9.978)
2009	<b>175.289</b>	(39.920)	-	-	-	-	<b>124.004</b>	(11.808)

Table App. I.5: Survey length frequencies currently available.

Year	West coast				South coast			
	Summer		Winter		Spring (Sept)		Autumn (Apr/May)	
	Sex-aggr.	By sex	Sex-aggr.	By sex	Sex-aggr.	By sex	Sex-aggr.	By sex
1985	†	-	†	-	-	-	-	-
1986	†	-	†	-	†	-	-	-
1987	†	-	†	-	†	-	-	-
1988	†	-	†	-	-	-	†	-
1989	-	-	†	-	-	-	-	-
1990	†	-	†	-	-	-	-	-
1991	†	-	-	-	-	-	†	-
1992	†	-	-	-	-	-	†	-
1993	†	†	-	-	-	-	†	†
1994	†	†	-	-	-	-	†	†
1995	†	†	-	-	-	-	†	†
1996	†	†	-	-	-	-	†	†
1997	†	†	-	-	-	-	†	†
1998	-	-	-	-	-	-	-	-
1999	†	†	-	-	-	-	†	†
2000	-	-	-	-	-	-	-	-
2001	-	-	-	-	†	-	-	-
2002	†	-	-	-	-	-	-	-
2003	†	-	-	-	†	-	†	-
2004	†	-	-	-	†	-	†	-
2005	†	-	-	-	-	-	†	-
2006	†	†	-	-	†	†	†	†
2007	†	†	-	-	†	†	†	†
2008	†	†	-	-	†	†	†	†
2009	†	†	-	-	-	-	†	†



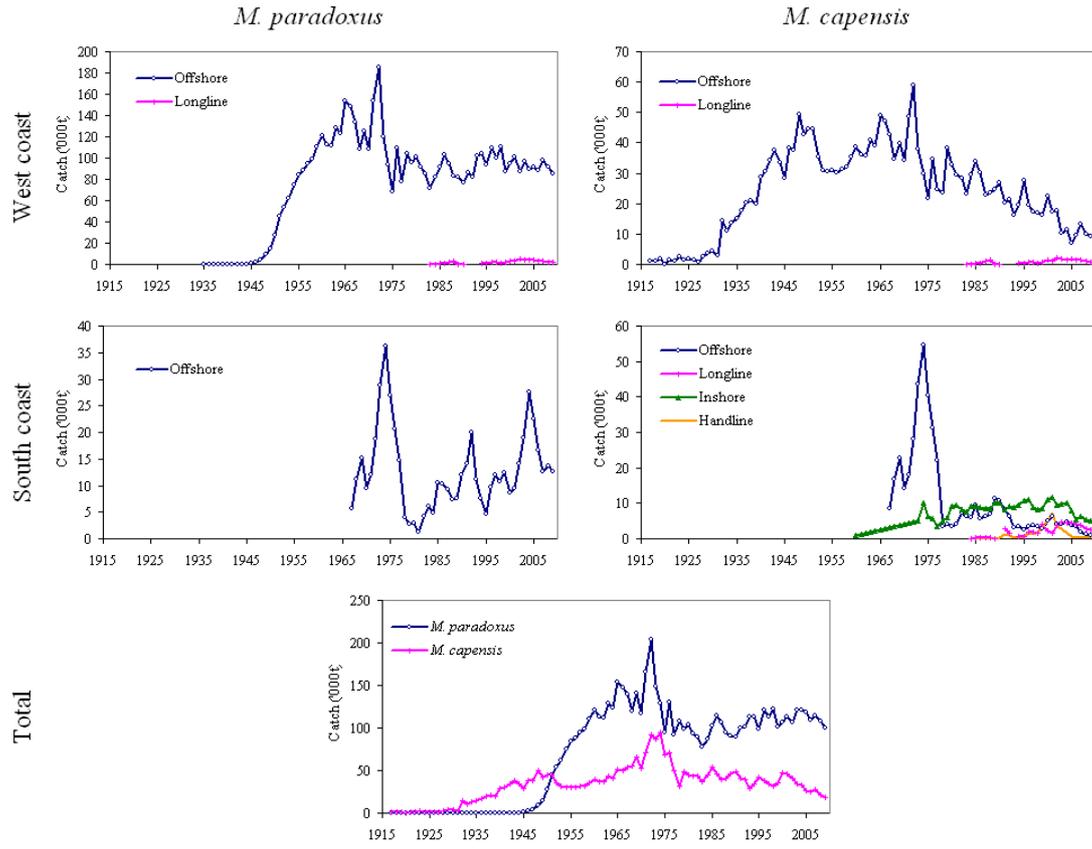


Fig. App.I.1: Annual catches, see text for details.

