

New Zealand Fisheries  
Assessment Report  
2009/7  
February 2009  
ISSN 1175-1584

## Assessment of hoki (*Macruronus novaezelandiae*) in 2008

R. I. C. C. Francis

**Assessment of hoki (*Macruronus novaezelandiae*) in 2008**

R. I. C. C. Francis

NIWA  
Private Bag 14901  
Wellington 6241

**Published by Ministry of Fisheries  
Wellington  
2009**

**ISSN 1175-1584**

©  
**Ministry of Fisheries  
2009**

Citation:  
Francis, R.I.C.C. (2009).  
Assessment of hoki (*Macruronus novaezelandiae*) in 2008.  
*New Zealand Fisheries Assessment Report 2009/7*. 80 p.

This series continues the informal  
New Zealand Fisheries Assessment Research Document series  
which ceased at the end of 1999.

## EXECUTIVE SUMMARY

**Francis, R.I.C.C. (2009). Assessment of hoki (*Macruronus novaezelandiae*) in 2008. New Zealand Fisheries Assessment Report 2009/7. 80 p.**

A new assessment is presented for hoki. This is similar to the 2007 assessment in using the same program (CASAL), stock structure (two stocks in four fishing grounds), and estimation procedure (Bayesian with lognormal errors, including a distinction between observation and process errors). Three data types were used: biomass indices (from trawl and acoustic surveys), proportions at age and sex (from trawl surveys and the four fisheries), and proportion spawning. Data new to this assessment were from a Cook Strait acoustic survey, two trawl surveys (Chatham Rise and sub-Antarctic), and proportions at age from the four fisheries.

A series of preliminary runs was carried out before the new year's data became available. From these it was concluded that (a) a new rebuilding criterion could be useful for managing hoki; (b) there is no need to change existing assumptions concerning changes in fisheries selectivities and the first random year in projections; (c) some concerns about proportion-spawning data were resolved by a reanalysis of existing data; (d) the assumed value of stock-recruit steepness should be reduced from 0.9 to 0.75; (e) catchability in the 2007 sub-Antarctic survey was higher than in 2006; and (f) model runs without natal fidelity should be dropped from the 2008 assessment.

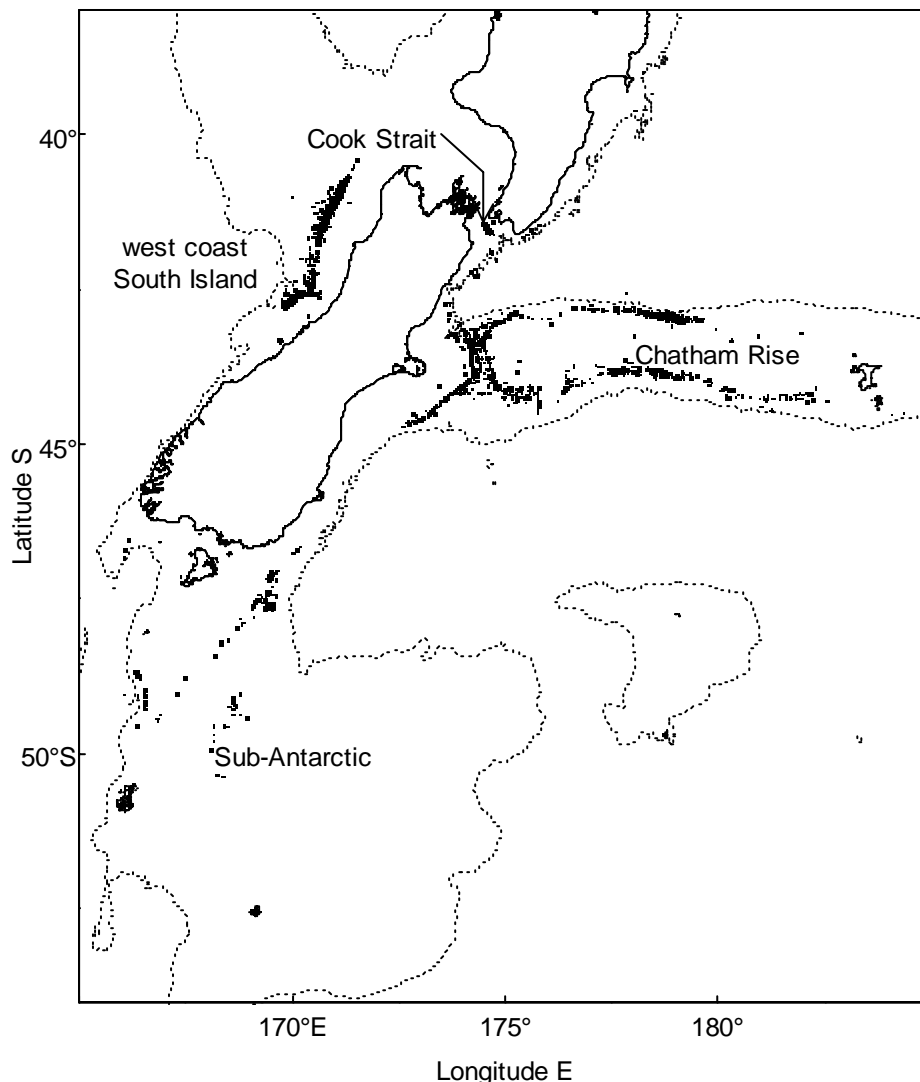
Initial runs using the new year's data helped in deciding what assumptions should be made for the 2008 assessment, and also in interpreting the assessment. It was decided that additional weight should be given to all trawl-survey biomass estimates to ensure a good fit to the decline shown in the sub-Antarctic survey series. However, no model runs were able to mimic the three-fold increase in the last biomass estimate from this series, and it was concluded that this increase was probably due to unusually high catchability in 2007. The assessment was found to be sensitive to the definition of spawning biomass (i.e., whether this calculated for both sexes or just for one) but not to the assumed value of stock-recruit steepness.

The Hoki Working Group agreed on two final model runs which were similar to two of the runs used in 2007. These provided two alternative explanations for the relative lack of old fish in the data: age-dependent natural mortality, and domed selectivities for the spawning fisheries. Both hoki stocks are estimated to be increasing after recently reaching their lowest levels since the fishery began. The West stock is much more depleted (28–30 % $B_0$ ) than the East stock (42–45 % $B_0$ ). The West stock experienced an extended period of poor recruitment from 1995 to 2001, but there is evidence of better (though still mostly below average) recruitment in subsequent years (2002–06). Projections suggest that continued fishing at current levels is likely to increase the biomass of the West stock and maintain the present biomass of the East stock.

Analyses carried out after the assessment illustrated improvements to model assumptions when natal fidelity is not assumed, provided information for an international database of stock-recruit data, and provided improved estimates of proportion spawning.

## 1. INTRODUCTION

Hoki (*Macruronus novaezelandiae*) is the most abundant commercial fish species in New Zealand waters, and has been our largest fishery since the mid 1980s. It is widely distributed throughout New Zealand's Exclusive Economic Zone in depths of 50–800 m, but most commercial fishing is at depths of 200–800 m. There are four main fisheries: two on spawning grounds (west coast South Island and Cook Strait), and two on feeding grounds (Chatham Rise and Sub-Antarctic) (Figure 1). Since the introduction of the QMS, hoki has been managed as a single fishstock, HOK 1 (ignoring HOK 10, which is purely administrative). Until recently, the TACC has oscillated between 200 000 t and its initial (1986–87) level of 250 000 t. In response to a series of poor recruitments the TACC was dropped to 180 000 t for 2003–04, to 100 000 t for 2004–05, and to 90 000 t in 2007–08 (Ministry of Fisheries 2008).



**Figure 1:** Southern New Zealand, showing the main hoki fishing grounds, the 1000 m contour (broken line), and the position of all 2006–07 tows from TCEPRs (Trawl Catch and Effort Processing Returns) in which at least 10 t of hoki was caught (dots).

Within HOK 1 two stocks are recognised — eastern and western — and these have been assessed separately since 1989. Originally, the two stocks were assessed in parallel models. More recently (since 1998 in NIWA assessments, and 2001 in industry-sponsored assessments) the stocks have been assessed simultaneously, using two-stock models. The complicated interactions inherent in a two-stock model, together with the large array of data sets that are available for HOK 1 — the 2004 NIWA assessment used more than 1800 individual observations spread over 15 data sets (Francis 2005) — make this one of the most complex of all New Zealand assessments. For most years in the 1990s there were two separate assessments — one funded by the Ministry of Fisheries (MFish) and one by industry — using different models and often reaching different conclusions.

This report documents the 2008 assessment of HOK 1, which is the seventh hoki assessment to use NIWA's general-purpose stock-assessment model CASAL (Bull et al. 2008). Since the last assessment (Francis 2008b) there has been another acoustic survey in Cook Strait in August 2007 (O'Driscoll, unpublished results) and two more trawl surveys – in the Sub-Antarctic in December 2007 (O'Driscoll & Bagley, unpublished results) and Chatham Rise in January 2008 (Stevens & O'Driscoll, unpublished results).

The work reported here addresses objective 1 of MFish project HOK2007/01: *To update the stock assessment of hoki in the year 2008, including estimates of biomass, risk and yields.*

## 2. PRELIMINARY ANALYSES

This section addresses several issues that arose during or after the 2007 hoki assessment (Francis 2008b) and which it was not possible to deal with at the time. These were considered in the context of that assessment, with the intention that this work might provide information useful in designing the 2008 assessment. Most of the analyses use data only for the years that were available during the 2007 assessment. The starting point for many of these runs was one of the three final runs from 2007 (Table 1). Some abbreviations are useful to simplify references to various data sets and parts of the model (Table 2).

**Table 1: Distinguishing characteristics for the three final model runs in the 2007 assessment.**

Label	Response to lack of old fish in the observations	Sex in model and selectivities length-based?	Natal fidelity assumed?	Biomass indices upweighted
4.4	<i>M</i> dependent on age	Yes	Yes	Trawl
4.5	Domed spawning selectivity	No	Yes	Trawl
4.7	<i>M</i> dependent on age	Yes	No	Trawl & acoustics

### 2.1 Long-term projections for the hoki rebuilding strategy

At the request of MFish, some projections were done to aid in the formulation of a rebuilding strategy for the W stock. It was assumed that, in keeping with the Draft Harvest Strategy Standard, the rebuilding plan would use the  $2T_{\min}$  criterion: that the time for the stock to rebuild to  $B_{\text{MSY}}$  should be no more than  $2T_{\min}$ , where  $T_{\min}$  is the number of years the stock would take to rebuild to  $B_{\text{MSY}}$  with zero catch. However, there were two potential difficulties in applying this criterion to hoki: uncertainty about future recruitment, and the fact that there is no agreed  $B_{\text{MSY}}$  for hoki (the latest Plenary Report gives the range 30–40%  $B_0$ , rather than a single value).

To explore the problem, twelve projections were done, covering all combinations of

- the three 2007 model runs, 4.4, 4.5, and 4.7,
- the two assumptions concerning future recruitment ('long-term' and 'recent'), and
- two scenarios for future catches:

0E0W	zero catch
65E25W	65 000 t from the E fisheries and 25 000 t from the W fisheries

**Table 2: Abbreviations used in describing the model and observations.**

Quantity	Abbreviation	Description
Stock	E	eastern stock
	W	western stock
Area	CR	Chatham Rise
	CS	Cook Strait
	SA	sub-Antarctic
	WC	west coast South Island
Fishery	Esp	E spawning fishery
	Wsp	W spawning fishery
	Ensp1, Ensp2	first and second parts of E non-spawning fishery
	Wnsp1, Wnsp2	first and second parts of W non-spawning fishery
Observation	CSacous	CS acoustic biomass index
	WCacous	WC acoustic biomass index
	CRsumbio, CRsumage	biomass index & propns at age from CR summer trawl survey
	SAsumbio, SASumage	biomass index & propns at age from SA summer trawl survey
	SAautbio, SAautage	biomass index & propns at age from SA autumn trawl survey
	pspawn	proportion spawning (estimated from SA autumn trawl survey)
	Espage, Wnspage, etc	proportions at age in catch from given fishery (from otoliths)
Migrations	EnspOLF, WnspOLF	proportions at age in catch from given fishery (from OLF <sup>1</sup> )
	Ertn, Wrtn	return migrations of E and W fish from spawning
	Whome	migration of juvenile fish from CR to SA
	Espmg, Wspmg	spawning migrations of E and W fish
Selectivity	Espsl, Wpspl, Enspsl, Wnspsl	selectivity in commercial fisheries
	CRsl, SASl	selectivity in trawl surveys

<sup>1</sup>OLF is a computer program that estimates proportions at age from length frequency data (Hicks et al. 2002).

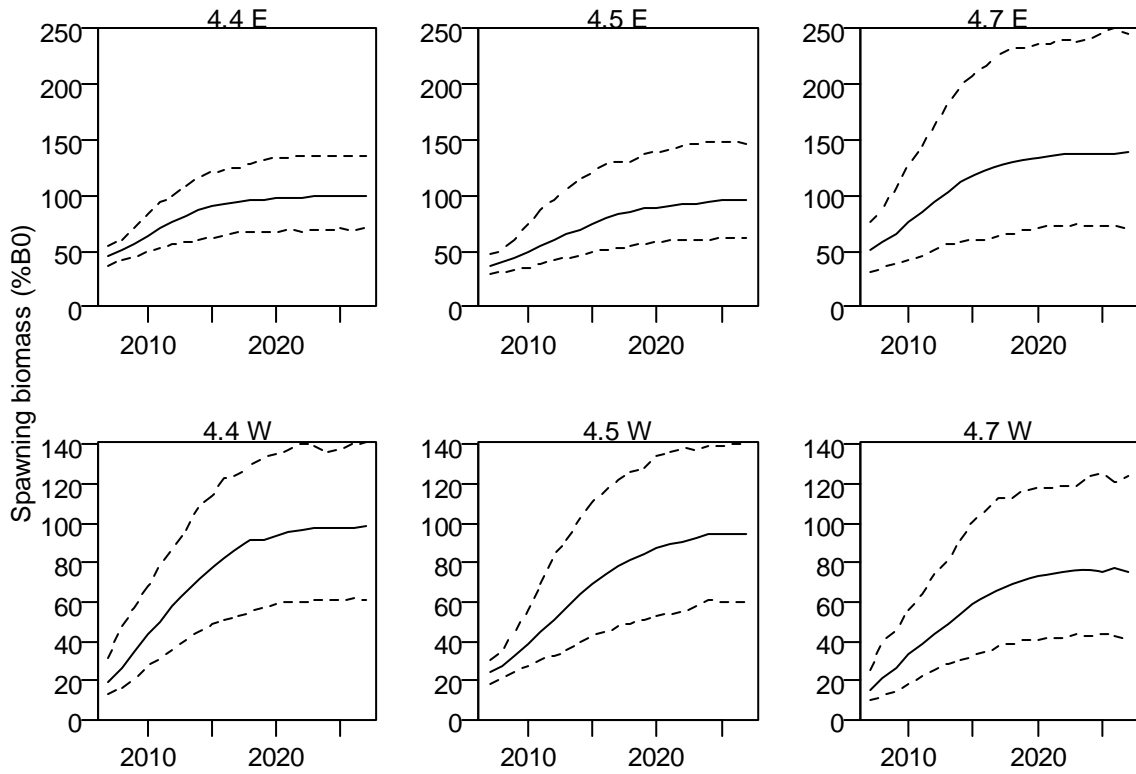
All projections were based on the 2007 stock assessment and covered a period of 20 y in order to allow the biomass to stabilise. For the second catch scenario, the catches were split between the spawning and non-spawning fisheries in the same ratios as occurred in 2007 (i.e., the 2006–07 year), leading to the following catches by fishery:

Ensp1	Ensp2	Wnsp1	Wnsp2	Esp	Wsp
25100	16800	1100	2500	23100	21400

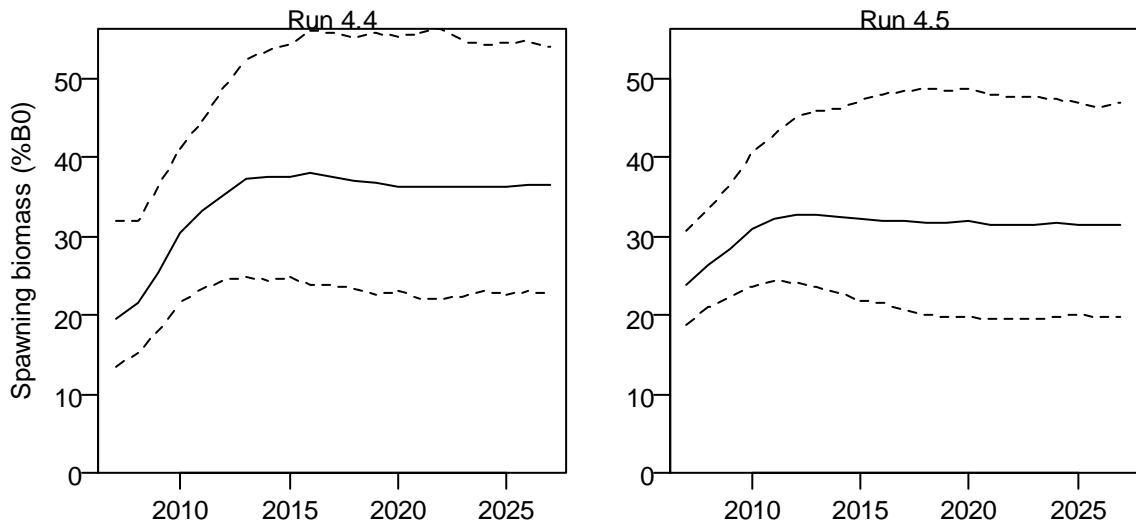
The projections revealed a problem with run 4.7. The definition of  $B_0$  (as the level of biomass that would occur, on average, if there was no fishing) led us to expect that the median biomass of each stock would tend  $B_0$  in those projections with the zero-catch scenario (OEOW) and long-term recruitment. This happened with runs 4.4 and 4.5, but not with run 4.7 (Figure 2). The reason for this is that run 4.7 assumes a single biological stock with two spawning grounds. Thus, for this run, the concept of  $B_0$  is really only properly defined for the combined stock (E + W); what is labelled  $B_0$  for the E and W stocks is actually  $B_{1972}$ . Because of this it was decided, initially, to ignore run 4.7 in developing the rebuilding plan.

The approach to rebuilding hoki was discussed in relationship to three alternative recruitment hypotheses.

*Hypothesis 1* is that there was a ‘regime shift’ in 1995 (the first of seven years of poor recruitment to the W stock), so we should expect that all recruitment in the foreseeable future will be like that seen since 1995 (i.e., we should use only the recent recruitment assumption in projections). Because recent recruitment has been lower, on average, than that in preceding years this hypothesis implies that  $B_0$  for this stock has decreased. We can estimate the new  $B_{0,W}$  from the zero-catch projections with recent recruitment, which suggest that  $B_{0W,new}$  is somewhere between 37%  $B_{0W,old}$  (the final median biomass for run 4.4) and 32%  $B_{0W,old}$  (that for run 4.5) (Figure 3). Now our estimates of current biomass from the 2007 assessment were 20%  $B_{0W,old}$  and 24%  $B_{0W,old}$  (for runs 4.4 and 4.5, respectively), which we can now express as 54%  $B_{0W,new}$  and 75%  $B_{0W,new}$ , respectively. Therefore, under this hypothesis the W stock is not depleted and there is no need for a rebuilding strategy.