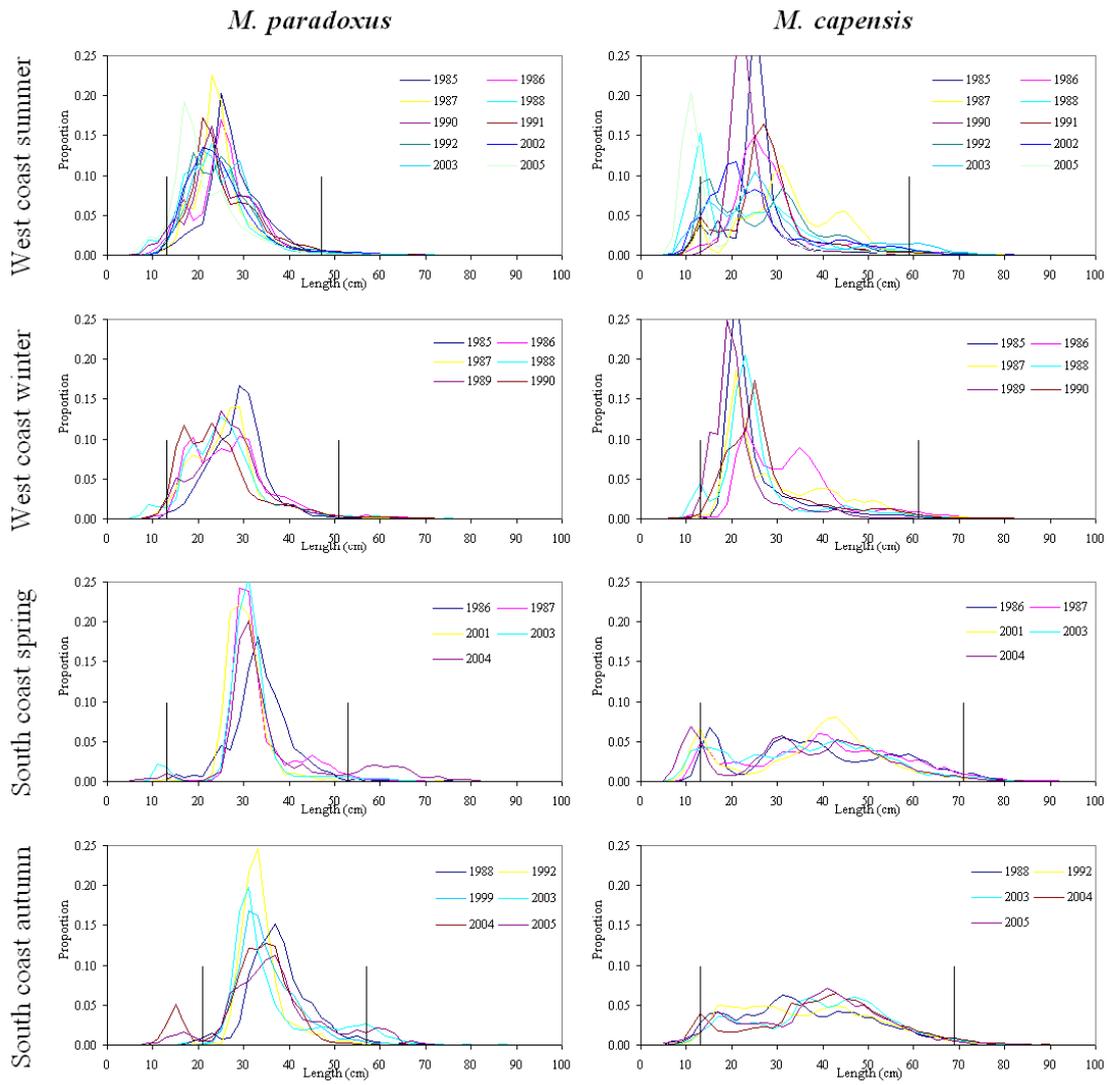


Fig. App.I.2: Sex-aggregated survey catch-at-length information. The vertical bars show the minus and plus groups used.

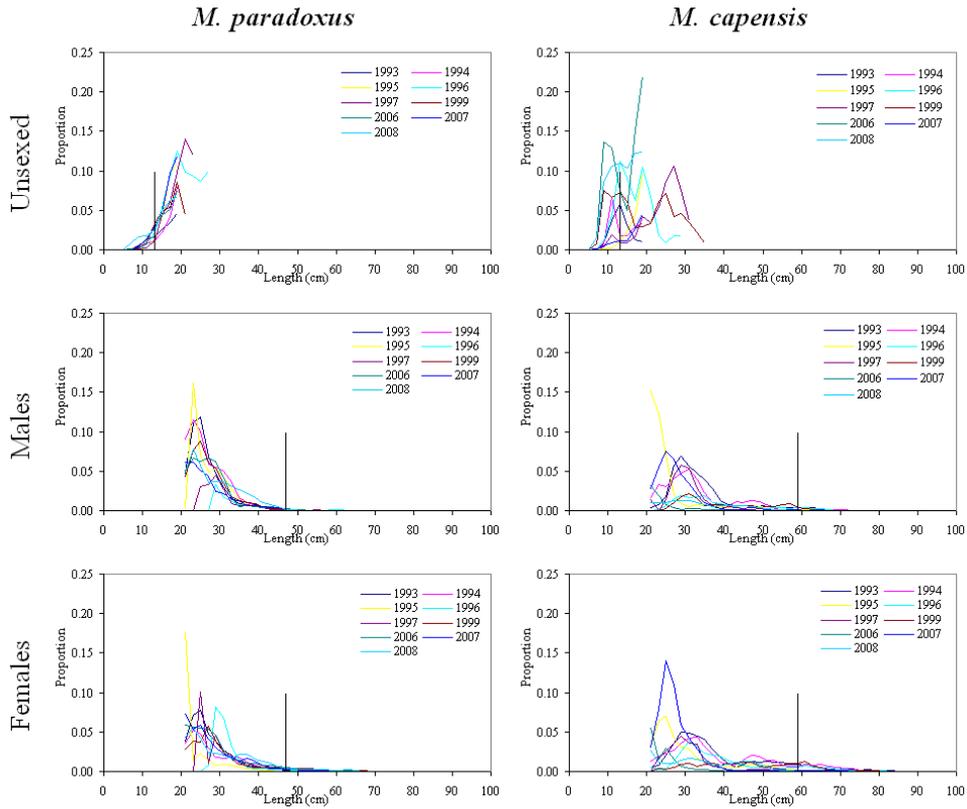


Fig App.I.3a: West coast summer gender-disaggregated survey catch-at-length information. The vertical bars show the minus and plus groups used.

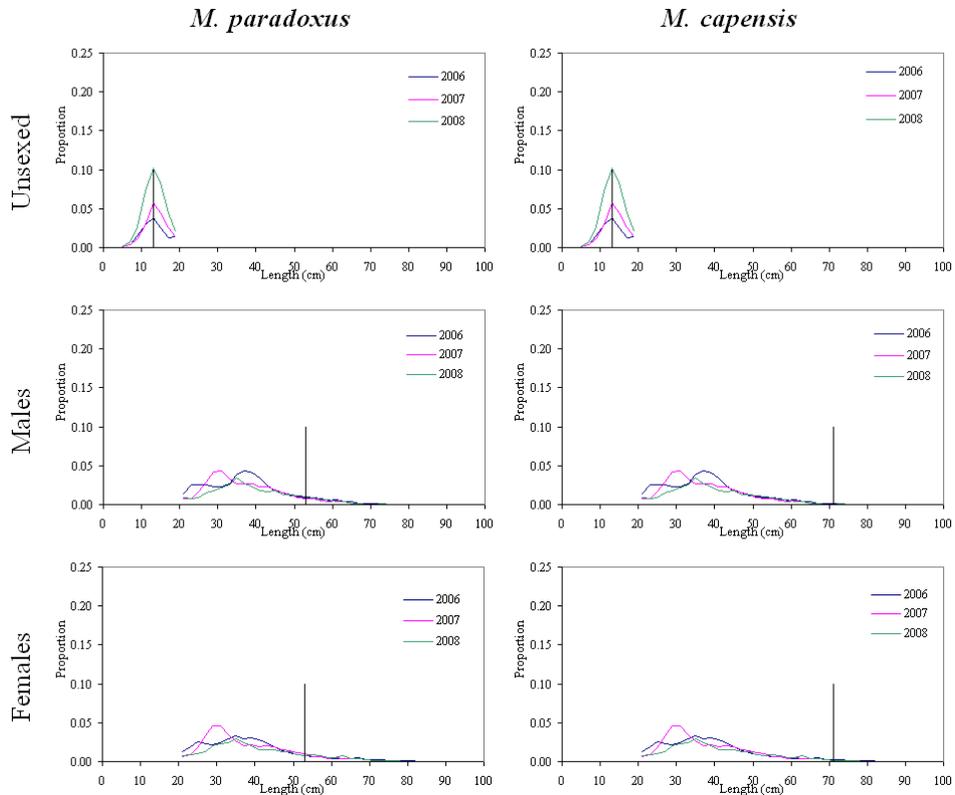


Fig App.I.3b: South coast spring gender-disaggregated survey catch-at-length information. The vertical bars show the minus and plus groups used.