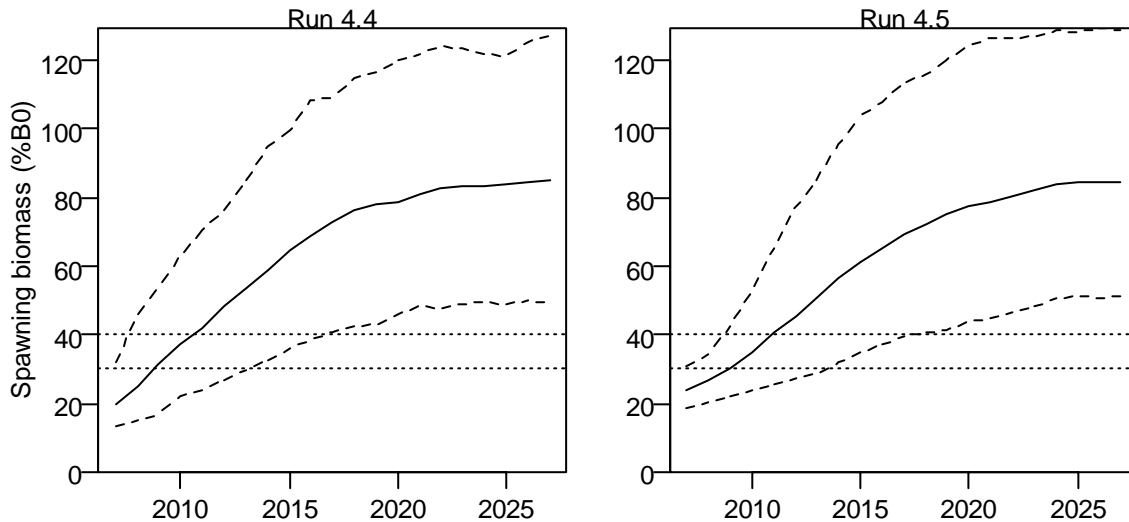
**Figure 2: Results of long-term projections for catch option '0E0W' and long-term recruitment: median spawning biomass (solid lines) with 95% confidence intervals (broken lines). Horizontal dotted lines span the range 30%–40%  $B_0$ .**



**Figure 3: Results of long-term projections for the W stock and runs 4.4 and 4.5 assuming recent recruitment and zero catches (i.e., catch option '0E0W'): median spawning biomass (solid lines) with 95% confidence intervals (broken lines).**



*Hypothesis 2* is that there has been no regime shift: the poor recruitment in 1995–2001 was just bad luck, there is no reason to expect more poor recruitment (i.e., we should use the long-term recruitment assumption), and there has been no change in  $B_{0W}$ . Under this hypothesis, we can see, from Figure 2, that  $T_{\min}$  for the W stock is somewhere between 2 y and 4 y, depending on which run we consider (4.4. or 4.5) and whether we take  $B_{MSY}$  as 30%  $B_0$  or 40%  $B_0$ . The 65E25W projections with long-term recruitment show that status quo catches will allow the W stock to rebuild to  $B_{MSY}$  well within  $2T_{\min}$  (Figure 4).



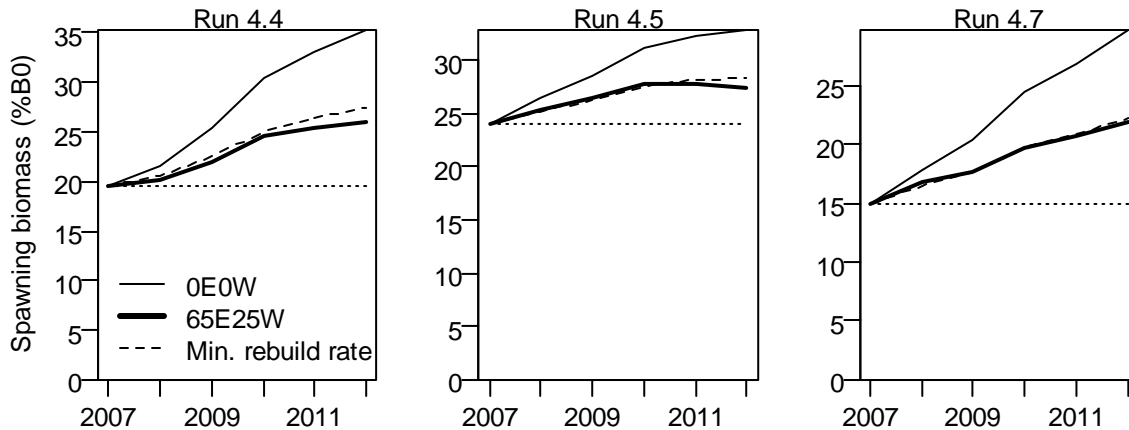
**Figure 4: Results of long-term projections for the W stock and runs 4.4 and 4.5 assuming long-term recruitment and catch option ‘65E25W’: median spawning biomass (solid lines) with 95% confidence intervals (broken lines). Horizontal dotted lines span the range 30%–40%  $B_0$ .**

*Hypothesis 3* is that there has been no regime shift (and so no change in  $B_{0W}$ ) and that recruitment will eventually improve, but the recent poor recruitment may persist for a few years. With this hypothesis we should use the recent recruitment assumption, but limit projections to a few years. The difficulty is that we don’t know how long the poor recruitment might persist, so we can’t use the  $2T_{\min}$  criterion because there is no way to calculate  $T_{\min}$ .

A new rebuilding criterion was proposed that is a generalisation of the  $2T_{\min}$  criterion. This requires that the rate of rebuilding of the stock must be at least half of that which would occur with zero catch, as illustrated in Figure 5. This plot shows that status quo catches (i.e., those in catch option 65E25W) are close to meeting this criterion, but that to make the criterion well defined we need to specify to which year it applies (note that for run 4.5 the criterion is met in 2008, 2009, and 2010, but not in 2011 or 2012). (Note that because this criterion does not depend on  $B_0$  or  $B_{MSY}$  it can be applied to run 4.7.) Given the practice in recent years of basing hoki assessment advice on 5-year projections it seems reasonable to apply the criterion to the year 5 y after the current year (i.e., 2012). That is, a catch option,  $C_{opt}$ , will satisfy the proposed criterion only if

$$B_{\text{current} + 5, C_{opt}} \geq 0.5(B_{\text{current}} + B_{\text{current} + 5, 0E0W})$$

where  $B_{\text{current} + 5, C_{opt}}$  is the projected median biomass 5 years after the current year under this catch option and  $B_{\text{current} + 5, 0E0W}$  is the corresponding biomass with zero catches. The largest annual catch from the W fisheries that will meet this constraint is 20 000 t for run 4.4, 21 000 t for run 4.5, and 22 000 t for run 4.7 (calculated by interpolation from the 2007 projection results in figure 43 of Francis (2008b)).



**Figure 5: Illustration of the rebuild-rate criterion for the W stock. In each panel the solid lines show the projected median biomass assuming recent recruitment with the 0E0W (light line) and 65E25W (heavy line) catch scenario. The new criterion is that the projected biomass should lie above the broken line, which is half-way between the 0E0W line and the no-rebuild line (dotted).**

## 2.2 More on step changes in fishery selectivities

After the 2007 assessment, the consequences of allowing temporal changes in some selectivities was investigated (see section 7.4, Francis 2008b). It was found that, although the data support allowing step changes in selectivities in both spawning fisheries, such changes had relatively little effect on the stock assessment. Thus it seemed advisable to apply Occam's razor and not include these step changes in future assessments. In these analyses, as in the stock assessment, it was assumed for all fisheries that males and females of the same length were equally selected. The Hoki Working Group asked that this analysis be repeated without this assumption. This was done for the same set of four step changes as were considered earlier (Table 3).

**Table 3: Step changes in selectivity that were investigated. Interpretation: the first line of the table describes a single step in 1999 for the Wsp fishery, which means that two selectivity curves would be estimated: WspsLE (for which 11 y of data were available, 1988–1998) and WspsLL (with 8 y of data, 1999–2006). In the new selectivity labels the last letter is 'E' for the early period, 'M' for the middle period (if any) and 'L' for the late period.**

Fishery	Step year(s)	Selectivities estimated (no. of years)
Wsp	1999	WspsLE (11 y), WspsLL (8 y)
Esp	1996, 2001	EspLE (8 y), EspLM (5 y), EspLL (6 y)
Ensp	2001	EnspLE (6 y), EnspLL (6 y)

Eight new model runs, labelled 1.1–1.8, were done. These were analogous to the runs done in the original analysis (5.7–5.14). That is, they were all the same as the 2007 final run 4.4, except that

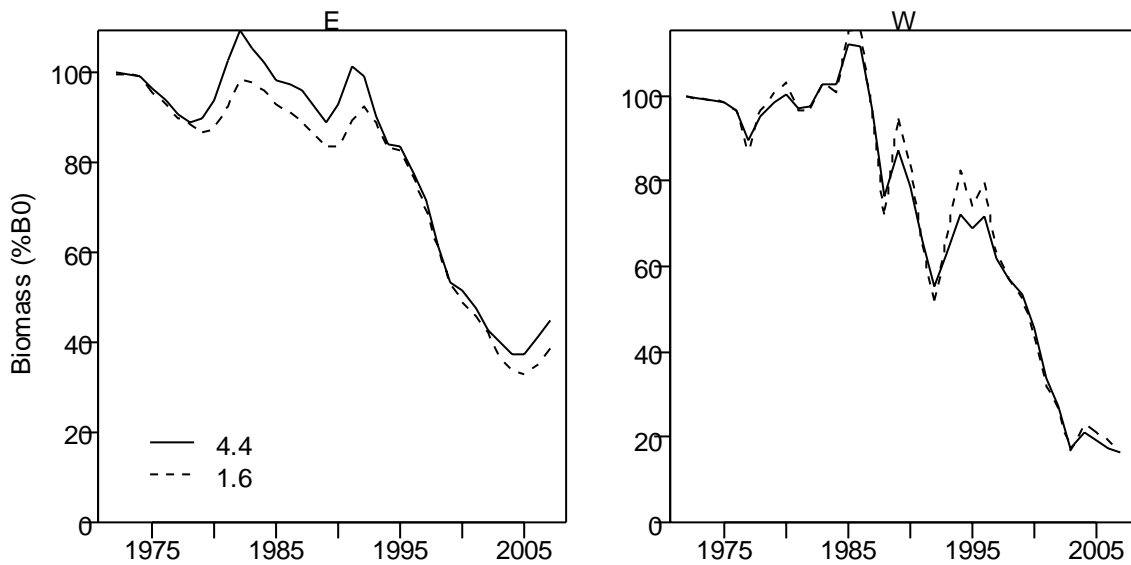
- all fishery selectivities were estimated separately for males and females, and
- year-to-year variation in the Wsp selectivity was disallowed from 1999 on,
- some selectivities had step changes, as in Table 3, and
- the constraint  $E_{psl} = W_{psl}$  that applies in run 4.4 was applied just to the earliest female selectivity in each of these two fisheries

For the new runs the effect, in terms of improvement of fit, of allowing step changes in selectivities was similar to that with the original runs, even though there were more additional parameters with the new runs (Table 4). As before, the improvement in fit was substantial for step changes in 1999 for the Wsp fishery and 1996 for Esp (runs 1.2 & 1.3), but relatively small for the other two step changes (runs 1.4 & 1.5). The best run (1.6) included the first two step changes; adding the other steps caused only a relatively small improvement in fit (see runs 1.7 & 1.8).

**Table 4: Comparison of the effects of step changes in selectivity from the original analysis (runs 5.7–5.14) with that from the new runs, 1.1–1.8, which are the same as the original runs except in estimating fishery selectivities separately by sex.**

Run		Selectivity steps	Additional parameters (relative to first run)		Improvement in fit (relative to first run)	
Original	New		Original	New	Original	New
5.7	1.1	None	0	0	0	0
5.8	1.2	Wsp(1999)	2	4	45	48
5.9	1.3	Esp(1996)	2	4	34	35
5.10	1.4	Esp(2001)	2	4	14	13
5.11	1.5	Ensp(2001)	3	6	10	11
5.12	1.6	Wsp(1999) Esp(1996)	4	8	79	83
5.13	1.7	Wsp(1999) Esp(1996) Esp(2001)	6	12	82	88
5.14	1.8	Wsp(1999) Esp(1996) Ensp(2001)	7	18	87	98

The new runs do not change the conclusions of the original analysis. Allowing step changes in selectivities does not seem worthwhile given the relatively small changes this makes to biomass trajectories (Figure 6) and projections (Figure 7). As in the original analysis, the step changes in the two spawning fisheries increased the probability of catching young fish (Figure 8).



**Figure 6: The effect of allowing step changes in selectivities on estimated biomass trajectories: comparison of trajectories from 2007 run 4.4 (with no steps, solid lines) and new run 1.6 (with a step at 1999 for Wsp and at 1996 for Esp, dotted lines).**

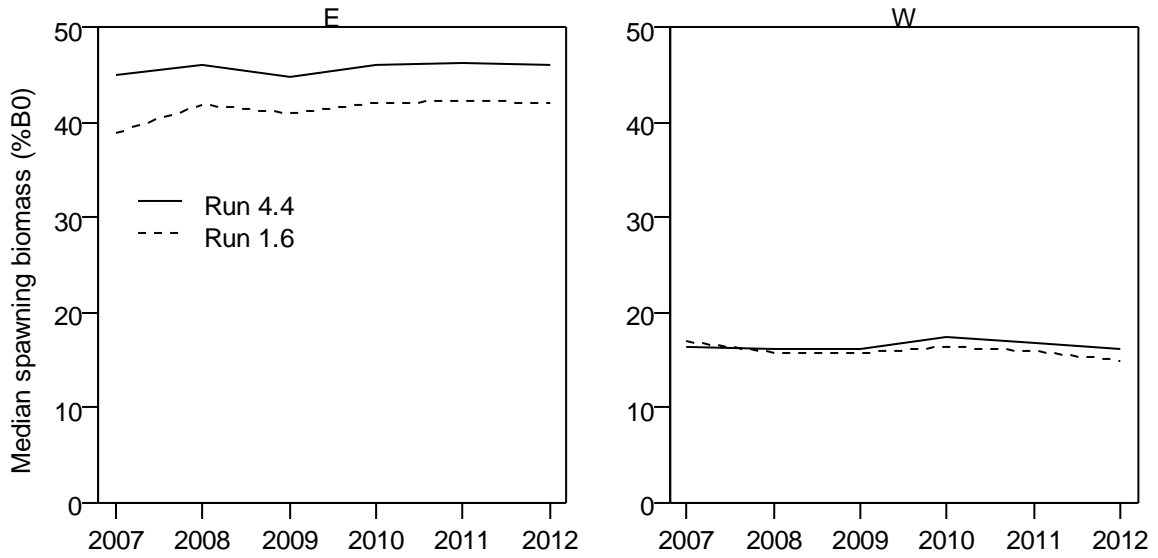


Figure 7: Effect of allowing step changes in selectivities on projected spawning biomass: MPD-based projection results for 2007 run 4.4 and new run 1.6, assuming status quo catches and recent recruitment.

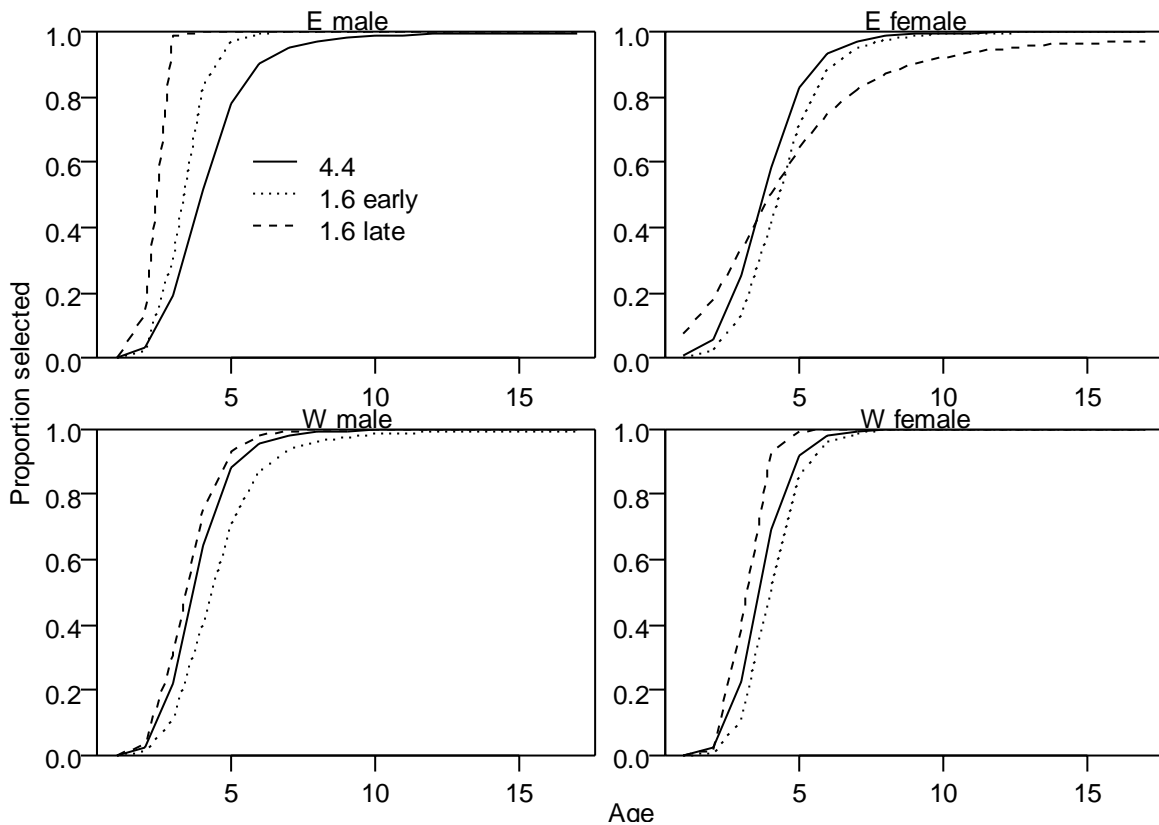


Figure 8: Comparison of the selectivities estimated in the two spawning fisheries in 2007 run 4.4 and new run 1.6 (which was the same as 4.4 except that it allowed a selectivity step for both of these fisheries).

### 2.3 Changing the first random year in projections

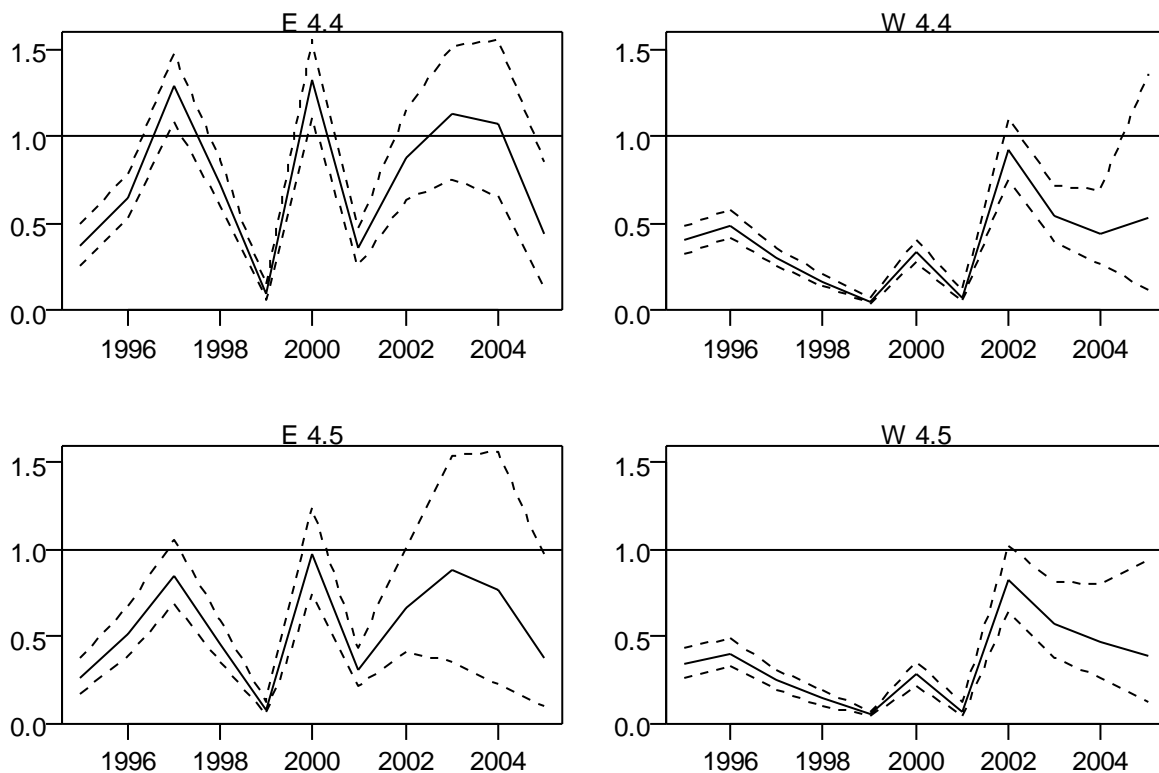
At a meeting in October 2007, the Working Group suggested the possibility of changing the first year with random (i.e., resampled) recruitment in projections. In the 2007 assessment, YCSs were estimated for the years 1975 to 2005, inclusive, and the first random year in projections was 2006. In this section I show the effect of changing the first random year to 2005, 2004, and 2003.