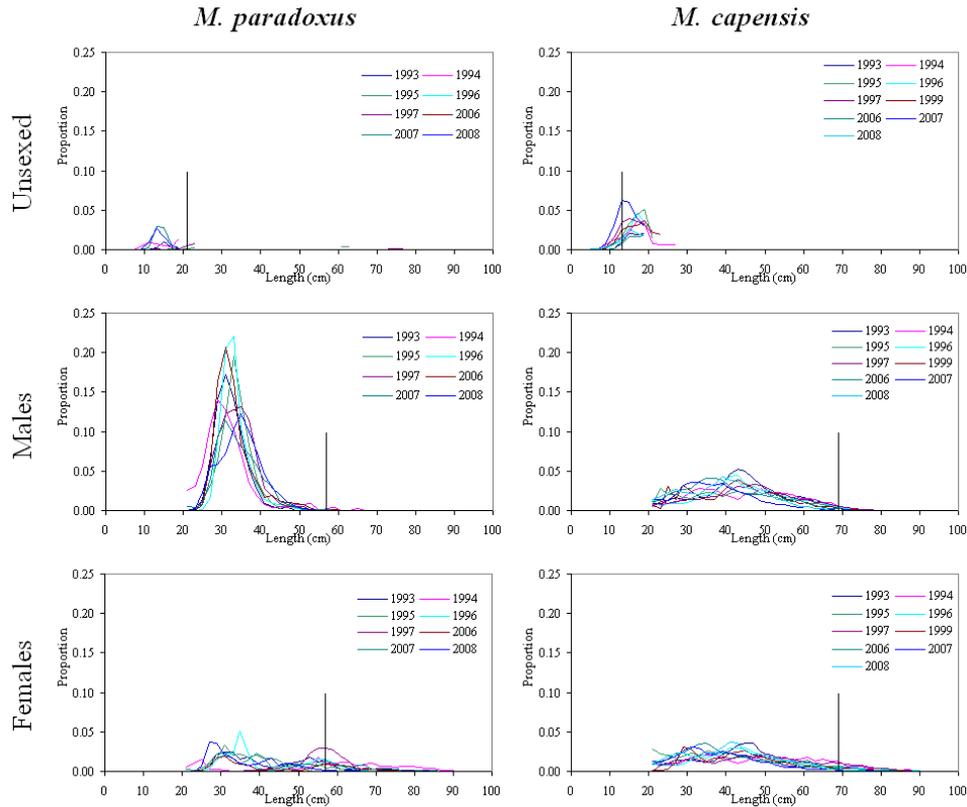


Fig App.I.3c: South coast autumn gender-disaggregated survey catch-at-length information. The vertical bars show the minus and plus groups used.

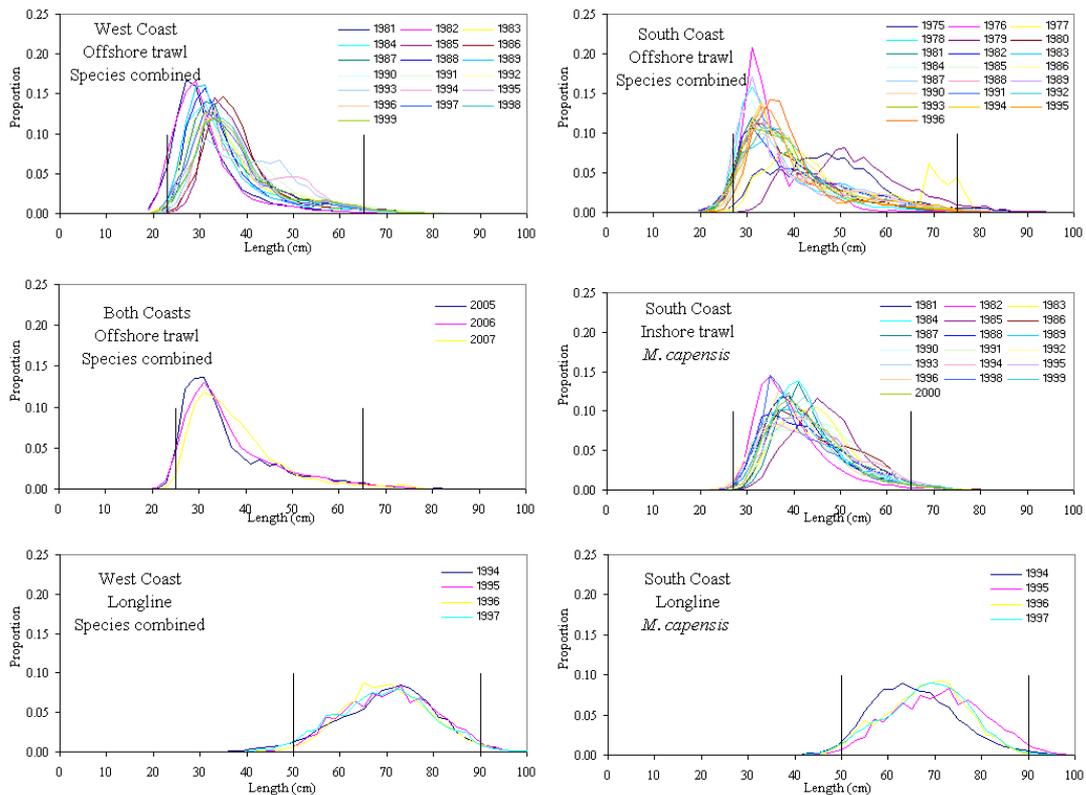


Fig App.I.4: Commercial catch-at-length information. The vertical bars show the minus and plus groups used.

## APPENDIX II – Gender-disaggregated, Age-Structured Production Model fitting to Age-Length Keys

The model used is a gender-disaggregated Age-Structured Production Model (ASPM), which is fitted directly to age-length keys (ALKs) and length frequencies. The model also involves assessing the two species as two independent stocks and is fitted to species-disaggregated data as well as species-combined data. The general specifications and equations of the overall model are set out below together with some key choices in the implementation of the methodology. Details of the contributions to the log-likelihood function from the different data considered are also given. Quasi-Newton minimisation is used to minimise the total negative log-likelihood function (implemented using AD Model Builder™, Otter Research, Ltd.).