The data available for these year classes were quite limited (Table 5), and this was reflected in the relatively wide confidence limits for their estimated strengths (Figure 9).

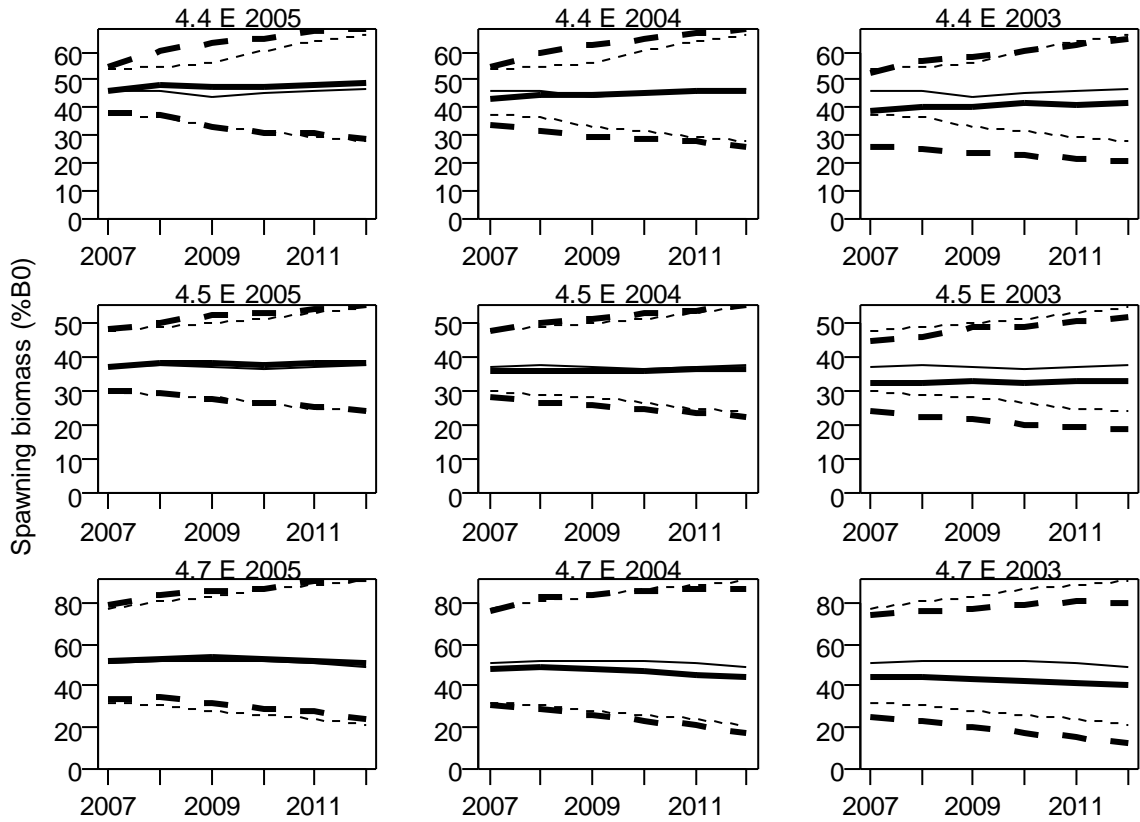
**Table 5: Ages (y) at which selected year classes were observed in the various at-age observations included in the 2007 assessment. The selected year classes are the three most recent amongst those that were estimated in the assessment. ‘-’, no data.**

Year class	Chatham Rise		Sub-Antarctic		WCSI	Cook Strait
	Survey	Fishery	Survey	Fishery		
2005	1+	-	-	-	-	-
2004	1+, 2+	1+	2+	-	2	2
2003	1+, 2+, 3+	1+, 2+	2+, 3+	2+	2, 3	2, 3

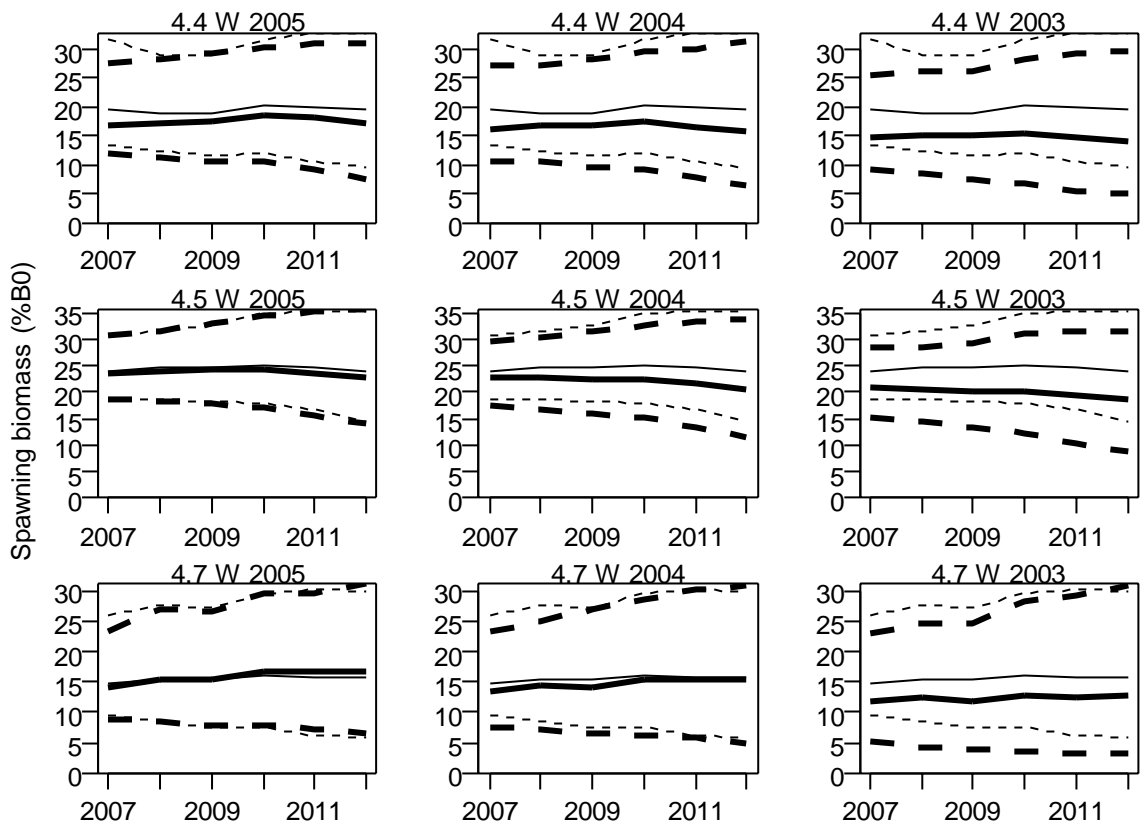
New projections were done, all assuming status quo catches and recent recruitment, with the first random year changed to 2005, then 2004, and 2003. The effect of these changes was generally not large. They produced no consistent effect on the main information in projection results – the slope of the median biomass line – with the slopes sometimes increasing slightly and sometimes decreasing slightly (Figure 10).



**Figure 9: Estimates of recent year-class strength (solid lines), with approximate 95% confidence intervals (broken lines), from the 2007 assessment.**



**Figure 10A:** The effect on projection results for the E stock of changing the first random year to 2005 (left column), 2004 (central column), or 2003 (right column). Each panel compares spawning biomass trajectories (medians as solid line, 95% confidence intervals as broken lines) with the adjusted first random year (heavy lines) with the original results, where the first random year was 2006 (light lines). All results assume status quo catches and recent recruitment.



**Figure 10B:** As in Figure 10A, but for the W stock.

The main effect of the changes was to shift the median biomass lines downwards. The shift is downwards because the mean YCS from the ‘recent’ period, from which future YCSs are selected (1999–2003), is less than that estimated for 2003–2005 (see Figure 9). This shift introduces an inconsistency, in that the 2007 biomass distribution shown in the projection results differs from that in the assessment.

Because of this inconsistency, and because the current approach does not appear to be biasing projection results, I conclude that it is better not to change the first random year in projections.

## 2.4 Reconsidering the observations of proportion spawning

Histological data from the three autumn sub-Antarctic surveys (TAN9204, TAN9304, TAN9805) have been used to estimate the proportion, by age, of female hoki that spawned in 1992, 1993, and 1998. Francis (2007) examined the three reports concerned with the calculation of these proportions spawning (Vignaux et al. 1995, Livingston et al. 1997, Livingston & Bull 2000) and found four points of concern:

1. The 1992 estimates for ages 3 and 5 are used in the assessment despite their being considered unreliable by Livingston & Bull (2000) and no comparable analysis of reliability was done for the 1998 estimates;
2. The decision not to use the 1993 estimates in the assessment because the 1993 survey was too late is inconsistent with the fact that these estimates are almost always higher than those for the other two years;
3. The method used to calculate the 1998 estimates appeared to differ from that used for the other years;
4. The 1998 estimates used data from a wider depth range (300–1000 m) than was used in the earlier surveys (300–800 m) and this may cause bias because preliminary analysis of subsequent data (Grimes & O’Driscoll 2005), which were from post-spawning observations, suggested that the proportion spawning may be lower for fish at greater depth (see figure 37 in Francis 2007).

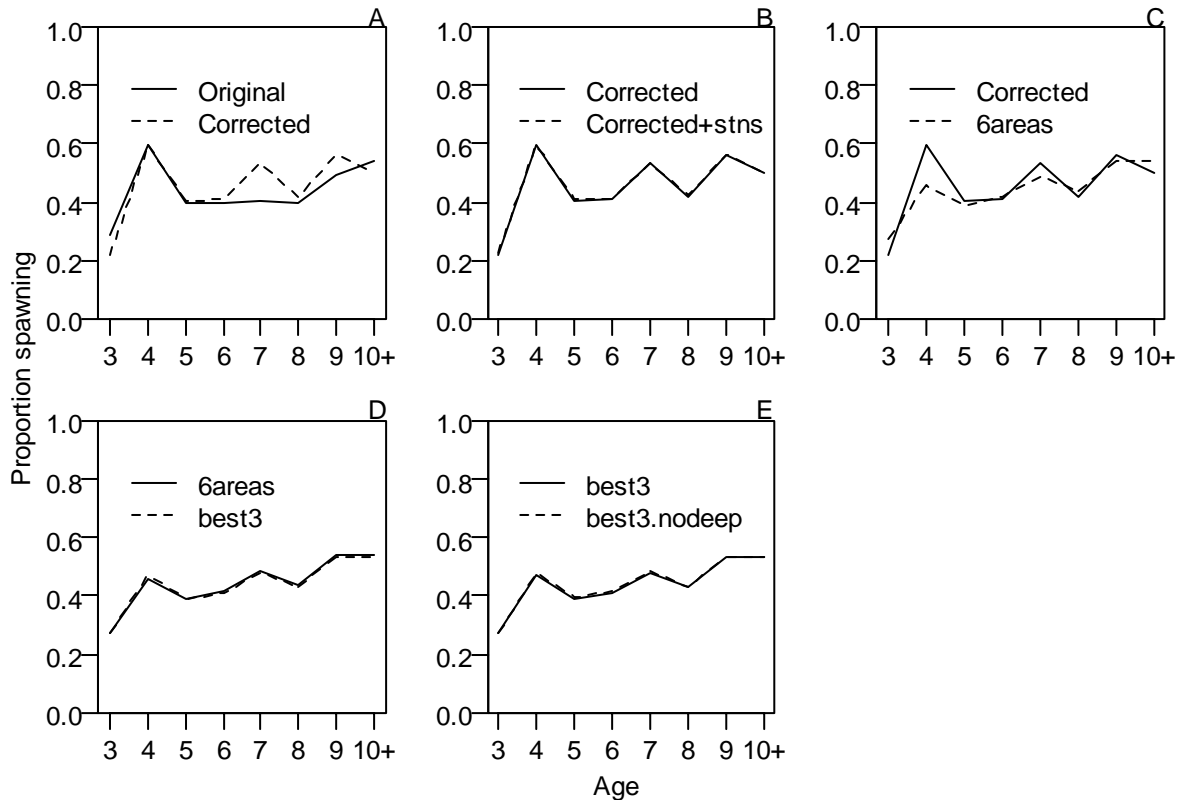
In this section I reanalyse some of these data sets with these points in mind.

### 2.4.1 Re-analysis of the 1998 data

I found both the 1998 data and the Splus code used to analyse them. From this code I reconstructed the equations used by Livingston & Bull (2000) which are as follows.

For fish of a given age,  $a$ , the proportion spawning (for females) was estimated as  $\sum_{is} (W_{sa} I_{isa}) / \sum_{is} W_{sa}$ , where  $s$  indexes strata,  $i$  indexes the females of age  $a$  in stratum  $s$  that were sampled biologically,  $W_{sa}$  is a stratum weight, and  $I_{isa}$  indicates whether a fish was expected to spawn ( $I_{isa} = 1$ ) or not ( $I_{isa} = 0$ ). The weights were supposed to be calculated as  $W_{sa} = A_s C_s P_{sa} / n_{sa}$ , where for stratum  $s$ ,  $A_s$  is the area ( $\text{km}^2$ ),  $C_s$  is the mean female catch rate ( $\text{no.km}^{-2}$ ),  $P_{sa}$  is the proportion of females that were of age  $a$ , and  $n_{sa}$  is the number of females of age  $a$  in the biological sample from the stratum.

Unfortunately, there was an error in the Splus code, and the  $n_{sa}$  used in this calculation was the total number of females (of all ages) in the biological sample from stratum  $s$ . Correcting this error made a visible difference to the estimates (Figure 11A).



**Figure 11: Pairwise comparisons of different estimates of proportion spawning (as described in Table 6) in 1998.**

**Table 6: Description of various alternative estimates of proportion spawning in 1998 (see text for more details).**

Label	Description
Original	As in Livingston & Bull (2000)
Corrected	With corrected $n_s$
Corrected+stns	With corrected $n_s$ and including stations 1, 74, 76, 77
6areas	Assuming proportion spawning does not vary within each of 6 regions
best3	Assuming proportion spawning does not vary within each of 3 regions
best3.nodeep	As for best3, but ignoring deep strata.

The catch rates,  $C_s$ , were calculated by taking the female catch rate, in  $\text{kg.km}^{-2}$ , at each station in the stratum, dividing by the mean weight of females at that station (calculated from the LF sample using length-weight parameters  $6.291 \times 10^{-6}$  and 2.829721), and averaging across all stations in the stratum. The proportions at age,  $P_{sa}$ , were calculated from a stratum age frequency (AF). This AF was constructed by applying an age-length key (calculated from all aged females with 3 cm length bins) to the female length frequency (LF) sample for each station to get an AF for the stn, and then summing AFs across all stations in stratum  $s$ . There were two potential problems with the calculation of  $C_s$  and  $P_{sa}$ , though neither had much effect on the final estimates of proportion spawning.

The first potential problem is that the analysis ignored four valid biomass stations:

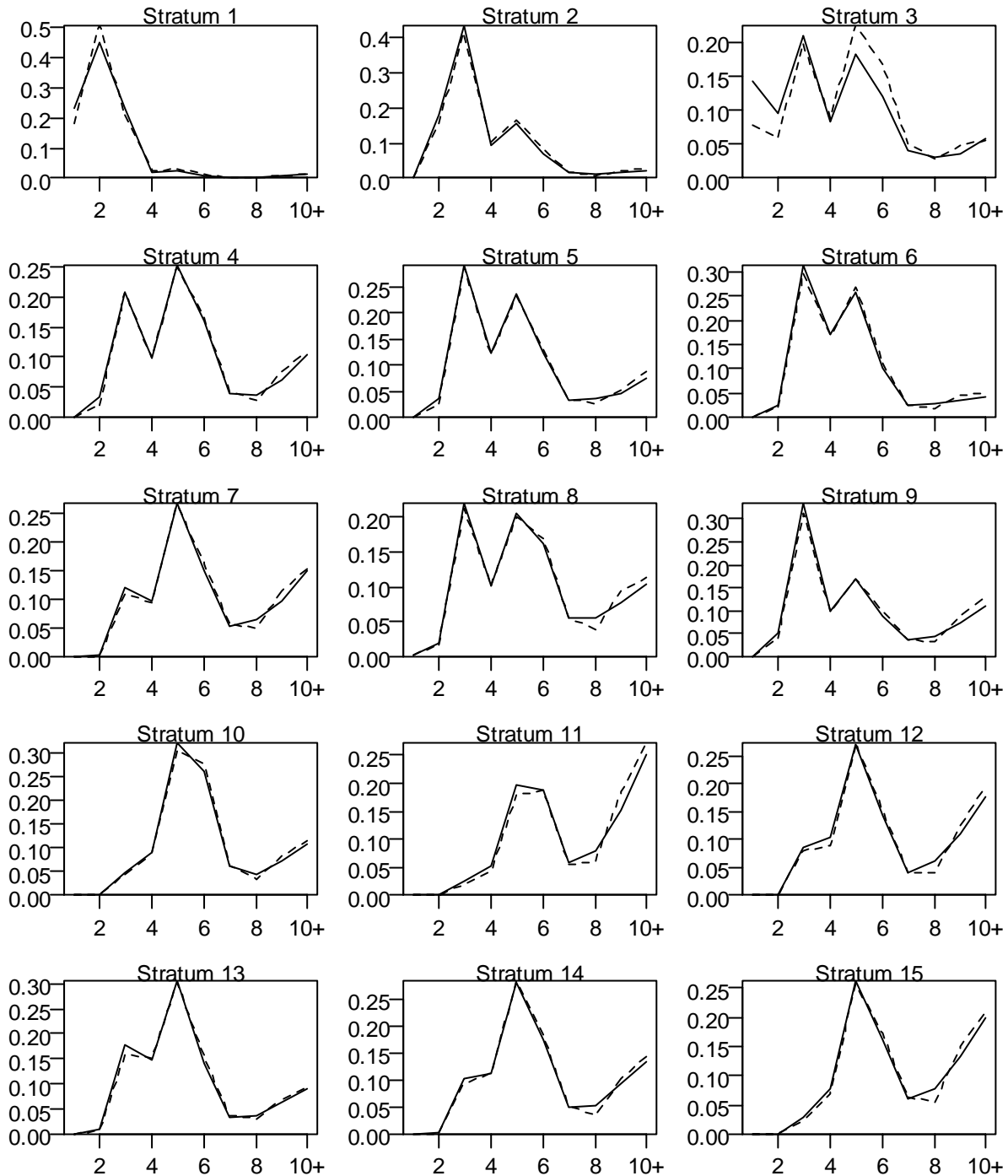
- station 9 in stratum 1, where no hoki were caught, and
- stations 74, 76, 77 in stratum 27, where there were said to be no female biological samples (according to the *trawl* database there were 2 female biologicals for station 74, but perhaps no histology was done for them).