### Population Dynamics

#### Numbers-at-age

The resource dynamics of the two populations (*M. capensis* and *M. paradoxus*) of the South African hake are modelled by the following set of equations:

Note: for ease of reading, the ‘species’ subscript *s* has been omitted below where not relevant.

$$N_{y+1,0}^g = R_{y+1}^g \quad (\text{App.II.1})$$

$$N_{y+1,a+1}^g = \left( N_{ya}^g e^{-M_a/2} - \sum_f C_{fya}^g \right) e^{-M_a^g/2} \quad \text{for } 0 \leq a \leq m - 2 \quad (\text{App.II.2})$$

$$N_{y+1,m}^g = \left( N_{y,m-1}^g e^{-M_{m-1}/2} - \sum_f C_{f,y,m-1}^g \right) e^{-M_{m-1}/2} + \left( N_{ym}^g e^{-M_m/2} - \sum_f C_{fym}^g \right) e^{-M_m/2} \quad (\text{App.II.3})$$

where

$N_{ya}^g$  is the number of fish of gender *g* and age *a* at the start of year  $y^l$ ,

$R_y^g$  is the recruitment (number of 0-year-old fish) of fish of gender *g* at the start of year *y*,

*m* is the maximum age considered (taken to be a plus-group),

$M_a$  denotes the natural mortality rate on fish of age *a* (gender independent for the moment), and

$C_{fya}^g$  is the number of hake of gender *g* and age *a* caught in year *y* by fleet *f*.

#### Recruitment

The number of recruits (i.e. new zero-year old fish) at the start of year *y* is assumed to be related to the corresponding **female** spawning stock size (i.e., the biomass of mature **female** fish) by means of the Beverton-Holt (Beverton and Holt, 1957) stock-recruitment relationship, parameterized in terms of the “steepness” of the stock-recruitment relationship, *h*, and the pre-exploitation equilibrium female spawning biomass,  $K^{\ominus sp}$ , and pre-exploitation recruitment,  $R_0$  and assuming a 50:50 sex-split at recruitment.

<sup>1</sup> In the interests of less cumbersome notation, subscripts have been separated by commas only when this is necessary for clarity.

$$R_y^g = \frac{4hR_0 B_y^{\circ,sp}}{K^{\circ,sp}(1-h) + (5h-1)B_y^{\circ,sp}} e^{(\zeta_y - \sigma_R^2/2)} \quad (\text{App.II.4})$$

where

$\zeta_y$  reflects fluctuation about the expected recruitment in year  $y$ ;

$B_y^{\circ,sp}$  is the female spawning biomass at the start of year  $y$ , computed as:

$$B_y^{\circ,sp} = \sum_{a=1}^m f_a w_a^{\circ} N_{ya}^{\circ} \quad (\text{App.II.5})$$

where

$w_a^g$  is the begin-year mass of fish of gender  $g$  and age  $a$ , and

$f_a$  is the proportion of fish of age  $a$  that are mature, and

$$R_0 = K^{\circ,sp} / \left[ \sum_{a=1}^{m-1} f_a w_a^{\circ} e^{-\sum_{a'=0}^{a-1} M_{a'}} + f_m w_m^{\circ} \frac{e^{-\sum_{a'=0}^{m-1} M_{a'}}}{1 - e^{-M_m}} \right] \quad (\text{App.II.6})$$

#### Total catch and catches-at-age

The fleet-disaggregated catch by mass, in year  $y$  is given by:

$$C_{fy} = \sum_g \sum_{a=0}^m w_{a+1/2}^g C_{fya}^g = \sum_g \sum_{a=0}^m w_{a+1/2}^g N_{ya}^g e^{-M_a/2} F_{fy} S_{fya}^g \quad (\text{App.II.7})$$

where

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

$C_{fya}^g$  is the catch-at-age, i.e. the number of fish of gender  $g$  and age  $a$ , caught in year  $y$  by fleet  $f$ ,

$F_{fy}$  is the fishing mortality of a fully selected age class, for fleet  $f$  in year  $y$  (independent of  $g$ ), and

$S_{fya}^g$  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 gender  $g$  at age  $a$  for year  $y$ , and fleet  $f$ ; when  $S_{fya}^g = 1$ , the age-class  $a$  is said to be fully selected. The south coast offshore trawl *M. paradoxus* female selectivity has been scaled down by a factor estimated in the model fitting (see below for details), for all other combinations of fleet and species, the male and female selectivities-at-length are assumed to be equal.

As it is not possible to estimate a sex-specific commercial selectivity-at-age, it is rather assumed that the selectivity-at-length is the same for both males and females (except for south coast offshore trawl *M. paradoxus* as explained above). The selectivity-at-length is converted to selectivity-at-age as follows:

$$S_{fya}^g = \sum_l S_{fyl} P_{a,l}^g \quad (\text{App.II.8})$$

where

$P_{a,l}^g$  is the proportion of fish of age  $a$  and gender  $g$  that fall in the length group  $l$  (i.e.,  $\sum_l P_{a,l}^g = 1$  for all ages  $a$ ).

The matrix  $P$  is calculated under the assumption that length-at-age is normally distributed about a mean given by the von Bertalanffy equation, i.e.:

$$l_a \sim N\left[l_\infty\left(1 - e^{-\kappa(a-t_0)}\right); \theta_a^2\right] \quad (\text{App.II.9})$$

where  $\theta_a$  is the standard deviation of length-at-age  $a$ , which is estimated directly in the model fitting for ages 0 to 7. For ages greater than 7,  $\theta_a = \theta_7$ .

The model estimate of the mid-year exploitable (“available”) component of biomass for each species and fleet 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:

$$B_{fy}^{ex} = \sum_g \sum_{a=0}^m w_{a+1/2}^g S_{fya}^g N_{ya}^g e^{-M_a/2} \left(1 - \sum_f S_{fya}^g F_{fy} / 2\right) \quad (\text{App.II.10})$$

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

$$B_y^{surv} = \sum_g \sum_{a=0}^{m_s} w_a^g S_a^{g,surv,sum} N_{ya}^g \quad (\text{App.II.11})$$

and in mid-year (winter):

$$B_y^{surv} = \sum_g \sum_{a=0}^m w_{a+1/2}^g S_a^{g,surv,win} N_{ya}^g e^{-M_a/2} \left(1 - \sum_f S_{fya}^g F_{fy} / 2\right) \quad (\text{App.II.12})$$

where

$S_a^{g,surv,sum/win}$  is the survey selectivity for age  $a$  for gender  $g$ , converted from survey selectivity-at-length in the same manner as for the commercial selectivity (see eqn App.II.8).

Survey selectivity-at-length has been assumed to be the same for males and females, except for *M. paradoxus* in the south coast spring and autumn surveys (see below).

Note that both the spring and autumn surveys are taken to correspond to winter (mid-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.,  $B_1^{g,sp} = K^{g,sp}$ , and year  $y=1$  corresponds to 1917 when catches are taken to commence.

### The likelihood function

The model is fit to CPUE and survey abundance indices, commercial and survey length frequencies, survey age-length keys, as well as to the stock-recruitment curve to estimate model parameters. Contributions by each of these to the negative of the log-likelihood ( $-\ln L$ ) are as follows<sup>2</sup>.

#### CPUE relative abundance data

The likelihood is calculated by assuming that the observed abundance index (here CPUE) is log-normally distributed about its expected value:

$$I_y^i = \hat{I}_y^i e^{\varepsilon_y^i} \quad \text{or} \quad \varepsilon_y^i = \ln(I_y^i) - \ln(\hat{I}_y^i) \quad (\text{App.II.13})$$

where

<sup>2</sup> Strictly it is a penalised log-likelihood which is maximised in the fitting process, as some contributions that would correspond to log-priors in a Bayesian estimation process are added.

$I_y^i$  is the abundance index for year  $y$  and series  $i$  (which corresponds to a specified species and fleet)

$\hat{I}_y^i = \hat{q}^i \hat{B}_{fy}^{ex}$  is the corresponding model estimate, where  $\hat{B}_{fy}^{ex}$  is the model estimate of exploitable resource biomass, given by equation App.II.10,

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

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

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 shallow water and z2) a “mixed zone” (Fig. App.II.1).

The total catch of hake of both species (*BS*) by fleet  $f$  in year  $y$  ( $C_{BS, fy}$ ) can be written as:

$$C_{BS, fy} = C_{C, fy}^{z1} + C_{C, fy}^{z2} + C_{P, fy} \quad (\text{App.II.14})$$

where

$C_{C, fy}^{z1}$  is the *M. capensis* catch by fleet  $f$  in year  $y$  in the *M. capensis* only zone (z1),

$C_{C, fy}^{z2}$  is the *M. capensis* catch by fleet  $f$  in year  $y$  in the mixed zone (z2), and

$C_{P, fy}$  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 ( $\gamma = B_{C, fy}^{ex, z2} / B_{C, fy}^{ex}$ ) (assumed to be constant throughout the period for simplicity) and  $\psi_{fy}$  be the proportion of the effort of fleet  $f$  in the mixed zone in year  $y$  ( $\psi_{fy} = E_{fy}^{z2} / E_{fy}$ ), so that:

$$C_{C, fy}^{z1} = q_C^{i, z1} B_{C, fy}^{ex, z1} E_{fy}^{z1} = q_C^{i, z1} (1 - \gamma) B_{C, fy}^{ex} (1 - \psi_{fy}) E_{fy} \quad (\text{App.II.15})$$

$$C_{C, fy}^{z2} = q_C^{i, z2} B_{C, fy}^{ex, z2} E_{fy}^{z2} = q_C^{i, z2} \gamma B_{C, fy}^{ex} \psi_{fy} E_{fy} \quad \text{and} \quad (\text{App.II.16})$$

$$C_{P, fy} = q_P^i B_{P, fy}^{ex} E_{fy}^{z2} = q_P^i B_{P, fy}^{ex} \psi_{fy} E_{fy} \quad (\text{App.II.17})$$

where

$E_{fy} = E_{fy}^{z1} + E_{fy}^{z2}$  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 ( $E_{fy}^{z1}$ ) and the effort in the mixed zone ( $E_{fy}^{z2}$ ), and

$q_C^{i, zj}$  is the catchability for *M. capensis* (*C*) for abundance series  $i$ , and zone  $zj$ , and

$q_P^i$  is the catchability for *M. paradoxus* (*P*) for abundance series  $i$ .

It follows that:

$$C_{C, fy} = B_{C, fy}^{ex} E_{fy} \left[ q_C^{i, z1} (1 - \gamma) (1 - \psi_{fy}) + q_C^{i, z2} \gamma \psi_{fy} \right] \quad (\text{App.II.18})$$

$$C_{P, fy} = B_{P, fy}^{ex} E_{fy} q_P^i \psi_{fy} \quad (\text{App.II.19})$$

From solving equations App.II.18 and App.II.19:

$$s_{fy} = \frac{q_C^{i,z1}(1-\gamma)}{\left\{ \frac{C_{C,fy} B_{P,fy}^{ex} q_P^i}{B_{C,fy}^{ex} C_{P,fy}} - q_C^{i,z2} \gamma + q_C^{i,z1}(1-\gamma) \right\}} \quad (\text{App.II.20})$$

and:

$$\hat{I}_y^i = \frac{C_{fy}}{E_{fy}} = \frac{C_{fy} B_{P,fy}^{ex} q_P^i \psi_{fy}}{C_{P,fy}} \quad (\text{App.II.21})$$

Zone 1 (z1): <i>M. capensis</i> only	Zone 2 (z2): Mixed zone
<i>M. capensis</i> : biomass ( $B_C^{z1}$ ), catch( $C_C^{z1}$ )	<i>M. capensis</i> : biomass ( $B_C^{z2}$ ), catch( $C_C^{z2}$ )  <i>M. paradoxus</i> : biomass ( $B_P$ ), catch( $C_P$ )
Effort in zone 1 ( $E^{z1}$ )	Effort in zone 2 ( $E^{z2}$ )

Fig. App.II.1: Diagrammatic representation of the two theoretical fishing zones.

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:

$$q_s^{SC} = r q_s^{WC} \quad (\text{App.II.22})$$

To correct for possible negative bias in estimates of variance ( $\sigma_y^i$ ) and to avoid according unrealistically high precision (and so giving inappropriately high weight) to the CPUE data, lower bounds on the standard deviations of the residuals for the logarithm of the CPUE series have been enforced; for the historic ICSEAF CPUE series (separate west coast and south coast series) the lower bound is set to 0.25, and to 0.15 for the recent GLM-standardised CPUE series, i.e.:  $\sigma^{ICSEAF} \geq 0.25$  and  $\sigma^{GLM} \geq 0.15$ .

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

$$-\ln L^{CPUE} = \sum_i \sum_y \left[ \ln(\sigma_y^i) + (\varepsilon_y^i)^2 / 2(\sigma_y^i)^2 \right] \quad (\text{App.II.23})$$

where

$\sigma_y^i$  is the standard deviation of the residuals for the logarithms of index  $i$  in year  $y$ .