Theoretically, all four stations should have been included in the calculation of the catch rates,  $C_s$ , and the last three should have been included in the calculation of the AF for stratum 27 (five other stations



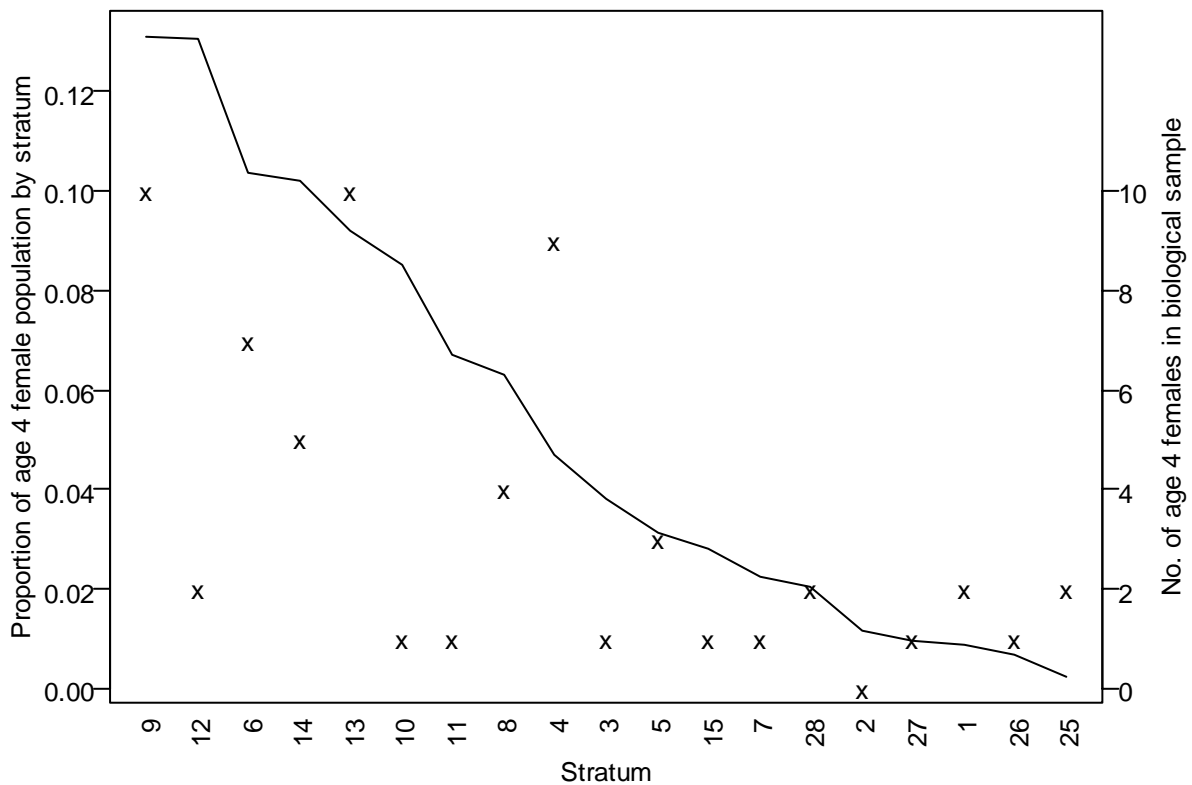
were included for this stratum). However, inclusion of these stations made very little difference to the estimates (Figure 11B).

The second potential problem concerns the calculation of the stratum AFs where, contrary to normal practice, no adjustments were made for area swept or sampling fraction (the proportion of the female hoki catch that was included in the LF sample). However, the effect of making these adjustments was slight, except for ages 5 and 6 in stratum 3 (differences at ages 1 and 2 are of no concern because the biological samples for these ages were too small to be useful) (Figure 12). These differences will not be very influential because only about 3% of 5- and 6-year old fish were in stratum 3.



**Figure 12: Comparison of stratum female age frequencies (AFs) estimated by Livingston & Bull (2000) (solid lines) with those estimated using the catch-at-age software (broken lines) which correctly adjusts for area swept and sampling fraction.**

A weakness in the above approach may be seen if we rewrite the above formula for proportion spawning as  $\sum_s (N_{sa} p_{sa}) / \sum_s N_{sa}$ , where  $N_{sa}$  is an estimate of the number of females of age  $a$  in stratum  $s$  in the autumn of 1998 and  $p_{sa}$  is an estimate of the proportion of those that would spawn in the following winter, calculated as  $p_{sa} = n'_{sa} / n_{sa}$ , where  $n'_{sa}$  is the number of spawners in the biological sample. The  $N_{sa}$  are relatively well estimated from the length samples ( $n = 4361$ ) and the age-length key. The problem is that 152 separate estimates of  $p_{sa}$  are required – one for each combination of the eight ages (3 to 10+) and 19 strata – and these are based on the relatively small female biological sample ( $n = 816$ ). Small sample sizes made many of these estimates very weak, as is illustrated for age 4 in Figure 13. Note that there were no females of this age in stratum 2, so this stratum had to be ignored in calculating the overall proportion spawning at age 4 (this happened for 12 of the 152 age-stratum combinations). Also, there were seven strata in which the estimate of  $p_{s4}$  was based on just one fish, and this included stratum 12 which, according to the AFs, contained about 13% of the population of age 4 females. It seems hard to believe that there is sufficient information in the biological data to justify the estimation of so many parameters.



**Figure 13: Estimated proportion,  $P_s$ , of the age 4 female population that was in each stratum in 1998 (line, left axis) and number of age 4 females in the biological sample from each stratum (points, right axis).**

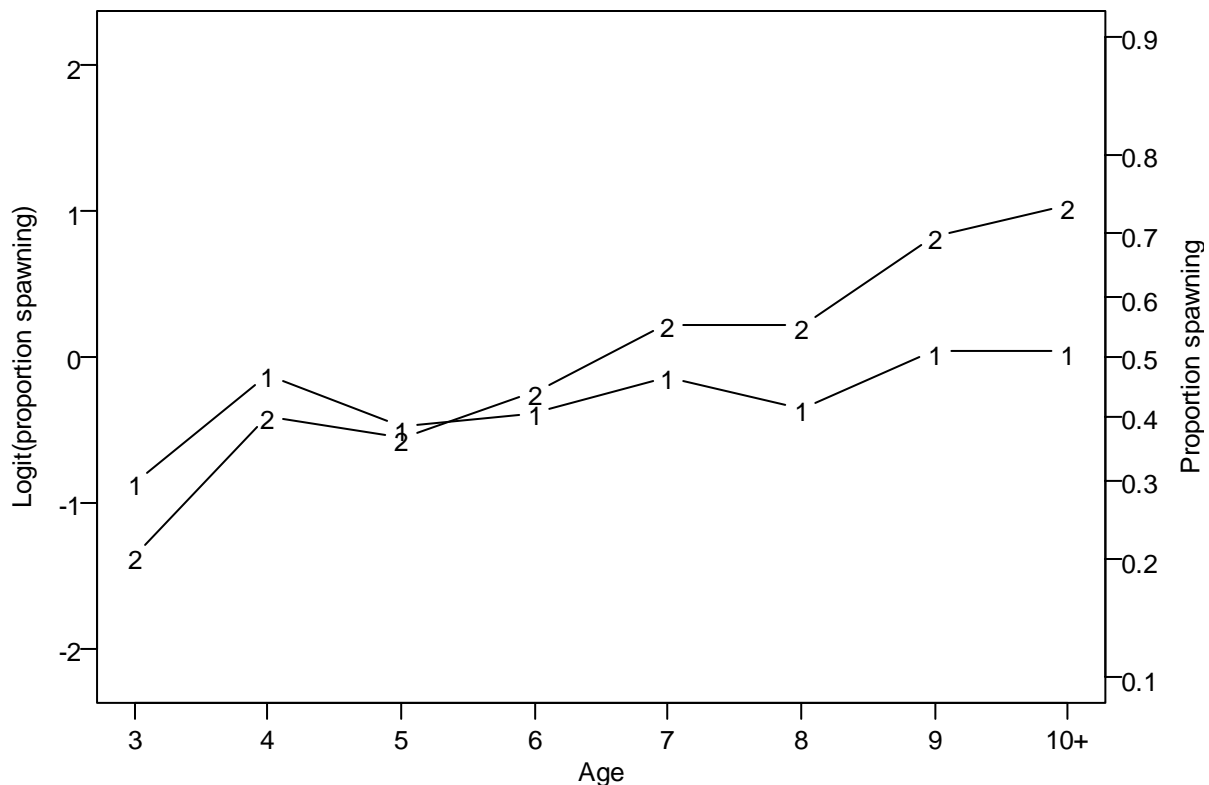
The way to reduce the number of parameters estimated is to add some assumptions. Two obvious assumptions to make are

- for a given age  $a$ ,  $p_{sa}$  is the same in ‘similar’ strata, and
- the relationship between  $p_{sa}$  and  $p_{s,a+1}$  is similar in different strata.

These assumptions were implemented for the 1998 survey data as follows. First, the strata were grouped into ‘regions’, where each region was a group of strata that were ‘similar’ in some way. Then it was assumed that  $p_{ra}$ , the proportion spawning at age  $a$  in the  $r$ th region, was given by

$$\text{logit}(p_{ra}) = c_{r1} + c_{r2} \text{logit}(p_{1a}) \quad (1)$$

where  $c_{r1}$  and  $c_{r2}$  are parameters to be estimated. (The logit transformation, defined by  $\text{logit}(p) = \log(p/(1-p))$ , is used here simply to ensure that  $p_{ra}$  always lies between 0 and 1). This assumption, illustrated in Figure 14, reduces the number of parameters to be estimated from (nominally) 152 to  $2m + 6$ , where  $m$  is the number of regions. [I also investigated a simpler model in which  $c_{r2} = 1$ , which means that the lines in Figure 14 would be parallel, but this was always inferior to equation (1)].



**Figure 14: Illustration of the assumptions about the proportions spawning in different ‘regions’ (where each region is a group of strata). The lines show the proportions spawning at age, plotted on a logit scale (left axis), in two regions (‘1’ and ‘2’). The assumption, embodied in equation (1), is that the gap between the two lines (in logit units) is a linear function of age. For comparison, the right axis shows the proportions in the natural scale.**

I tried two ways of grouping the strata: geographic and bathymetric (Table 7). For each, the  $p_{sa}$  were estimated by maximum likelihood and the model fits were compared to that of the original method using the AIC (Akaike 1974) [note that the log-likelihood for each model is given by  $\sum_{sa} (n'_{sa} \log(p_{sa}) + (n_{sa} - n'_{sa}) \log(1 - p_{sa}))$ ]. Both groupings were better than the original method, but the geographical grouping was slightly better (upper part of Table 8). Compared to the ‘Corrected’ estimates of proportion spawning, those from the ‘6areas’ vary much less between adjacent ages and the high estimate at age 4 is much reduced (see Figure 11C).

It seemed possible that an even better model could be found by reducing the number of areas. I investigated all possible ways of doing this by grouping areas (there are 202 possibilities!) and found that the best result ('best3') was obtained with just three groups of areas (middle part of Table 8). The overall estimates of proportion spawning differ very little between the '6areas' and 'best3' models (see Figure 11D), but the latter model is made more plausible by the fact that the areas that are grouped are adjacent, forming a northwest group (Puysegur and Stewart-Snares Shelf) a central group (Pukaki, NWC and SWC) and the single area SEC in the southeast. The estimated proportions spawning differ markedly between these three areas, decreasing towards the northwest (Figure 15A). Results from the bathymetric grouping of strata (the '3depths' model) show higher proportions spawning in shallower water (Figure 15B).

**Table 7: Two ways of grouping the strata from the 1998 survey into 'regions' (the geographic regions are from table 1 of Livingston & Bull (2000)).**

Geographic regions			Bathymetric regions		
Label	Description	Strata	Label	Description	Strata
NWC	Northwest Campbell Plateau	5 6 8 9	Shallow	300–600 m	1 3 6 9 12 13 14
Puk	Pukaki Rise	11 12 27	Medium	600–800 m	2 4 5 7 8 10 11 15
Puys	Puysegur Bank	1 2 25	Deep	800–1000 m	25 26 27 28
SEC	Southeast Campbell Plateau	13 14 15			
SSS	Stewart-Snares Shelf	3 4 28			
SWC	Southwest Campbell Plateau	7 10 26			

**Table 8: Comparison of methods of estimating  $p_{sa}$ , the proportion of females of age  $a$  from stratum  $s$  that will spawn in the following winter. The best method is that with the lowest AIC.**

Method	Description	Parameters estimated	AIC	
Original	As in Livingston & Bull (2000)	140	1155.1	
6areas	Strata grouped geographically into 6 areas	18	1089.9	
3depths	Strata grouped bathymetrically into 3 depth ranges	12	1090.5	
				Groups of areas
best1	6 geographical areas grouped into 1 group	8	1101.7	NWC.Puk.Puys.SEC.SSS.SWC
best2	6 geographical areas grouped into 2 groups	10	1085.6	NWC.Puk.SEC.SWC Puys.SSS
best3	6 geographical areas grouped into 3 groups	12	1084.2	NWC.Puk.SWC Puys.SSS SEC
best4	6 geographical areas grouped into 4 groups	14	1085.4	NWC.Puk.SWC Puys SEC SSS
best5	6 geographical areas grouped into 5 groups	16	1087.3	NWC.Puk Puys SEC SSS SWC
best3.free	As in best3 but no constraints between areas	24	1085.0	as for best3

We should remember that the apparently clear trend in proportion spawning from northwest to southeast (see Figure 15A) is in part an artefact of our model assumptions. In particular it is caused by our assumption (embodied in equation (1) and illustrated in Figure 14) that the three lines in this plot have the same shape. For model 'best3.free', I dropped that assumption but still used the three areas of model 'best3'. This model is not quite as good (in terms of AIC) as 'best3' (see last line, Table 8), but it gives us a clearer picture of what we can reliably infer about the between-area differences in proportion spawning (Figure 16). The sample sizes used to calculate standard errors in Figure 16 ranged from 5 to 124, with median 27 (Table 9).

**Table 9: Sizes of biological samples, by age and area, for models best3 and best3.free.**

Area	Age							
	3	4	5	6	7	8	9	10+
SEC	5	16	28	26	8	9	12	31
NWC.Puk.SSC	46	31	124	83	30	35	65	83
Puys.SSS	23	16	40	41	11	12	15	26

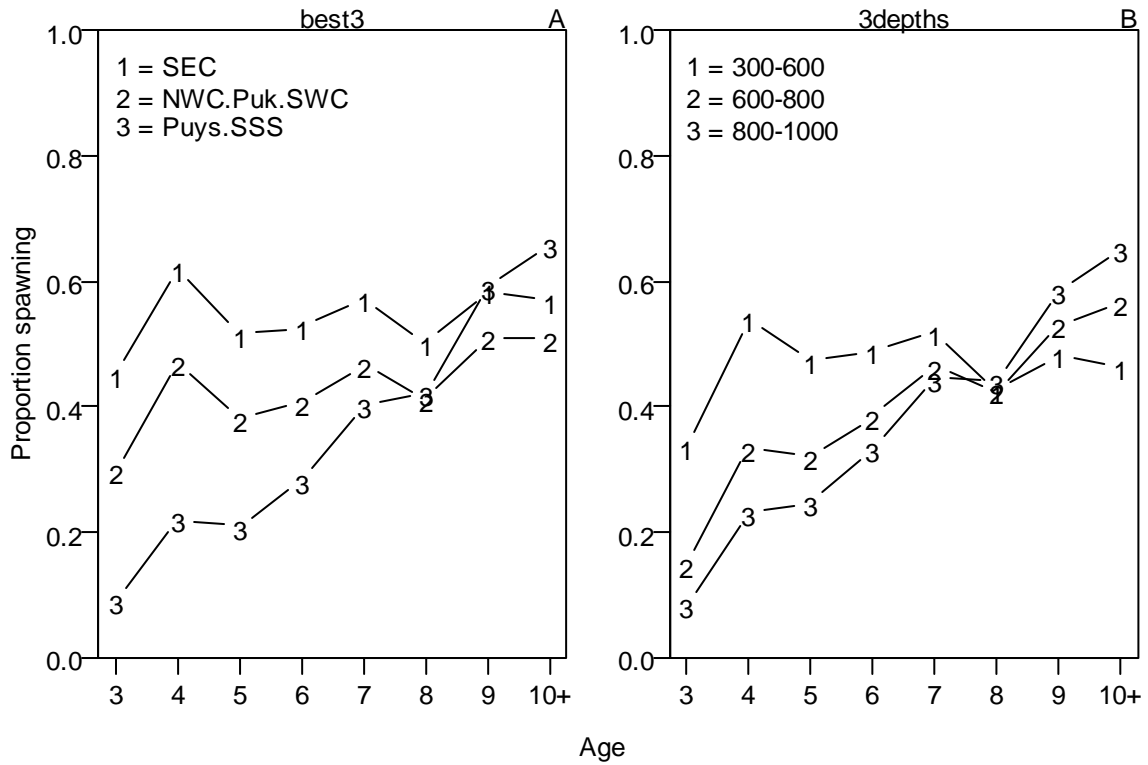


Figure 15: Estimates of proportion spawning by region for two of models in Table 8: A, best3; B, 3depths.

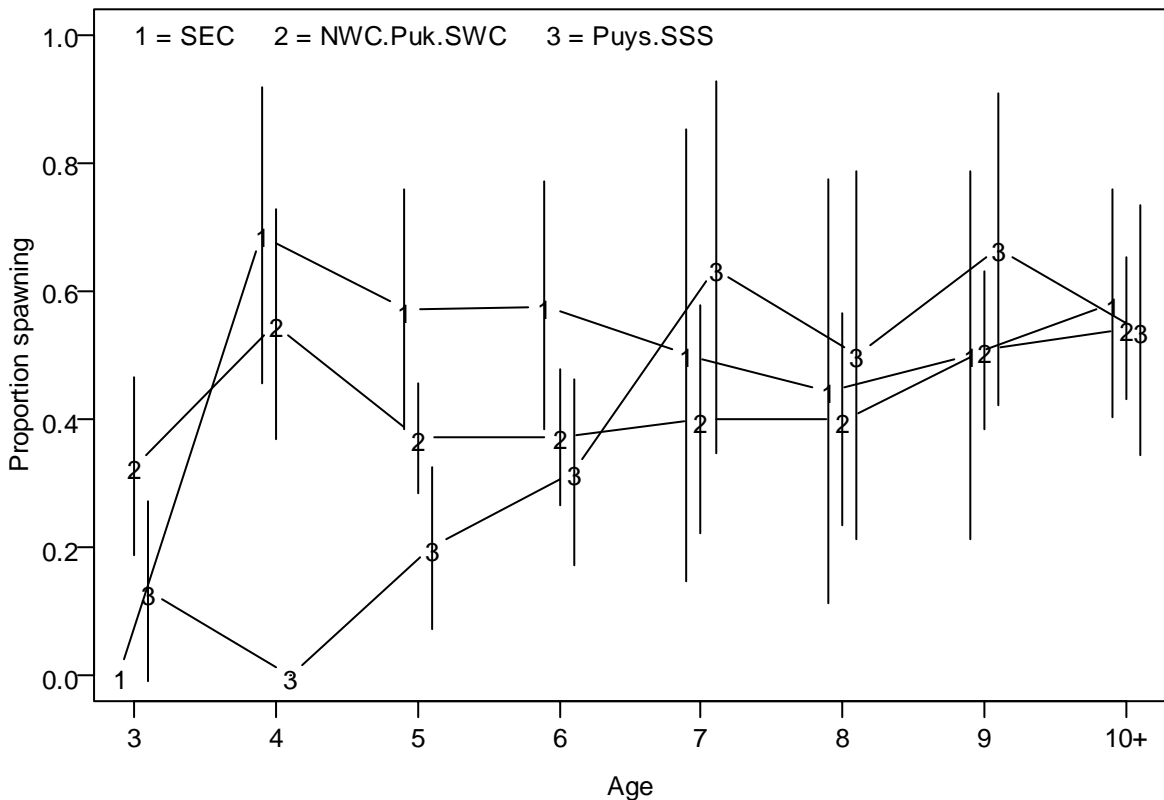


Figure 16: Estimates of proportion spawning by region and age for model 'best3.free'. Vertical bars are approximate 95% confidence intervals (i.e., +/- 2 s.e.s).

The final question to be addressed for the 1998 survey concerns the deep strata (800–1000 m), which were included in this survey but not in those from 1992 and 1993. Did this cause the 1998 estimates of proportion spawning to be biased relative to the earlier estimates? To answer this question I calculated the estimates labelled ‘best3.nodeep’ using the same equation as before,  $\sum_s (N_{sa} p_{sa}) / \sum_s N_{sa}$ , and the same values of  $N_{sa}$  and  $p_{sa}$  as for ‘best3’, but excluding the deep strata from the summations. This made very little difference to the estimates (see Figure 11E), so I conclude that any bias was minimal.

#### 2.4.2 Re-analysis of the 1992 and 1993 data

Because these data were not found in time they were not reanalysed before the assessment (but see Section 7.3 for a post-assessment analysis).

#### 2.4.3 Re-analysis of the post-spawning data

I analysed the post-spawning data of Grimes & O’Driscoll (2005) using the same approach as applied to the 1998 pre-spawning data. This data set contains only 286 histological samples, and so is only about one third the size of the earlier one (for which  $n = 816$ ). For the post-spawning data, the two best models grouped the data by depth (model 3depths), and in three groups of geographical areas (model best3) (Table 10).

The results show some similarity with those from the 1998 data. The trend with depth is in the same direction in both cases, although the estimates for the deep strata are much more extreme with the more recent data (compare Figure 17A and Figure 15A). The geographical grouping for the two data sets are similar except that area SEC had a high proportion of spawners in 1998 and a medium proportion in the later data (compare Figure 17B and Figure 15B). The geographical groups produced for the post-spawning data are less plausible than those for the 1998 data because one group, SSS.SEC, involves non-contiguous areas.

**Table 10: Comparison of methods of estimating  $p_{sa}$ , the proportion of females of age  $a$  from stratum  $s$  that spawned in the preceding winter. The best method is that with the lowest AIC.**

Method	Description	Parameters estimated	AIC	
Free	As in Livingston & Bull (2000)	96	368.35	
6areas	Strata grouped geographically into 6 areas	18	302.58	
3depths	Strata grouped bathymetrically into 3 depth range	12	287.68	
				Groups of areas
best1	6 geographical areas grouped into 1 group	8	314.90	NWC.Puk.Puys.SEC.SSS.SWC
best2	6 geographical areas grouped into 2 groups	10	296.44	NWC.Puk.SWC Puys.SSS.SEC
best3	6 geographical areas grouped into 3 groups	12	296.36	NWC.Puk.SWC Puys SSS.SEC
best4	6 geographical areas grouped into 4 groups	14	297.24	NWC.Puk Puys SSS.SEC SWC

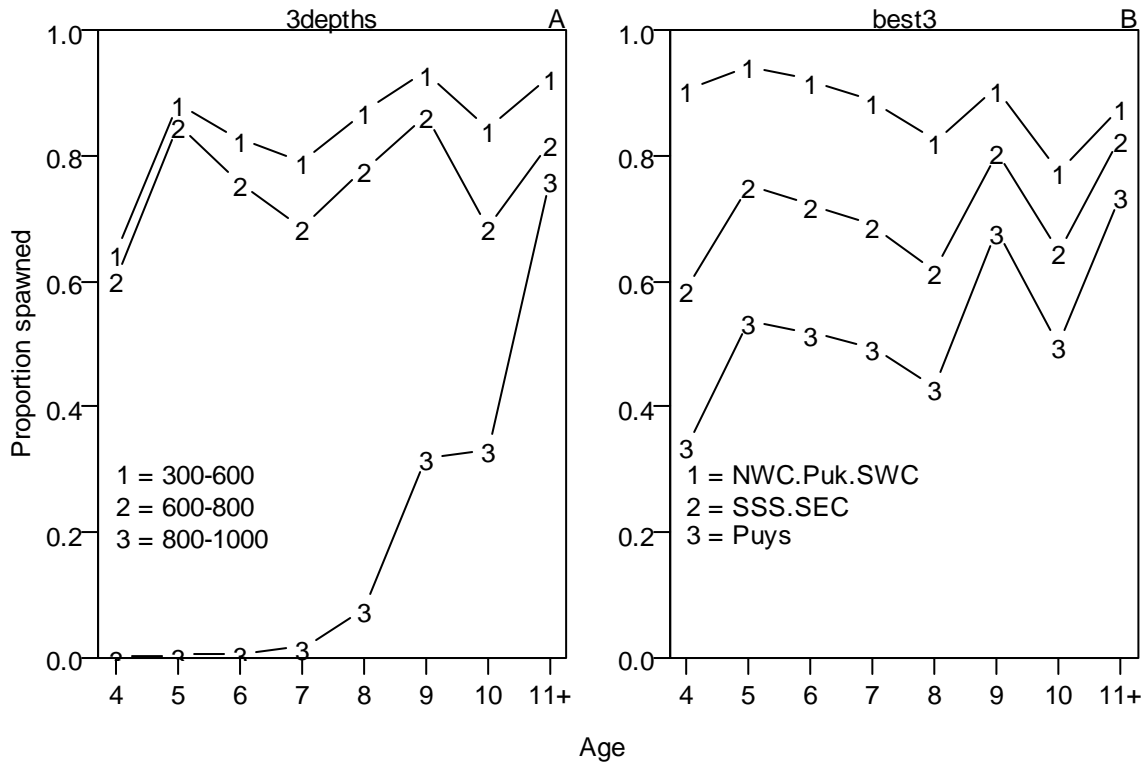


Figure 17: Estimates of proportion spawned by for the two best models of Table 10. Note that the proportions spawned at ages 4–11+ in this plot are comparable to the proportions spawning at ages 3-10+ in Figure 15.

#### 2.4.4 Conclusions concerning the proportion spawning data

The new approach to estimating proportion spawning seems more robust than, and thus preferable to, that used previously. It should be applied to the 1992 and 1993 data, if these can be found.

Meanwhile, we can ask what has been resolved concerning the four points of concern listed at the beginning of Section 4. As to the first point (unreliability of 1992 estimates for ages 3 & 5), the additional assumptions used in the new estimation method should remove (or at least reduce) this concern. The second point (the inconsistency of the decision not to use the 1993 estimates) has not been addressed. The third point (possible methodological differences between data sets) will be addressed if we are able to find the earlier data. Finally, the influence of the deep strata (the fourth point) has been shown to be minimal.

[The 1992 and 1993 data were found after the assessment was completed. A reanalysis of these data, and further conclusions about proportion spawning are contained in Section 7.3 below.]

## 2.5 Investigating stock-recruit steepness

In a recent review of the use of MSY-based reference points in New Zealand, the reviewers suggested that profiling be used to understand what information there is in stock assessments about the steepness of the stock-recruit relationship (which is defined as the proportion of mean virgin recruitment that is produced when the spawning biomass is reduced to 20% of its unfished level). In this section, I review the values of steepness that have been used in previous hoki assessments, construct steepness profiles based on the 2007 assessment, look at the effect of steepness of the assessment, and consider some ancillary information that might be useful in reconsidering the best value to use for hoki.

### 2.5.1 Historical values of steepness for hoki

Three fixed values of steepness have been used in New Zealand hoki assessments: 0.95, 0.75, and 0.9 (in order of use). The first value was first used in 1990 (Sullivan & Cordue 1990), or possibly in 1989 (Sullivan & Coombs (1989) used a Beverton-Holt stock-recruit relationship, but did not give its parameters). This was reduced to 0.75 in 1992 (Sullivan & Cordue 1992, Pikitch et al. 1993), presumably following the recommendation of Francis (1992) concerning default steepness values. Finally, a steepness of 0.9 has used been used since the 1994 assessment (Cordue 1994), and this was derived from a steepness prior constructed by Punt et al. (1994) using estimates from 45 stocks of gadiform species (0.9 is approximately the median of that prior, which is discussed further below in Section 2.5.4).

The only assessment where a prior distribution has been used for steepness, rather than a fixed value, seems to be that of Punt et al. (1994), who did not describe the resulting posterior for this parameter.

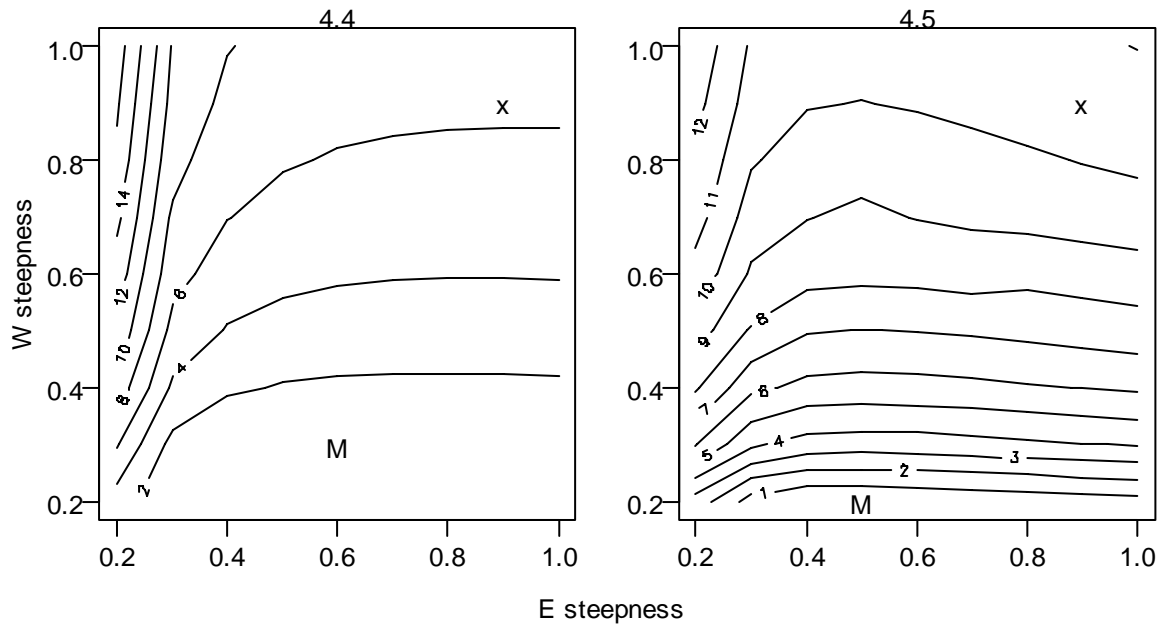
In some assessments, sensitivity analyses considered alternative values of steepness, but these were always 0.75 and/or 0.95.

### 2.5.2 Profiling steepness

I constructed posterior profiles for steepness for 2007 runs 4.4 and 4.5 using a no-information prior distribution which was uniform on the interval (0.2,1) (run 4.7, was not considered because the stock-recruit relationship is not well determined for this run). Best fits were found for all combinations of E and W steepness over the set of values 0.2, 0.3,..., 1.0. For both runs, the resulting joint posterior profiles show that the ‘best’ estimates of steepness were considerably lower than the value of 0.9 assumed in the assessment – 0.6 (in run 4.4) or 0.5 (4.5) for the E stock, and 0.3 or 0.2, respectively, for the W stock (Figure 18). If we assume that steepness is the same for both stocks the ‘best’ value is still very low: 0.3 for run 4.4, and 0.2 for 4.5.

The profiles are quite flat. That is, the data and priors do not provide strong evidence to prefer the lower values of steepness. They decrease the objective function by only 6 or 10 points (for runs 4.4 and 4.5, respectively), and only about half of this decrease comes from observations (Table 11).



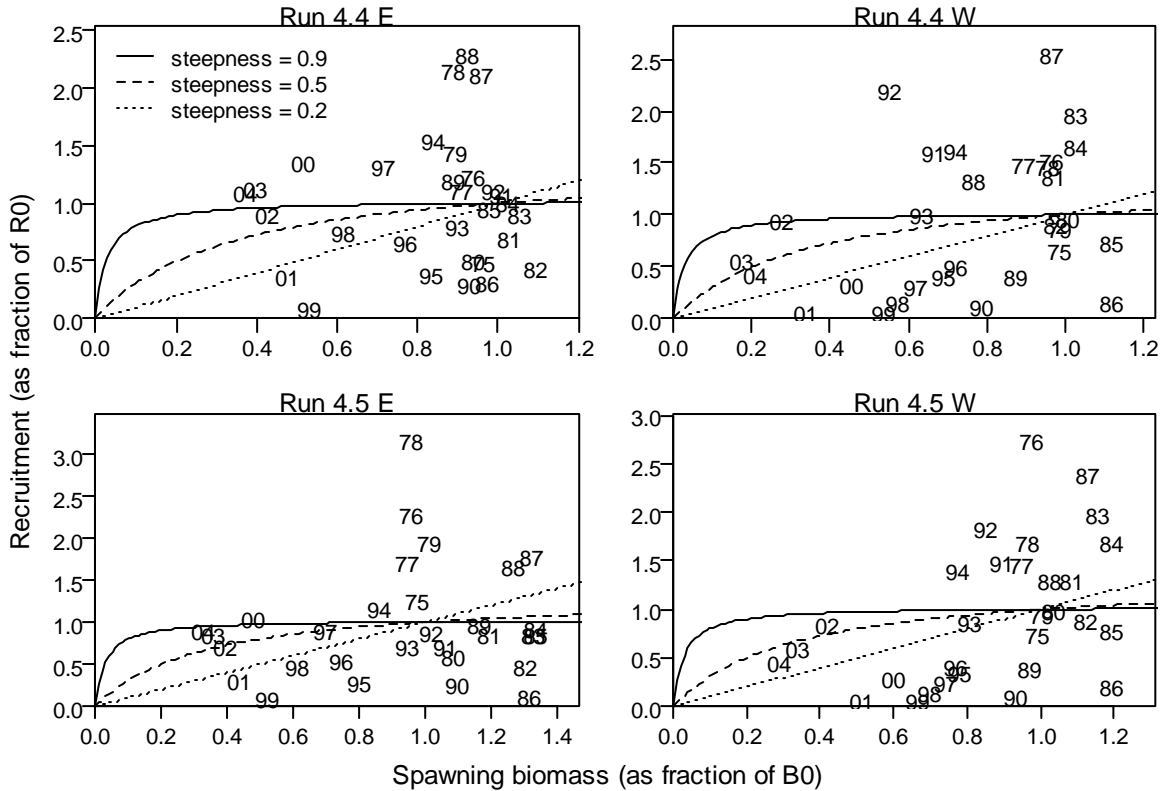


**Figure 18: Joint posterior profiles of the stock-recruit parameter steepness for the E (x-axis) and W (y-axis) stocks in runs 4.4 (left panel) and 4.5 (right panel). The contour lines are relative to value of the profiles at the maxima, which are plotted as ‘M’, and ‘x’ indicates the steepness values assumed in the stock assessments (0.9 for both stocks).**

**Table 11: Gain (decrease in objective function), by component, when ‘best’ values of steepness are used instead of the assumed value of 0.9. Gains are rounded to the nearest integer, and components with zero gain for both runs are omitted from the table.**

Component of objective function	4.4	4.5
SAsumbio	1	2
CRsumage	0	-1
SAsumage	2	2
Enspage	0	-3
Espage	1	2
Wnspage	-2	-1
Wspage	1	3
prior_on_Bmean_prop_stock1	1	0
prior_on_selectivity[Wspsl].shift_a	1	0
prior_on_recruitment[W].YCS	2	5
prior_on_Mmale	-1	0
All	6	10

Stock-recruitment plots based on the 2007 assessment show why the above profiles suggest that the assumed value of steepness was too high. For the W stock, all year classes since 1995 lie near or below the assumed stock-recruit curve, with steepness 0.9, and this is true for most year classes for the E stock (Figure 19).



**Figure 19:** Plots of recruitment against spawning biomass for runs 4.4 (upper panels) and 4.5 (lower panels). Each plotted point relates to a year class (with the plotting symbol being the last two digits of the year it was spawned) and shows its strength (on the vertical axis) and the biomass of the spawners that produced it (on the horizontal axis). Also shown (as lines) are Beverton-Holt stock-recruitment curves with steepnesses of 0.9, 0.5, and 0.2.

### 2.5.3 The effect of steepness

I evaluated the effect of steepness on the assessment by considering how the 2007 assessment results would have differed if the ‘best’ values of steepness from the above profiles had been used in place of the assumed value of 0.9. The effect of these lower steepness values is clearly negative for the W stock – both in terms of current biomass (Figure 20) and projections (Figure 21) – but slight for the E stock. It is of interest that, for both runs, the new MPD estimates of current biomass for the W stock lie below the range of the 95% confidence intervals calculated using the assumed values of steepness (Table 12).

**Table 12:** Comparison of estimates of the current biomass of the W stock (expressed as % $B_0$ ) from the 2007 assessment (where steepness was assumed to be 0.9 for both stocks) with those using the ‘best’ values of steepness.

Run	From the 2007 assessment		Using ‘best’ steepness
	MPD estimate	95% confidence interval	MPD estimate
4.4	16	(13,32)	11
4.5	23	(19,31)	17

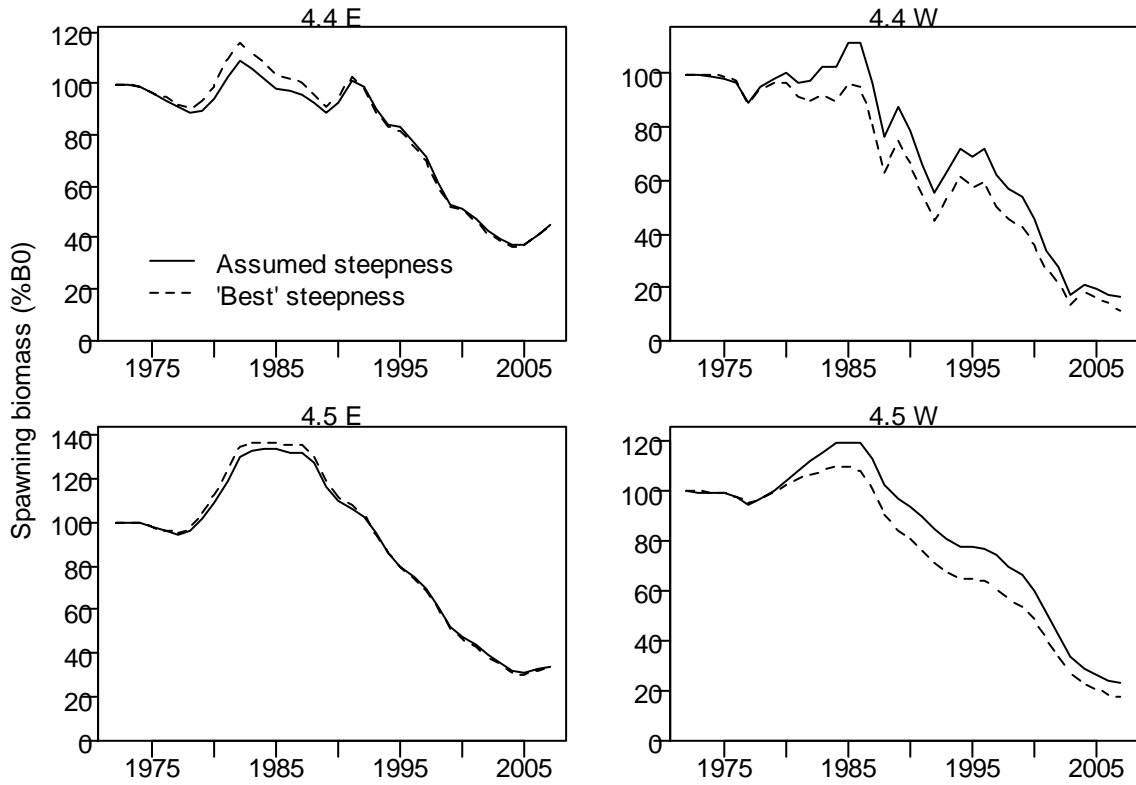


Figure 20: Effect on estimated biomass trajectories of using the 'best' values of steepness for each stock, rather than the assumed values (0.9 for both stock) for runs 4.4 (upper panels) and 4.5 (lower panels).