Homoscedasticity of residuals for CPUE series is customarily assumed<sup>3</sup>, so that  $\sigma_y^i = \sigma^i$  is estimated in the fitting procedure by its maximum likelihood value:

$$\hat{\sigma}^i = \sqrt{1/n_i \sum_y (\ln(I_y^i) - \ln(\hat{I}_y^i))^2} \quad (\text{App.II.24})$$

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

<sup>3</sup> There are insufficient data in any series to enable this to be tested with meaningful power.

In the case of the species-disaggregated CPUE series, 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:

$$\ln \hat{q}^i = \frac{\sum_y (\ln I_y^i - \ln \hat{B}_{srfy}^{ex}) / (\sigma_y^i)^2}{\sum_y 1 / (\sigma_y^i)^2} \quad (\text{App.II.25})$$

In the case of the species-combined CPUE,  $q_C^{i,z1}$ ,  $q_C^{i,z2}$ ,  $q_P^i$  and  $\gamma$  are directly estimated in the fitting procedure.

### Survey abundance data

Data from the research surveys are treated as relative abundance indices in a similar manner to the species-disaggregated CPUE series above, with survey selectivity function  $S_a^{surv, sum/win}$  replacing the commercial selectivity  $S_{fya}$  (see equations App.II.11 and App.II.12 above, which also take account of the begin- or mid-year nature of the survey).

An estimate of sampling variance is available for most surveys and the associated  $\sigma_y^i$  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 App.II.23). The procedure adopted takes into account an additional variance  $(\sigma_A^i)^2$  which is treated as another estimable parameter in the minimisation process. This procedure is carried out enforcing the constraint that  $(\sigma_A^i)^2 > 0$ , i.e. the overall variance cannot be less than its externally input component.

In June 2003, the trawl gear on the *Africana* was changed and a different value for the multiplicative bias factor  $q$  is taken to apply to the surveys conducted with the new gear. Calibration experiments have been conducted between the *Africana* with the old gear (hereafter referred to as the “old *Africana*”) and the *Nansen*, and between the *Africana* with the new gear (“new *Africana*”) and the *Nansen*, in order to provide a basis to relate the multiplicative biases of the *Africana* with the two types of gear ( $q_{old}$  and  $q_{new}$ ). A GLM analysis assuming negative binomial distributions for the catches made (Brandão *et al.*, 2004) provided the following estimates:

$$\begin{aligned} \Delta \ln q^{capensis} &= -0.494 & \text{with } \sigma_{\Delta \ln q^{capensis}} &= 0.141 & \text{i.e. } (q^{new} / q^{old})^{capensis} &= 0.610 & \text{and} \\ \Delta \ln q^{paradoxus} &= -0.053 & \text{with } \sigma_{\Delta \ln q^{paradoxus}} &= 0.117 & \text{i.e. } (q^{new} / q^{old})^{paradoxus} &= 0.948 \end{aligned}$$

where

$$\ln q_{new}^s = \ln q_{old}^s + \Delta \ln q^s \quad \text{with } s = capensis \text{ or } paradoxus \quad (\text{App.II.26})$$

No plausible explanation has yet been found for the particularly large extent to which catch efficiency for *M. capensis* is estimated to have decreased for the new research survey trawl net. It was therefore recommended (BENEFIT, 2004) that the ratio of the catchability of the new to the previous *Africana* net be below 1, but not as low as the ratio estimated from the calibration experiments.  $\Delta \ln q^{capensis}$  is therefore taken as -0.223, i.e.  $(q^{new} / q^{old})^{capensis} = 0.8$ .

The following contribution is therefore added as a penalty (or a prior in a Bayesian context) to the negative log-likelihood in the assessment:

$$- \ln L^{q-ch} = (\ln q_{new} - \ln q_{old} - \Delta \ln q)^2 / 2 \sigma_{\Delta \ln q}^2 \quad (\text{App.II.27})$$

A different length-specific selectivity is estimated for the “old *Africana*” and the “new *Africana*”.

### Commercial proportions at length

Commercial proportions at length cannot be disaggregated by species and gender. The model is therefore fit to the proportions at length as determined for both species and gender combined.

The catches at length are computed as:

$$C_{fyl} = \sum_s \sum_g \sum_{a=0}^m N_{sya}^g F_{sfy} S_{sfyl}^g P_{sal}^g e^{-M_{sa}/2} \quad (\text{App.II.28})$$

With the predicted proportions at length:

$$\hat{P}_{yl}^i = C_{fyl} / \sum_{l'} C_{fyl'} \quad (\text{App.II.29})$$

The contribution of the proportion at length data to the negative of the log-likelihood function when assuming an “adjusted” lognormal error distribution is given by:

$$-\ell n L^{\text{length}} = 0.1 \sum_y \sum_l \left[ \ell n \left( \sigma_{len}^i / \sqrt{p_{yl}^i} \right) + p_{yl}^i \left( \ell n p_{yl}^i - \ell n \hat{p}_{yl}^i \right)^2 / 2 \left( \sigma_{len}^i \right)^2 \right] \quad (\text{App.II.30})$$

where

the superscript ‘*i*’ refers to a particular series of proportions at length data which reflect a specified fleet and coast, and species (or combination thereof) and

$\sigma_{len}^i$  is the standard deviation associated with the proportion at length data, which is estimated in the fitting procedure by:

$$\hat{\sigma}_{len}^i = \sqrt{\sum_y \sum_l p_{yl}^i \left( \ln p_{yl}^i - \ln \hat{p}_{yl}^i \right)^2 / \sum_y \sum_l 1} \quad (\text{App.II.31})$$

The initial 0.1 multiplicative factor is a somewhat arbitrary downweighting to allow for correlation between proportions in adjacent length groups. The coarse basis for this adjustment is the ratio of effective number of age-classes present to the number of length groups in the minimisation, under the argument that independence in variability is likely to be more closely related to the former.