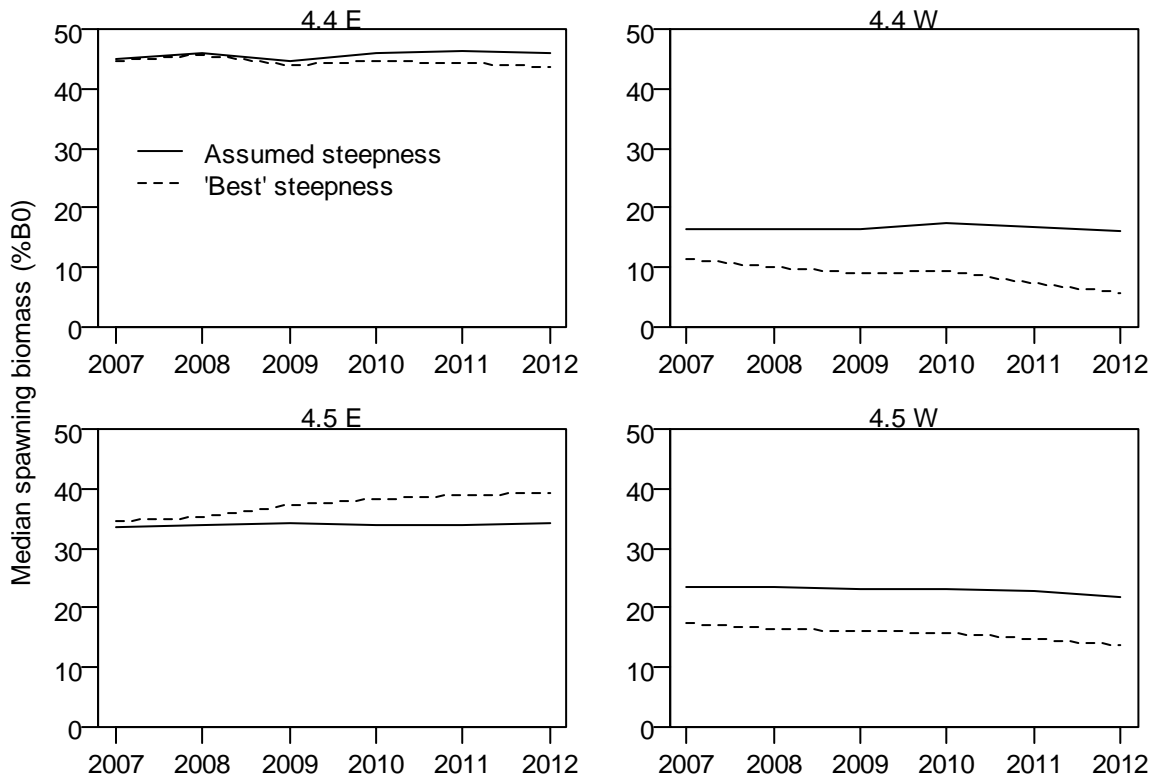


Figure 21: Effect on projected spawning biomass of using the 'best' values of steepness for each stock rather than the assumed values (0.9 for both stock): MPD-based projection results with status-quo catches (106kt55%E) and recent recruitment.

## 2.5.4 Ancillary information on steepness

Two papers by Ram Myers and colleagues provide some useful ancillary information. Both use the extensive spawner-recruit database described by Myers et al. (1995) (see also <http://fish.dal.ca/welcome.html>). This database was also used by Punt et al. (1994) to construct the steepness prior mentioned above in Section 2.5.1. Myers et al. (1999) provided more recent estimates of steepness which are markedly lower, typically by about 0.2 (Table 13). This reduction appears to be primarily a result of a change in the estimation methodology, which raises the issue of bias.

Myers et al. (1999) describe their estimates of steepness as ‘conservative’, which is, perhaps, a polite way of saying ‘negatively biased’. However, the same estimation procedure was used in the other paper I discuss below (Myers et al. 2002), despite the need, expressed by the authors, to “obtain objective priors of steepness that would be rigorous enough for a court of law”. Thus these authors clearly believed that their methodology was more defensible than the original one. Further, I wonder whether the original estimates procedure may have been biased in the other direction, given that almost half (20 of 45) of the steepness estimates used by Punt et al. (1994) were exactly 1.000.

Another issue of interest with these data is taxonomic. Many authorities [e.g., FAO Species Catalogue (Cohen et al. 1990), ITIS (<http://www.itis.gov/index.html>), MarBEF (<http://www.marbef.org/data/index.php>), and FishBase (<http://www.fishbase.org/search.php>)] include the genera *Merluccius* and *Macruronus* (which includes hoki) in the hake family (Merlucciidae). However, some [e.g., Encyclopædia Britannica Online (<http://www.britannica.com>)] classify them as part of the cod family (Gadidae) [as did Myers et al. (1999)]. This is relevant because the few estimates of steepness for *Merluccius* species tend to be lower than those for the other gadiforms.

**Table 13: Median estimates, from two sources, of steepness for some stocks of gadiform species.**

Family	Species	Punt et al. (1994) <sup>1</sup>		Myers et al. (1999)	
		Number of stocks	Median steepness <sup>2</sup>	Number of stocks	Median steepness <sup>3</sup>
Gadidae	Blue whiting ( <i>Micromesistius poutassou</i> )	2	0.88	2	0.71
	Whiting ( <i>Merlangius merlangus</i> )	2	0.99	5	0.81
	Walleye pollock ( <i>Theragra chalcogramma</i> )	2	1.00	2	0.55
	Atlantic cod ( <i>Gadus morhua</i> )	20	0.97	21	0.84
	Haddock ( <i>Melanogrammus aeglefinus</i> )	8	0.93	9	0.74
	Pollock or saithe ( <i>Pollachius virens</i> )	5	1.00	5	0.81
Gadidae or Merlucciidae <sup>4</sup>	Pacific hake ( <i>Merluccius productus</i> )	1	0.59	1	0.32
	European hake ( <i>Merluccius merluccius</i> )	1	0.67	–	–
	Cape hake ( <i>Merluccius capensis</i> / <i>M. paradoxus</i> )	1	1.00	–	–
	Silver hake ( <i>Merluccius bilinearis</i> )	3	0.51	3	0.39
	(Argentine) hake ( <i>Merluccius hubbsi</i> )	–	–	1	0.82
	All gadiforms in Myers et al. (1999)	–	–	49	0.79

<sup>1</sup> Estimates (listed in table 4 of Punt et al. (1994)) taken from Myers et al. (1995); <sup>2</sup> Simple median; <sup>3</sup> Median of prior from mixed model; <sup>4</sup> There is not consensus as to which family includes the genus *Merluccius* – see text.

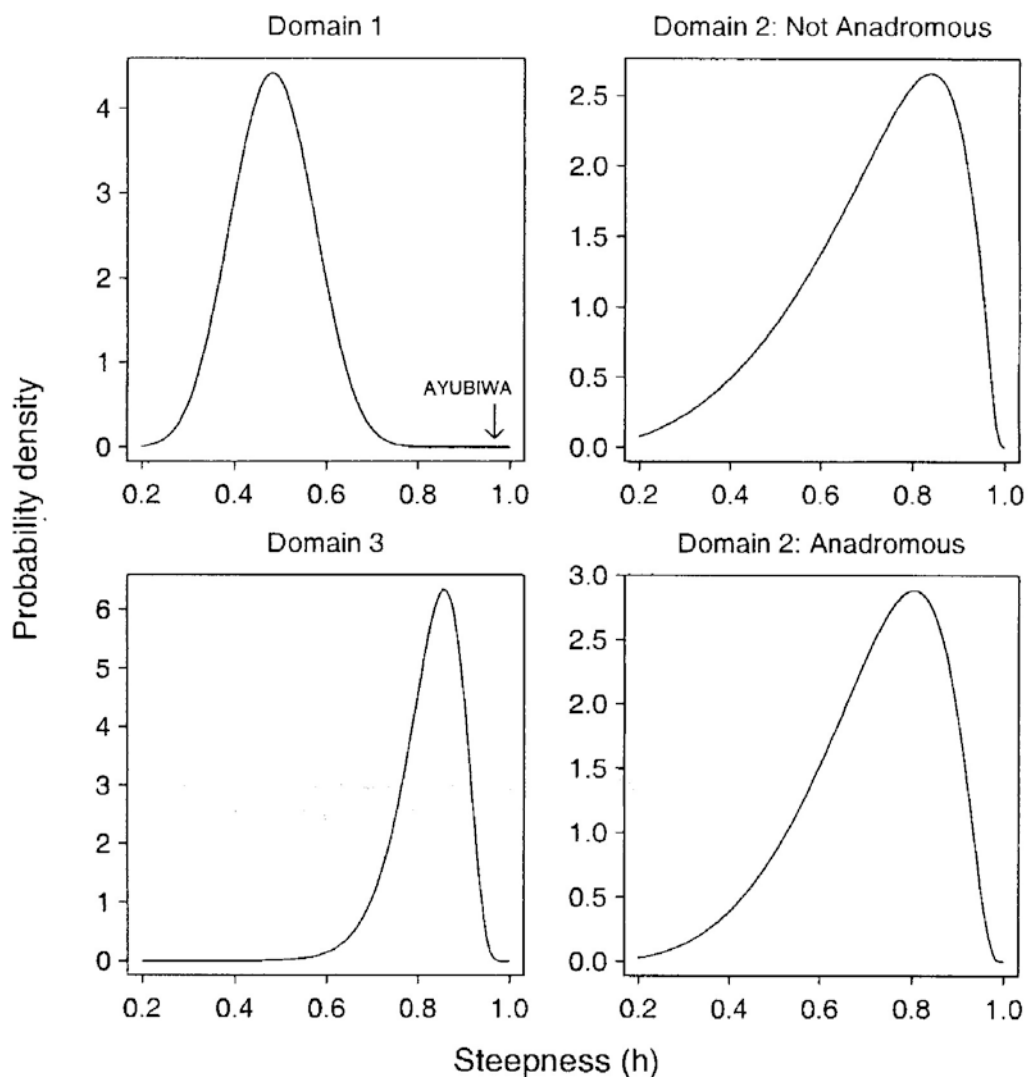
The second paper, by Myers et al. (2002), considered the problem of constructing steepness priors using a meta-analysis of data from 246 fish populations. Two of their analyses could be relevant for hoki. In one analysis, they used an expert (J.H. Cowan) to divide their populations into four groups on ecological grounds (using information on natural mortality, longevity, type of reproduction, habitat, fecundity, age at maturity, and environmental temperature). From their description of the distinguishing characteristics of these groups it seems clear that hoki can be excluded from two of them (groups 1 and 2A), but not from the other two (2NA and 3) (Table 14). Myers et al. constructed steepness priors for each of their four groups. The two that might be appropriate for hoki both have medians less than 0.9, but are quite different, with that for group 3 being mostly restricted between 0.7 and 0.95, while that for group 2NA allowing lower values (Figure 22). I sought further information

about the classification and now have a list of the stocks included in each of the four groups (J.H. Cowan, Louisiana State University, pers. comm.). A point of interest is that all five of the *Merluccius* stocks included in this analysis were classified in group 2NA.

**Table 14: Life history characteristics of four species groups defined in Myers et al. (2002), and associated information for hoki.**

Group/species	Age at maturity <sup>1</sup> (y)	Natural mortality (y <sup>-1</sup> )	Fecundity (millions of eggs)	Anadromous?
1	< 2	≥ 0.3	< 0.1	–
2A	2–5	0.2–0.5	0.1–0.75	Yes
2NA	2–5	0.2–0.5	0.1–0.75	No
3	> 4	< 0.2	> 0.5	–
Hoki	?	0.27 <sup>2</sup>	> 1 <sup>3</sup>	No

<sup>1</sup> Age at which 75% of females reach maturity; <sup>2</sup> Median value from run 4.5; <sup>3</sup> Schofield & Livingston (1998)



**Figure 22: Copy of figure 1 of Myers et al. (2002) showing the steepness priors they calculated for each of the four species groups of Table 14.**

The above steepness profiles (see Figure 18) assumed a uniform prior for steepness. I recalculated these profiles using the priors for groups 2NA and 3. The latter prior made a big difference to the profiles, with the ‘best’ values of steepness being 0.8 in all cases (Table 15).

**Table 15: Effect of steepness priors on the ‘best’ estimates of steepness (i.e., those at the maxima of the calculated profiles).**

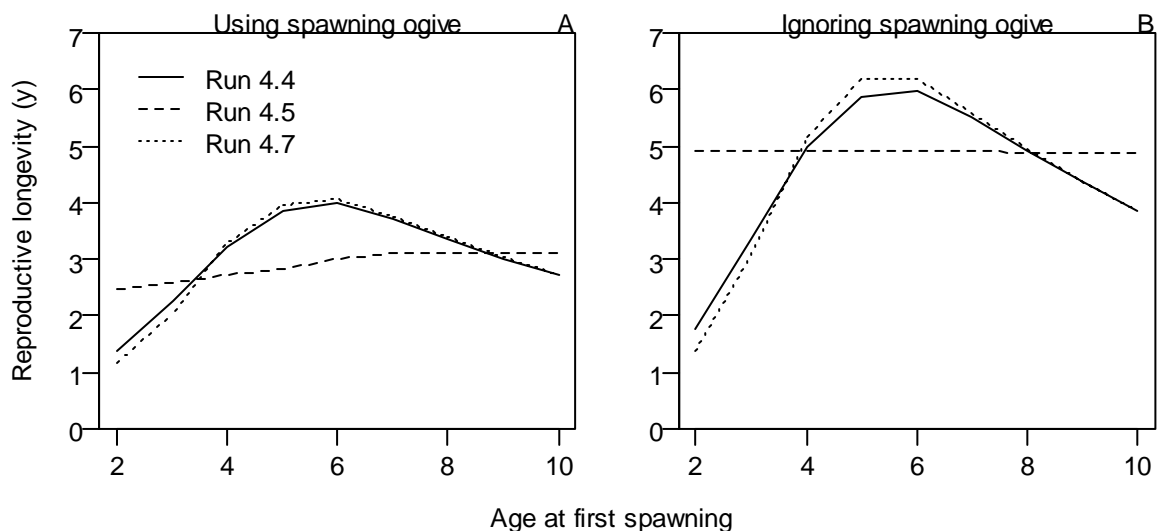
Prior	E stock		W stock		Combined <sup>1</sup>	
	Run 4.4	Run 4.5	Run 4.4	Run 4.5	Run 4.4	Run 4.5
Uniform	0.6	0.5	0.3	0.2	0.3	0.2
As for group 2NA	0.8	0.8	0.3	0.2	0.8	0.8
As for group 3	0.8	0.8	0.8	0.8	0.8	0.8

<sup>1</sup> Assuming equal steepness for the two stocks

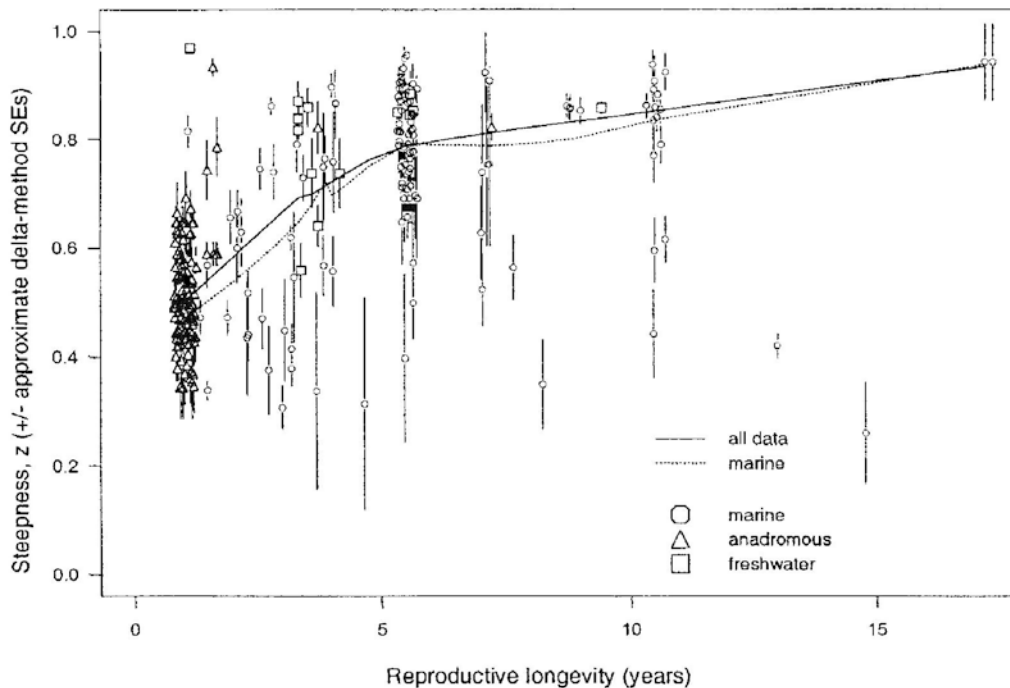
In a second analysis, Myers et al. (2002) considered a range of covariates, seeking those which were most strongly related to steepness. Two covariates are relevant to hoki – reproductive longevity and habitat type (i.e., marine, anadromous, or freshwater) – with steepness increasing with reproductive longevity, and being slightly lower for marine fish. Reproductive longevity was defined as the expected number of years of spawning after reproduction begins. The equation used by Myers et al. (2002) to calculate this quantity was very simple because it did not allow for either age-dependent natural mortality or the possibility that not all mature fish spawn each year. For hoki, I generalised their equation to the following:  $L_{a1} = 1 + \sum_{a>a1} \left[ P_a \exp\left(\sum_{a'=a1}^{a-1} M_{a'}\right) \right]$ , where  $L_{a1}$  is the reproductive longevity for a fish which spawns first at age  $a1$  and, for a fish of age  $a$ ,  $M_a$  is the (instantaneous) natural mortality and  $P_a$  is the probability that it will spawn (from the spawning migration ogive). I applied this equation using the parameter estimates from the 2007 assessment for females from the E stock (the calculations are more complex for the W stock because of the juvenile migration from CR to SA). The estimated reproductive longevity depended strongly on the age at first spawning. For example, for runs 4.4 and 4.7 it increased from about 1.3 y at age 2, to about 4 y at age 6, and then declined to 2.7 y for those fish which don't start spawning until age 10 (Figure 23A). If we assume that most female hoki start spawning between ages 3 and 6 then it seems likely that, roughly speaking, the average reproductive longevity lies between 3 y and 4 y.

These estimates will not be comparable to those of Myers et al. (2002) for any species which does not spawn annually. To allow for the possibility that there are many such species I redid my calculations ignoring the spawning ogive, and got values that were typically 1–2 y higher (Figure 23B). With these estimates, average reproductive longevity for hoki seems likely to lie between 5 y and 6 y.

The results of Myers et al. (2002) suggest that, for marine fish, a typical value of steepness rises from about 0.6 to about 0.8 as reproductive longevity increases from 3 y to 6 y (see dotted line in Figure 24).



**Figure 23: Estimates of reproductive longevity (expected number of years spawning after reproduction begins) for E female hoki as a function of their age at first spawning, the model run (4.4, 4.5, or 4.7): A, using the spawning ogive, and B, ignoring the spawning ogive.**



**Figure 24:** Copy of figure 3 of Myers et al. (2002): “Best linear unbiased estimates of steepness versus reproductive longevity for species with oviparous reproduction by habitat type (anadromous, freshwater, or marine). Lines are lowess fits”.

### 2.5.5 Conclusions on steepness for hoki

The above material suggested strongly that our assumed value of steepness is too high, and that a value between 0.6 and 0.8 would be more appropriate. The Hoki Working Group decided that a value of 0.75 should be used for the 2008 assessment, with sensitivity analyses to explore the effect of lower and higher values.

### 2.6 Catchability in TAN0714 survey

Because the biomass estimate from the last sub-Antarctic trawl survey (TAN0714, in December 2007) was higher by a factor of three than that from the previous survey (Figure 25), some concern was expressed, at a recent Working Group meeting, that the catchability of hoki may have been abnormally high at the time. To investigate that possibility, I examined the estimates of hoki numbers at age from all surveys in this series.