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

### Survey proportions at length

The survey proportions at length 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 App.II.30). In this case however, data are disaggregated by species, and for some surveys further disaggregated by gender.

$$p_{syl}^{g,surv} = \frac{C_{syl}^{g,surv}}{\sum_{l'} C_{syl'}^{g,surv}} \quad \text{is the observed proportion of fish of species } s, \text{ gender } g \text{ and length } l \text{ from}$$

survey *surv* in year *y*,

$\hat{p}_{syl}^{g,surv}$  is the expected proportion of fish of species *s*, gender *g* and length *l* in year *y* in the survey *surv*, given by:

$$\hat{P}_{syl}^{g,surv} = \frac{\sum_a S_{sl}^{g,surv,sum} P_{sal}^g N_{sya}^g}{\sum_{l'} \sum_a S_{sl'}^{g,surv,sum} P_{sal'}^g N_{sya}^g} \quad (\text{App.II.32})$$

for begin-year (summer) surveys, or

$$\hat{P}_{syl}^{g,surv} = \frac{\sum_a S_{sl}^{g,surv,sum} P_{sal}^g N_{sya}^g e^{-M_{sa}/2} \left(1 - \sum_f S_{sfya}^g F_{sfy} / 2\right)}{\sum_{l'} \sum_a S_{sl'}^{g,surv,sum} P_{sal'}^g N_{sya}^g e^{-M_{sa}/2} \left(1 - \sum_f S_{sfya}^g F_{sfy} / 2\right)} \quad (\text{App.II.33})$$

for mid-year (autumn, winter or spring) surveys.

### Age-length keys

Under the assumption that fish are sampled randomly with respect to age within each length-class, the contribution to the negative log-likelihood for the ALK data (ignoring constants) is:

$$-\ln L^{ALK} = -w \sum_i \sum_l \sum_a \left[ A_{i,l,a}^{obs} \ln(\hat{A}_{i,l,a}) - A_{i,l,a}^{obs} \ln(A_{i,l,a}^{obs}) \right] \quad (\text{App.II.34})$$

where

$w$  is a downweighting factor to allow for overdispersion in these data compared to the expectation for a multinomial distribution with independent data; for the moment this weight factor is set to 0.001 (for values much higher than this, this component of the likelihood dominates those related to trends in abundance indices, with the result that fits to these trends deteriorate markedly, which indicates that there is some conflict amongst these sources of data; for the fit obtained an estimate of the overdispersion of these data suggests that this downweighting factor should be somewhat greater at 0.013),,

$A_{i,a,l}^{obs}$  is the observed number of fish of age  $a$  that fall in the length class  $l$ , for ALK  $i$  (a specific combination of survey, year, species and gender),

$\hat{A}_{i,a,l}$  is the model estimate of  $A_{i,a,l}^{obs}$ , computed as:

$$\hat{A}_{i,a,l} = W_{i,l} \frac{C_{i,l} A_{a,l}}{\sum_{a'} C_{i,l} A_{a',l}} \quad (\text{App.II.35})$$

where

$W_{i,l}$  is the number of fish in length class  $l$  that were aged for ALK  $i$ ,

$A_{a,l} = \sum_a P(a'|a) A_{a,l}$  is the ALK for age  $a$  and length  $l$  after accounting for age-reading error,

with  $P(a'|a)$ , the age-reading error matrix, representing the probability of an animal of true age  $a$  being aged to be that age or some other age  $a'$ .

Age-reading error matrices have been computed for each reader and for each species in Rademeyer (2009).

When multiple readers age the same fish, these data are considered to be independent information in the model fitting.

### Stock-recruitment function residuals

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

$$-\ln L^{SR} = \sum_s \sum_{y=y1}^{y2} \zeta_{sy}^2 / 2\sigma_R^2 \quad (\text{App.II.36})$$

where

$\zeta_{sy}$  is the recruitment residual for species  $s$ , and year  $y$ , which is assumed to be log-normally distributed with standard deviation  $\sigma_R$  and which is estimated for year  $y1$  to  $y2$  (see equation App.II.4) (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); and

$\sigma_R$  is the standard deviation of the log-residuals, which is input.

The stock-recruitment residuals are estimated for years 1985 to 2006, with recruitment for other years being set deterministically (i.e. exactly as given by the estimated stock-recruitment curve) as there is insufficient catch-at-age information to allow reliable residual estimation for earlier years. A limit on the recent recruitment fluctuations is set by having the  $\sigma_R$  (which measures the extent of variability in recruitment – see equation – App.II.36) decreasing linearly from 0.25 in 2004 to 0.1 in 2009, effectively forcing recruitment over the last years to lie closer to the stock-recruitment relationship curve.

### **Model parameters**

#### Estimable parameters

The primary parameters estimated are the species-specific female virgin spawning biomass ( $K_s^{\varphi sp}$ ) and “steepness” of the stock-recruitment relationship ( $h_s$ ). The standard deviations  $\sigma^i$  for the CPUE series residuals (the species-combined as well as the GLM-standardised series) as well as the additional variance ( $\sigma_A^i$ )<sup>2</sup> for each survey abundance series are treated as estimable parameters in the minimisation process. Similarly, in the case of the species-combined CPUE,  $q_C^{i,z1}$ ,  $q_C^{i,z2}$ ,  $q_P^i$  and  $\gamma$  are directly estimated in the fitting procedure.

The species- and gender-specific von Bertalanffy growth curve parameters ( $L_{inf}$ ,  $\kappa$  and  $t_0$ ) are estimated directly in the model fitting process, as well as age-specific  $\beta$  values used to compute the standard deviation of the length-at-age  $a$ .

The following parameters are also estimated in the model fits undertaken (if not specifically indicated as fixed).

#### *Natural mortality:*

Natural mortality ( $M_{sa}$ ) is assumed to be age-specific and is estimated using the following functional form:

$$M_{sa} = \begin{cases} M_{s2} & \text{for } a \leq 1 \\ \alpha_s^M + \frac{\beta_s^M}{a+1} & \text{for } 2 \leq a \leq 5 \\ M_{s5} & \text{for } a > 5 \end{cases} \quad (\text{App.II.37})$$

$M_{s0}$  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  $M_{s1}$ .

Upper bounds of 1.0 and 0.5 on ages 2 and 5 respectively are implemented in the New Reference Case assessment to maintain biological realism (likelihood maximisation prefers higher values for large ages in many cases, but this seems unlikely given there are few known predators of large hake).

*Commercial fishing selectivity-at-length:*

The fishing selectivity-at-length (gender independent) for each species and fleet,  $S_{sfl}$ , is estimated in terms of a logistic curve given by:

$$S_{sfl} = \left[ 1 + \exp\left(-\left(l - l_{sf}^c\right) / \delta_{sf}^c\right) \right]^{-1} \quad (\text{App.II.38})$$

where

$l_{sf}^c$  cm is the length-at-50% selectivity,

$\delta_{sf}^c$  cm<sup>-1</sup> defines the steepness of the ascending limb of the selectivity curve.

The selectivity is sometimes modified to include a decrease in selectivity at larger lengths, as follows:

$$S_{sfl} \rightarrow S_{sfl} e^{-s_{sfl}(l - l_{slope})} \quad \text{for } l > l_{slope}, \quad (\text{App.II.39})$$

where

$s_{sfl}$  measures the rate of decrease in selectivity with length for fish longer than  $l_{slope}$  for the fleet concerned, and is referred to as the “selectivity slope”.