A simple plot comparing estimated numbers from the same cohort in consecutive years suggests that the catchability of hoki was markedly higher in 2007 than in 2006 (Figure 26A). From the comparison of just these two surveys we cannot say whether hoki catchability was unusually low in 2006 or unusually high in 2007 (or both). Comparisons involving surveys back to 2003 suggest the latter, because they show no strong catchability differences between adjacent surveys. The imbalance between the 2006 and 2007 surveys is similar in scale, but opposite in direction, to that shown between adjacent pairs of surveys between 2000 and 2003. However, the biomass of the W stock was estimated to have fallen substantially between 2000 and 2003 (Francis 2008b) so at least part of the imbalance shown in the earlier years would be due to changes in abundance, rather than catchability. Comparing catchabilities of surveys two years apart also supports the hypothesis that catchability was unusually high in 2007 (Figure 26B).

There is some reason to believe that apparently high catchability in the 2007 survey was real, and not caused by gear effects an artefact (Bagley et al. 2009). However, it is not possible, from plots such as this, to say how unusual this was, or whether this amount of variation is more than is allowed for in the assessment (an assumed process-error c.v. of 0.2 is to allow for year-to-year variations in trawl-survey catchability). An analysis of residuals from the 2008 stock-assessment model may be more informative.

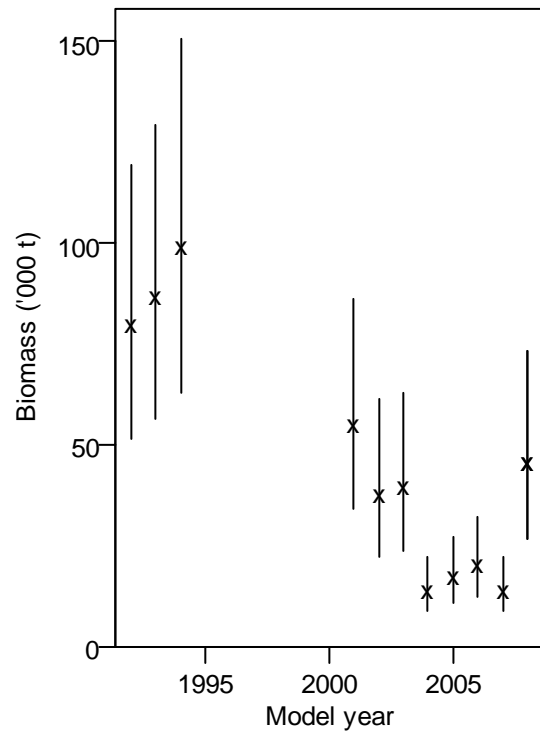


Figure 25: Biomass estimates ('x', with 95% confidence intervals as vertical lines) from trawl-survey time series SAsumbio. The plotted confidence intervals allow for both observation and process error. Plotted years are as in the model (so the last survey, in late 2007, is plotted in model year 2008).

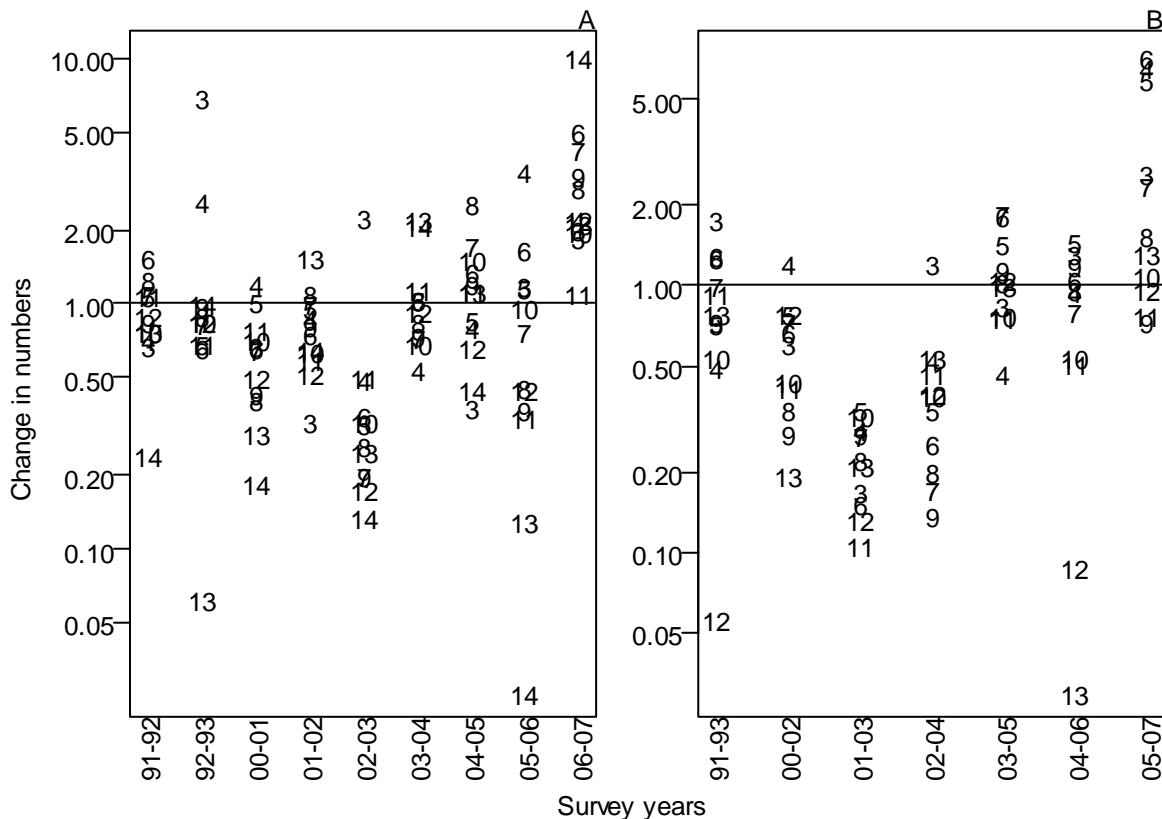


Figure 26: Changes, between pairs of surveys in the sub-Antarctic summer series, in estimated numbers of selected cohorts: A, surveys in consecutive years and B, surveys two years apart. Each plotted point indicates how the estimated number in a cohort changed between two surveys; the plotting symbol is the age of the cohort in the earlier survey. For example, the top right point in panel A shows that the estimated number in the cohort that was aged 14 in the 2006 survey increased by a factor of 10 between the 2006 and 2007 surveys.



## 2.7 The problem of defining $B_0$ in run 4.7

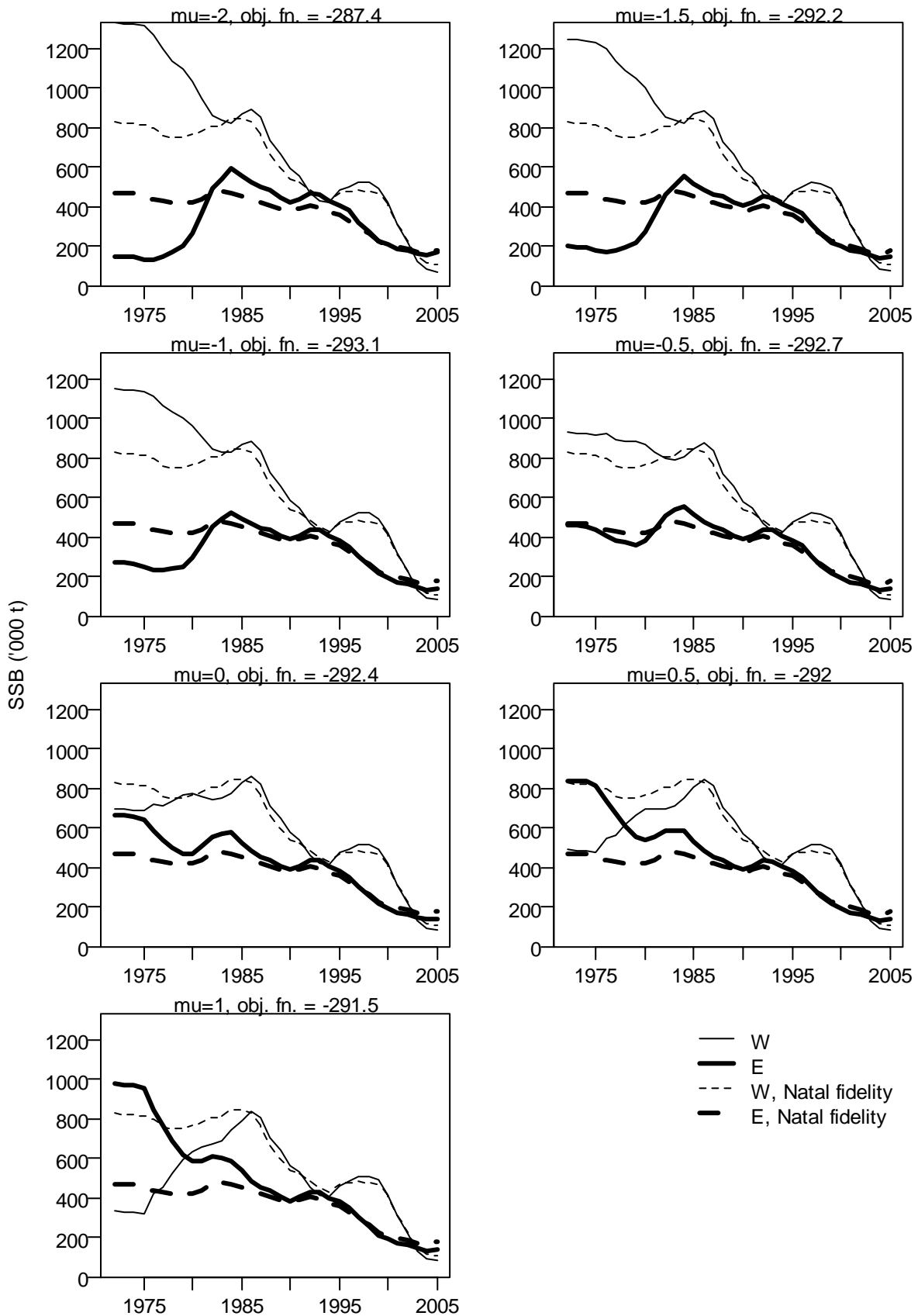
The above long-term projections for hoki (see Section 2.1) drew attention to the problem of defining  $B_0$  for the E and W stocks in model runs like 4.7, in which natal fidelity is not assumed. The existence of this problem was well known. However, the recent results shed new light on the problem and emphasised the deficiencies in current assumptions for this type of model. I will briefly recapitulate the problem as it stood in 2007, then describe what we can learn from the recent work, and finally suggest a way ahead for these types of runs.

There are several important differences between models with and without natal fidelity (Table 16). A result of these differences is that  $B_0$  is clearly defined for both stocks (E and W) in the former models, but only for the total stock ( $T = E + W$ ) in the latter. Where it is defined,  $B_0$  is always interpretable as the mean unfished spawning biomass (i.e., the biomass that would occur if there was no fishing and recruitment was average in every year). One of the most interesting results from initial explorations of models with natal fidelity was that the observations contain almost no information about the relative sizes of the initial biomasses in E and W. This is shown in Figure 27 (note that as  $mu$  increases, the initial biomasses change substantially, but the objective function value changes little, as do the biomass trajectories over the years after 1988, when the first observations occur). It was decided to constrain the initial biomasses by applying the same prior to  $B_{initE}$  and  $B_{initW}$  as was applied, in runs with natal fidelity, to  $B_{0E}$  and  $B_{0W}$  (see pE and pE' in Table 16) and then to assume  $B_0 = B_{init}$  for the E and W stocks. Although this latter assumption is not ideal, it seemed better to include model runs of this sort in the assessment (and thus include uncertainty about natal fidelity) than to exclude them.

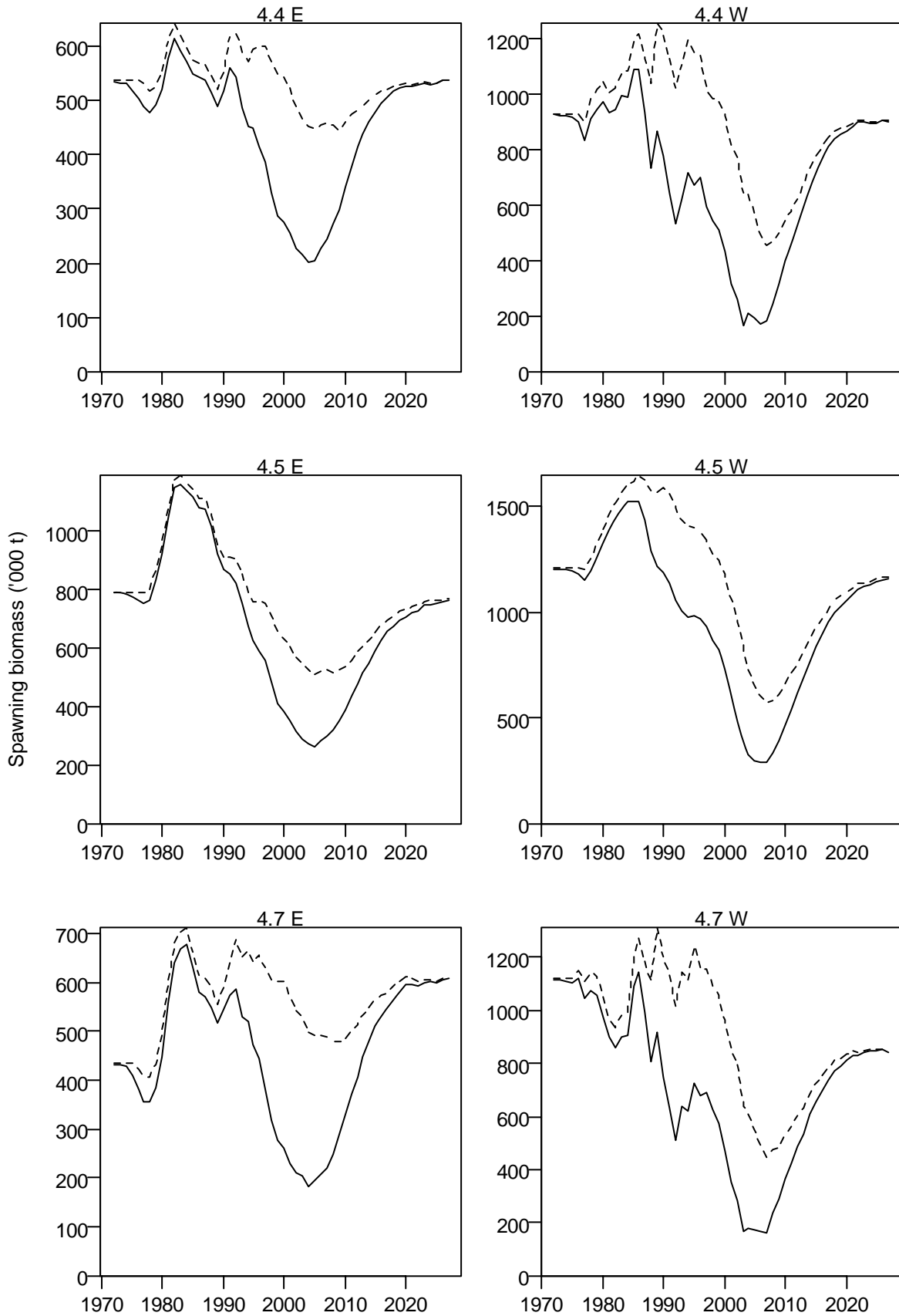
**Table 16: Some important differences between hoki models with and without natal fidelity.  $B$  = spawning biomass; YCS, year-class strength;  $R$ , recruitment;  $m$ , a multiplier which allows year-to-year variation in the strength of the annual migration of juveniles from Chatham Rise to sub-Antarctic;  $y$  indexes years; E, W, and T refer to the east, west, and total stocks. [In the interests of clarity, some aspects of the models have been simplified for presentation in this table. Full model descriptions were given by Francis (2008b).]**

	Models with natal fidelity	Models without natal fidelity
Key parameters	$B_{0E}, B_{0W}, YCS_{y,W}, YCS_{y,E}$	$B_{0T}, YCS_{y,T}, m_y$
Derived parameters	$R_{yE}, R_{yW}$	$R_{yT}$
Constraints	Prior on pE [= $B_{0E}/(B_{0E}+B_{0W})$ ] $B_{initE} = B_{0E}, B_{initW} = B_{0W}$	Prior on pE' [= $B_{initE}/(B_{initE}+B_{initW})$ ] Restrict variation, but not mean, of $m_y$

What the recent work on long-term projections has provided for the first time is a natural (though cumbersome) way of defining  $B_{0E}$  and  $B_{0W}$  for runs without natal fidelity: the median spawning biomass after a long (20-year) projection period without fishing, assuming that future conditions are like those in the past (i.e., in the projections, future values of  $R_{yT}$  and  $m_y$  were sampled from past values). To check that 20 years is long enough to reach equilibrium I calculated, for each stock and run, what the biomass trajectories would have been had these stocks never been fished. The calculated trajectories show that 20 years is indeed long enough, and that, with its new definition,  $B_0$  can be quite different from  $B_{initial}$  (Figure 28). Further, the new definition for runs without natal fidelity is consistent with the usual definition in runs with natal fidelity (note that initial and final biomasses are very similar for runs 4.4 and 4.5).



**Figure 27: Results from early explorations of models without natal fidelity. In each panel, the solid lines are biomass trajectories for the W (light lines) and E (heavy lines) populations in a series of runs that differed only in the assumed mean,  $\mu$ , of the migration multipliers,  $m_j$ . The value of the objective function for each run is shown above the plot. The broken lines, the same in each panel, are the biomass trajectories for a run with natal fidelity. (This plot was figure 57 in Francis 2006).**



**Figure 28: Median SSB trajectories from 1972 to 2027, by stock and run, assuming no catches (dashed line) or historic catches to 2007 and then no catches (solid line).**

Another point of interest is that the new definition of  $B_0$  makes the estimates of current stock status (i.e.,  $B_{\text{current}}$  as  $\%B_0$ ) from run 4.7 more like those from the other two runs (Table 17).

**Table 17: Estimates of median biomass for the three final runs of the 2007 assessment, and also for a modified version of run 4.7 with a new definition of  $B_0$  (see text for details).**

Run	$B_0$ ('000 t)		$B_{\text{current}}$ ( $\%B_0$ )	
	E	W	E	W
4.4	535	930	46	20
4.5	792	1207	37	24
4.7	432	1115	51	15
Modified 4.7	610	841	36	19

### 2.7.1 Proposed way ahead

The new definition of  $B_0$  for runs without natal fidelity seems promising, but it has not been thoroughly explored and it may have some unforeseen consequences, or pose implementation difficulties. However, we have now clearly seen the weakness of the old definition. Accordingly, I propose excluding runs without natal fidelity from the 2008 stock assessment, and using the time before the 2009 assessment to develop, and fully test, a better set of assumptions for such runs.

## 3. MODEL ASSUMPTIONS AND INPUTS FOR 2008

This section provides a fairly detailed summary of all model assumptions and inputs for the 2008 assessment. A complete description is contained, for the final runs only, in the files given in Appendix 1 (which should be read in conjunction with the CASAL manual (Bull et al. 2008)).

The model uses Bayesian estimation. In describing the model assumptions I will sometimes need to distinguish between different types of model runs: MPD versus MCMC, or initial versus final. MPD runs are so called because they estimate the Mode of the Posterior Distribution, which means they provide a point estimate, whereas MCMC (or full Bayesian) runs provide a sample from the posterior distribution using a Markov Chain Monte Carlo technique (this sample is sometimes referred to as a chain). MCMC runs are more informative, but much more time consuming to produce. For this reason only MPD runs were used for most of the initial exploratory analyses (Section 4). These runs were used to define the assumptions for the final model runs (Section 5), which were full Bayesian, and whose results provide the formal stock assessment.

The model is based on the fishing year, which is labelled by its second part, so 1990 refers to the 1989–90 fishing year. This convention is applied throughout, so that, for instance, the most recent sub-Antarctic survey, carried out in November–December 2006, is referred to as the 2007 survey.