Periods of fixed and changing selectivity have been assumed for the offshore trawl fleet to take account of the change in the selectivity at low ages over time in the commercial catches, likely due to the phasing out of the (illegal) use of net liners to enhance catch rates. Details of the fishing selectivities used in the assessment are shown in Table App.II.2 below.

On the south coast, for *M. paradoxus*, the female offshore trawl selectivity (only the trawl fleet is assumed to catch *M. paradoxus* on the south coast) is scaled down by a factor estimated in the model fitting procedure. Although there is no gender information for the commercial catches, the south coast spring and autumn surveys catch a much higher proportion of male *M. paradoxus* than female (ratios of about 7:1 and 3.5:1 for spring and autumn respectively). This is assumed to reflect a difference in distribution of the two genders which would therefore affect the commercial fleet similarly.

Table App.II.2: Details for the commercial selectivity-at-length for each fleet and species combination, as well as indications of what data are available.

	<i>M. paradoxus</i>	<i>M. capensis</i>	data available
1. West coast offshore			
1917-1976	set equal to 1989	set equal to 1989	
1977-1984	two logistic + slope parameters estimated	differential shift compared to 1993+ as for paradoxus, slope 1/3 of inshore	species combined
1985-1992	linear change between 1984 and 1993 selectivity		species combined
1993-2009	two logistic + slope parameters estimated	same as SC inshore but shifted to the right by 10 cm, slope 1/3 of inshore	species combined
2. South coast offshore			
1917-1976	set equal to 1989	set equal to 1989	
1977-1984	two logistic + slope parameters estimated	differential shift compared to 1993+ as for paradoxus, slope 1/3 of inshore	species combined
1985-1992	linear change between 1984 and 1993 selectivity		species combined
1993-2009	two logistic + slope parameters estimated	same as SC inshore but shifted to the right by 10 cm, slope 1/3 of inshore	species combined
3. South coast inshore		two logistic + slope parameters estimated	<i>M. capensis</i>
4. West coast longline	two logistic + slope parameters estimated	same as South Coast longline	species combined
5. South coast longline		two logistic + slope parameters estimated	<i>M. capensis</i>
6. South coast handline		average of South Coast longline and inshore	

*Survey fishing selectivity-at-length:*

The survey selectivities are estimated directly for seven pre-determined lengths for *M. paradoxus* and *M. capensis*. When the model was fitted to proportion-at-age rather than proportion-at-length, survey selectivities were estimated directly for each age (i.e. seven age classes); the seven length correspond to these ages. The lengths at which selectivity is estimated directly are survey specific (linear between the minus and plus groups) and are given in Table App.I.3 below. Between these lengths, selectivity is assumed to change linearly. The slope from lengths  $l_{minus}$  to  $l_{minus}+1$  is assumed to continue exponentially to lower lengths to length 1, and similarly the slope from lengths  $l_{plus}$  to  $l_{plus}-1$  for *M. paradoxus* and *M. capensis* to continue for greater lengths.

For the south coast spring and autumn surveys, gender-specific selectivities are estimated for *M. paradoxus*. Furthermore, the female selectivities are scaled down by a parameter estimated for each survey to allow for the male predominance in the survey catch.

Table App.II.3: Lengths (in cm) at which survey selectivity is estimated directly.

<i>M. paradoxus</i>	West coast summer	13	18	23	28	32	37	42	47
	West coast winter	13	18	24	29	35	40	46	51
	South coast spring	21	26	30	35	39	44	48	53
	South coast autumn	21	26	31	36	42	47	52	65
<i>M. capensis</i>	West coast summer	13	20	26	33	39	46	52	59
	West coast winter	13	17	21	30	40	47	54	61
	South coast spring	13	19	28	38	46	54	63	71
	South coast autumn	13	19	28	36	44	52	61	69

*Stock-recruitment residuals:*

Stock-recruitment residuals  $\zeta_{sy}$  are estimable parameters in the model fitting process. They are estimated separately for each species from 1985 to the present, and set to zero pre-1985 because there are no catch-at-length data for that period to provide the information necessary to inform estimation.

**Input parameters and other choice for application to hake***Age-at-maturity:*

The proportion of female fish of species  $s$  and length  $l$  that are mature is assumed to follow a gender-independent logistic curve with the parameter values given below (from Fairweather and Leslie, 2008, “stage 2, >40cm”):

	<i>M. paradoxus</i>	<i>M. capensis</i>
$l_{50}$	43.43cm	43.64cm
$\Delta$	4.33cm	6.23cm

Maturity-at-length is then converted to maturity-at-age as follows:

$$f_{sa}^g = \sum_l f_{sl} P_{a,l}^g \quad (\text{App.II.40})$$

*Weight-at-length:*

The weight-at-length for each species and gender is calculated from the mass-at-length function, with values of the parameters for this function listed below (from Fairweather, 2008, taking the average of the west and south coasts):

	<i>M. paradoxus</i>		<i>M. capensis</i>	
	<i>Males</i>	<i>Females</i>	<i>Males</i>	<i>Females</i>
$\alpha$ (gm/cm <sup><math>\beta</math></sup> )	0.007541	0.005836	0.006307	0.005786
$\beta$ (no units)	2.9882	3.0653	3.0612	3.0851

*Minus- and plus-groups*

Because of a combination of gear selectivity and mortality, a relatively small number of fish in the smallest and largest length classes are caught. In consequence, there can be relatively larger errors (in terms of variance) associated with these data. To reduce this effect, the assessment is conducted with minus- and plus-groups obtained by summing the data over the lengths below and above  $l_{minus}$  and  $l_{plus}$  respectively. The minus- and plus-group used are given in Table App.II.42 (and plotted in Figs.I.2 and 3). Furthermore, the proportions at length data (both commercial and survey) are summed into 2cm length classes for the model fitting.

Table App.II.3: Minus- and plus-groups taken for the surveys and commercial proportion at length data.

SURVEY DATA	<i>M. paradoxus</i>		<i>M. capensis</i>	
	Minus	Plus	Minus	Plus
West coast summer	13	47	13	59
West coast winter	13	51	13	61
South coast spring	13	53	13	71
South coast autumn	21	57	13	69
COMMERCIAL DATA	Minus	Plus		
West coast offshore, species combined	23	65		
South coast offshore, species combined	27	75		
South coast inshore, <i>M. capensis</i>	27	65		
West coast longline, species combined	51	91		
South coast longline, <i>M. capensis</i>	51	91		
Both coasts offshore, species combined	25	65		