### 3.1 Model structure and catches

Two stocks are assessed. Fish from the eastern (E) stock spawn in Cook Strait (CS) and have their home grounds in Chatham Rise (CR); the western (W) stock spawn in west coast South Island (WC) and have their home grounds in sub-Antarctic (SA) (these areas are shown in Figure 1). Soon after being spawned, all juveniles move to CR. In previous assessments two alternative assumptions concerning the juveniles have been modelled. The original assumption is that the juveniles show natal fidelity – that is, they grow up to spawn on the ground where they were spawned. Under this assumption, the stock to which a fish belongs is determined at birth. At some time before age 8 all W juveniles migrate to their home ground, SA. The alternative assumption, used first in 2006, is that there is no natal fidelity. Each year, some juveniles choose to move from the nursery ground (CR) to SA, and thereby become W fish. Those fish remaining in CR by age 8 become E fish. It was decided,

pending the resolution of some technical problems (see Section 2.7), to avoid models without natal fidelity in the 2008 assessment.

The model partition divides the population into two sexes, 17 age groups (1 to 17+), four areas corresponding to the four fisheries (CR, WC, SA, and CS), and two stocks (E and W). The annual cycle (Table 18) is the same as in 2007.

**Table 18: Annual cycle of the assessment model, showing the processes taking place at each time step, their sequence within each time step, and the available observations (excluding catch at age). This is unchanged from that used in 2005. *M* fraction is the proportion of natural mortality which occurs within the time step. An age fraction of, say, 0.25 for a time step means that a 2+ fish is treated as being of age 2.25 in that time step. The last column (“propn. mort.”) shows the proportion of that time step’s mortality that is assumed to have taken place when each observation is made.**

Step	Approx. months	Processes	<i>M</i> fraction	Age fraction	Observations label	Observations propn. mort.
1	Oct-Nov	migrations Wrtn: WC→SA, Ertn: CS→CR	0.17	0.25	–	
2	Dec-Mar	recruitment at age 1+ to CR (for both stocks) part1, non-spawning fisheries (Ensp1, Wnsp1)	0.33	0.60	SAsum CRsum	0.5 0.6
3	Apr-Jun	migration Whome: CR→SA part2, non-spawning fisheries (Ensp2, Wnsp2)	0.25	0.90	SAaut pspawn	0.1
4	End Jun	migrations Wspmg: SA→WC, Espmg: CR→CS	0.00	0.90	–	
5	Jul-Sep	increment ages spawning fisheries (Esp, Wsp)	0.25	0.0	CSacous WCacous	0.5 0.5

As in 2007, the catches used in the model (Table 19) were calculated by apportioning the official total catch for each year amongst the six fisheries using the method described in Table 20. The catches from 2001 to 2006 were slightly revised using the most recent data from MFish and the assumed catches for the current year (2008) were scaled from those for 2007 to agree with the new TACC (90 000 t) and catch split (65 000 t E, 25 000 t W). The proportion of the catch taken from the western fisheries increased between 1996 and 2002, but has since dropped as fishers shifted effort from West Coast South Island (Wsp) to Cook Strait (Esp) to reduce pressure on the W stock (Figure 29).

Catches used in the assessment are shown in Table 1 and Figure 1.

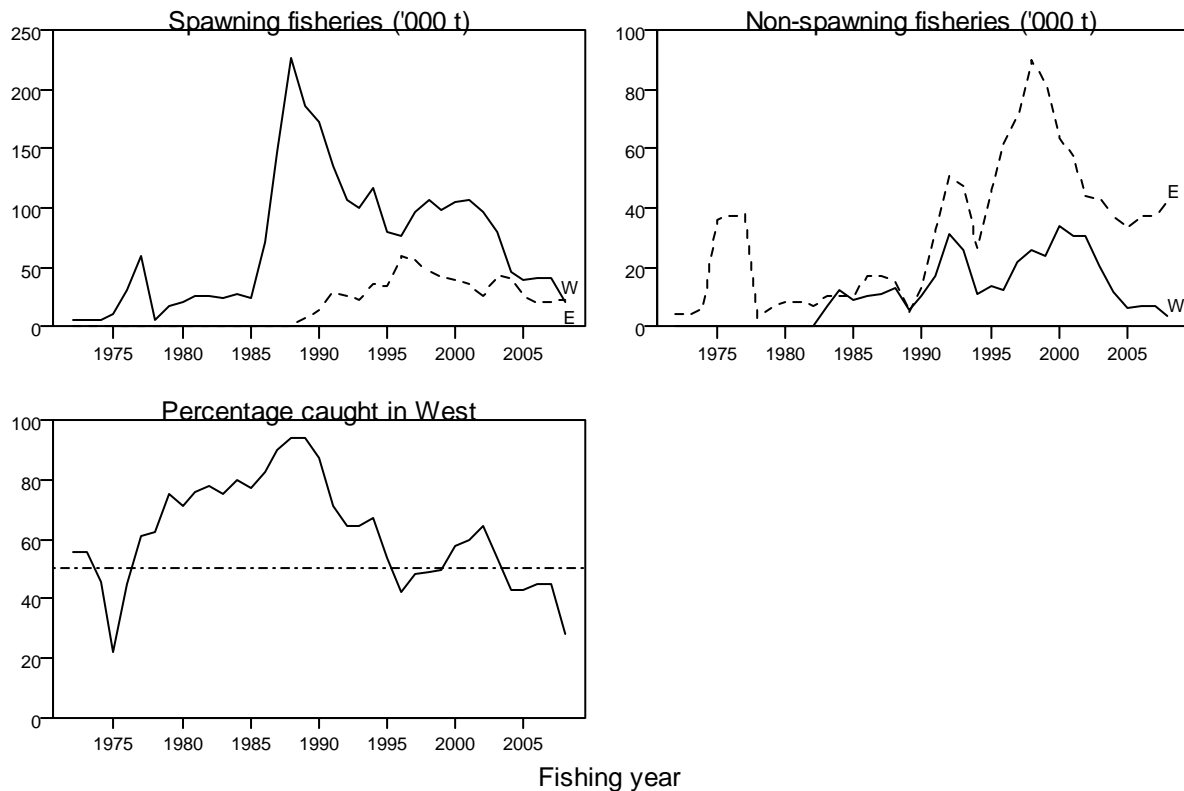
**Table 19: Catches (t) by fishery and fishing year (1972 means fishing year 1971–72), as used in the assessment.**

Year	Fishery					
	Ensp1	Ensp2	Wnsp1	Wnsp2	Esp	Wsp
1972	1500	2500	0	0	0	5000
1973	1500	2500	0	0	0	5000
1974	2200	3800	0	0	0	5000
1975	13100	22900	0	0	0	10000
1976	13500	23500	0	0	0	30000
1977	13900	24100	0	0	0	60000
1978	1100	1900	0	0	0	5000
1979	2200	3800	0	0	0	18000
1980	2900	5100	0	0	0	20000
1981	2900	5100	0	0	0	25000
1982	2600	4400	0	0	0	25000
1983	1500	8500	3200	3500	0	23300
1984	3200	6800	6700	5400	0	27900
1985	6200	3800	3000	6100	0	24900
1986	3700	13300	7200	3300	0	71500
1987	8800	8200	5900	5400	0	146700
1988	9000	6000	5400	7600	600	227000
1989	2300	2700	700	4900	7000	185900
1990	3300	9700	900	9100	14000	173000
1991	17400	14900	4400	12700	29700	135900
1992	33400	17500	14000	17400	25600	107200
1993	27400	19700	14700	10900	22200	100100
1994	16000	10600	5800	5500	35900	117200
1995	29600	16500	5900	7500	34400	80100
1996	37900	23900	5700	6800	59700	75900
1997	42400	28200	6900	15100	56500	96900
1998	55600	34200	10900	14600	46700	107100
1999	59200	23600	8800	14900	40500	97500
2000	43100	20500	14300	19500	39000	105600
2001	37100	20100	13300	17300	35400	106700
2002	25100	18800	16900	13500	25000	96300
2003	24400	19000	12400	7900	41900	79000
2004	17900	19100	6400	5400	41000	46200
2005	19300	13900	4400	2000	26400	38500
2006	22200	14800	2000	4800	20500	40200
2007	22500	15100	2000	4800	20700	40700
2008	25100	16800	1100	2500	23100	21400

**Table 20: Method of dividing annual catches into the six model fisheries (Esp, Wsp, Ensp1, Ensp2, Wnsp1, and Wnsp1). The small amount of catch reported in the areas west coast North Island and Challenger (typically 100 t per year) was ignored (which means that this catch is pro-rated across all fisheries).**

Area	Oct–Mar	Apr–May	Jun–Sep
West coast South Island; Puysegur	Wsp	Wsp	Wsp
Sub-Antarctic	Wnsp1	Wnsp2	Wnsp2
Cook Strait; Pegasus	Ensp1	Ensp2	Esp
Chatham Rise; east coasts of South Island & North Island; null <sup>1</sup>	Ensp1	Ensp2	Ensp2

<sup>1</sup> no area stated



**Figure 29:** Annual catches by fishery for the spawning (top left panel) and non-spawning (top right panel) fisheries, and annual percentage of catch caught in western fisheries (Wsp, Wnsp1, Wnsp2) (bottom panel).

The fixed biological parameters in the model (Table 21) are unchanged from those used in 2007.

**Table 21:** Fixed biological parameters used by the model. Sources: a, Horn & Sullivan (1996) by sex, and Francis (2005) for both sexes combined; b, Francis (2003); c, assumed.

Type	Symbol	All fish	W stock			E stock			Source
			Male	Female	Both	Male	Female	Both	
Growth	$L_{\infty}$		92.6	104.0	102.1	89.5	101.8	100.8	a
	$k$		0.261	0.213	0.206	0.232	0.161	0.164	
	$t_0$		-0.5	-0.6	-0.96	-1.23	-2.18	-2.16	
Length-weight [W(kg)=aL(cm) <sup>b</sup> ]	$a$	$4.79 \times 10^{-6}$							b
	$b$	2.89							
Proportion by sex at birth		0.5							c

### 3.2 Ogives

The ogives used in the model are the same as in 2007: six selectivity ogives (one for each of the four fisheries — Espsl, Wpspl, Enspsl, Wnspsl — and one each for trawl surveys in areas CR and SA – CRsl, SAsl), and three migration ogives (for migrations Whome, Espmg, and Wspmg). As in previous years, two alternative sets of ogive assumptions were used for the final runs (Table 22).

The home migration ogive, Whome, applied only to the W juveniles in CR and was the same in every year. At age 8, all W juveniles remaining in CR were forced to migrate to SA. In previous years this ogive has had a different interpretation in models without natal fidelity (Francis 2008b).

**Table 22: Ogive assumptions for the two final runs. In the ogive constraints,  $O_{7,F,E}$  refers to the ogive value at age 7 for female fish from the E stock, etc.**

Runs	Ogive type	Description	Constraints
2.3	Spawning selectivity	Length-based, logistic	same for M and F, same for E and W
	Non-spawning selectivity	Length-based, double-normal	same for M and F, must be domed <sup>1</sup>
	Survey selectivity	Length-based, double-normal	same for M and F, must be domed <sup>1</sup>
	Spawning migration	Free, ages 1–8	$O_{8,M,E} = O_{8,M,W}$ , $O_{8,F,E} = O_{8,F,W} \geq 0.6$ $O_A = O_8$ for $A > 8$
2.3	Home migration	Free, ages 1–7	same for M and F, =1 for age > 7
2.4	Spawning selectivity	Age-based, double-normal	same for E and W
	Non-spawning selectivity	Age-based, double-normal	
	Survey selectivity	Age-based, double-normal	
	Home migration	Free, ages 1–7	=1 for age > 7
	Spawning migration	Free, ages 1–8	$O_A = O_8$ for $A > 8$

<sup>1</sup> see figure 11, and associated text, of Francis et al. (2003) for an explanation of what this means

As in previous years, the model attempted to estimate annual changes in  $W_{spsl}$  (the selectivity ogive for W spawning fishery). Following the recommendation of Francis (2006) these changes were restricted to years for which there were  $W_{spage}$  data (i.e., from 1988 onwards). The changes were driven by the median day of the fishery (Table 23). Annual changes in the selectivity for the other fisheries were not estimated because these were shown not to improve model fits in 2003 (Francis 2004).

**Table 23: Median catch day by year for  $W_{sp}$ , as used in estimating annual changes in the selectivity  $W_{spsl}$ . The mean value was used for all years (including 2007) for which there was catch but no  $W_{spage}$  data.**

1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
299	302	298	301	306	304	308	307	312	310	311	309
2000	2001	2002	2003	2004	2005	2006	2007	Mean			
309	309	308	309	307	309	310	307	305			

### 3.3 Other structural assumptions

For each stock, the population at the start of the fishery was assumed to have a stable age structure with biomass,  $B_0$ , and constant recruitment,  $R_0$ . The Francis parameterisation of recruitment was used. Thus, recruitment at age 1 in year  $y$  in each stock was given by  $R_y = R_{mean} \times YCS_{y-2} \times SR(SSB_{y-2})$ , where  $YCS_y$  is the year-class strength for fish spawned in year  $y$ ,  $SR$  is a Beverton-Holt stock-recruit relationship with assumed steepness 0.75 (reduced from 0.9 – see Section 2.5),  $R_{mean}$  is the expected recruitment (ignoring the stock-recruit relationship), and  $SSB_y$  is the mid-season spawning stock biomass in year  $y$ .  $R_0$  is calculated as  $R_{mean} Y_{mean}$ , where  $Y_{mean}$  is the mean  $YCS$  over the years 1975 to 2002, inclusive (so  $R_0$  is mean recruitment over those years, ignoring the effect of the stock-recruit relationship).

Thirty-two  $YCS$ s were estimated for each stock, for years 1975 to 2006, inclusive.  $YCS$ s for the initial years (1970 to 1974) were fixed at 1. The E and W  $YCS$ s for 2006 were constrained (by a penalty function) to be equal for MPD runs, but this constraint was removed for the full Bayesian runs.

The maximum exploitation rates assumed were the same as in previous years: 0.3 in each part of the two non-spawning fisheries (which is approximately equivalent to 0.5 for the two parts combined), and 0.67 for both spawning fisheries. A penalty function was used to strongly discourage model estimates for which these maximum exploitation rates were exceeded.



As in previous years, the model's expected age distributions had ageing error applied to them before they were compared with the observed distributions (i.e., before they were used to calculate the objective function value).

### 3.4 Observations

Three types of observations were used in the model: biomass indices (Table 24), proportions at age (and sex) (Table 25, Figure 30), and proportion spawning (Table 26). Biomass indices new to this assessment came from an acoustic survey in Cook Strait in August 2007 (O'Driscoll, unpublished results), and trawl surveys of the Sub-Antarctic in December 2007 (O'Driscoll & Bagley, unpublished results) and Chatham Rise in January 2008 (Stevens & O'Driscoll, unpublished results).

The proportions-at-age data fall into three groups. The first group — trawl survey (CRsumage, SAsumage, SAautage) and spawning catch at age (Wspage, Espage) — is the most substantial and reliable. These data are otolith-based, and use an age-length key to transform proportions at length to proportions at age. The non-spawning otolith-based data (Enspage, Wnspage) are available only since sufficient otoliths have been collected from these fisheries. Because the fisheries are spread over many months, these proportions at age must be estimated directly (rather than via an age-length key). The third group of data (EnspOLF, WnspOLF), which is OLF-based (Hicks et al. 2002), is less reliable because of the difficulty of inferring age distributions from length data alone.

Although both trawl surveys provide information about year-class strengths (YCSs) the CR survey is more reliable for recent year classes. This is apparent in two ways. First, the between-survey consistency of estimated numbers at age for the youngest ages is much greater in CR than in SA (Figure 31). Second, the correlation between these estimates and model estimates of YCS is not strong until age 4 for the SA survey, but is quite strong at age 1 for the CR survey (Figure 32).

The 1998 proportions-spawning data have been changed to the 'best3' estimates (see Section 2.4) but the 1993 proportions are unchanged (Table 26).

The way the proportions-at-age data enter the model varies amongst data sets (Table 27). As in 2002 (and all subsequent years), all proportions less than 0.0001 were replaced by 0.0001 (for reasons, see Francis et al. (2003)). For the otolith-based data sets the maximum ages were set as high as was possible without allowing the percentage of data points requiring this adjustment to exceed 2%.

**Table 24: Biomass indices ('000 t) used in the assessment, with observation and total c.v.s (respectively) in parentheses. Bold values are new to this assessment.**

	<u>CRsumbio</u>	<u>SASumbio</u>	<u>SAautbio</u>	<u>CSacous</u>	<u>WCacous</u>
1988	–	–	–	–	417 (0.22,0.60)
1989	–	–	–	–	249 (0.15,0.38)
1990	–	–	–	–	255 (0.06,0.40)
1991	–	–	–	180 (0.13,0.41)	340 (0.14,0.73)
1992	120 (0.08,0.21)	80 (0.07,0.21)	68 (0.08,0.22)	–	345 (0.14,0.49)
1993	186 (0.10,0.22)	87 (0.06,0.21)	–	583 (0.15,0.52)	550 (0.07,0.38)
1994	146 (0.10,0.22)	100 (0.09,0.22)	–	592 (0.06,0.91)	–
1995	120 (0.08,0.21)	–	–	427 (0.12,0.61)	–
1996	153 (0.10,0.22)	–	89 (0.09,0.22)	202 (0.09,0.57)	–
1997	158 (0.08,0.22)	–	–	295 (0.12,0.40)	654 (0.10,0.60)
1998	87 (0.11,0.23)	–	68 (0.11,0.23)	170 (0.10,0.44)	–
1999	109 (0.12,0.23)	–	–	243 (0.10,0.36)	–
2000	72 (0.12,0.23)	–	–	–	396 (0.14,0.60)
2001	60 (0.10,0.22)	56 (0.13,0.24)	–	220 (0.12,0.30)	–
2002	74 (0.11,0.23)	38 (0.16,0.26)	–	320 (0.13,0.35)	–
2003	53 (0.09,0.22)	40 (0.14,0.24)	–	225 (0.17,0.34)	–
2004	53 (0.13,0.24)	14 (0.13,0.24)	–	–	–
2005	85 (0.12,0.23)	18 (0.12,0.23)	–	132 (0.11,0.32)	–
2006	99 (0.11,0.23)	21 (0.13,0.24)	–	126 (0.17,0.34)	–
2007	70 (0.08,0.22)	14 (0.11,0.23)	–	<b>216</b> (–,0.46)	–
2008	<b>77</b> (0.11,0.23)	<b>46</b> (0.16,0.26)	–	–	–

**Table 25: Description of the proportions-at-age observations used in the assessment. These data derive either from otoliths or from the length-frequency analysis program OLF (Hicks et al. 2002). Data new to this assessment are in bold type.**

Area	Label	Data type	Years	Source of age data
WC	Wspage	Catch at age	1988– <b>07</b>	otoliths
SA	WnspOLF	Catch at age	1992–94, 96, 99–00	OLF
	Wnspage	Catch at age	2001–04, 06– <b>07</b>	otoliths
	SAsumage	Trawl survey	1992–94, 2001– <b>08</b>	otoliths
	SAautage	Trawl survey	1992, 96, 98	otoliths
CS	Espage	Catch at age	1988– <b>07</b>	otoliths
CR	EnspOLF	Catch at age	1992, 94, 96, 98	OLF
	Enspage	Catch at age	1999– <b>07</b>	otoliths
	CRsumage	Trawl survey	1992– <b>08</b>	otoliths

**Table 26: Proportion spawning data, pspawn. These are estimates, from the 1992 and 1998 SAaut surveys, of the proportion, by age, of females that were expected to spawn in the following winter.**

Year	<u>Age</u>						
	3	4	5	6	7	8	9+
1992	0.19	0.45	0.27	0.57	0.68	0.69	0.64
1998	0.27	0.47	0.39	0.41	0.48	0.43	0.53

**Table 27: Age ranges used for at-age data sets. In all cases the last age was treated as a plus group.**

Data set	<u>Age range</u>	
	Lower	Upper
Espage, Wspage, SAsumage, SAautage	2	15
Wnspage	2	13
CRsumage, Enspage	1	13
WnspOLF	2	6
EnspOLF	1	6
pspawn	3	9

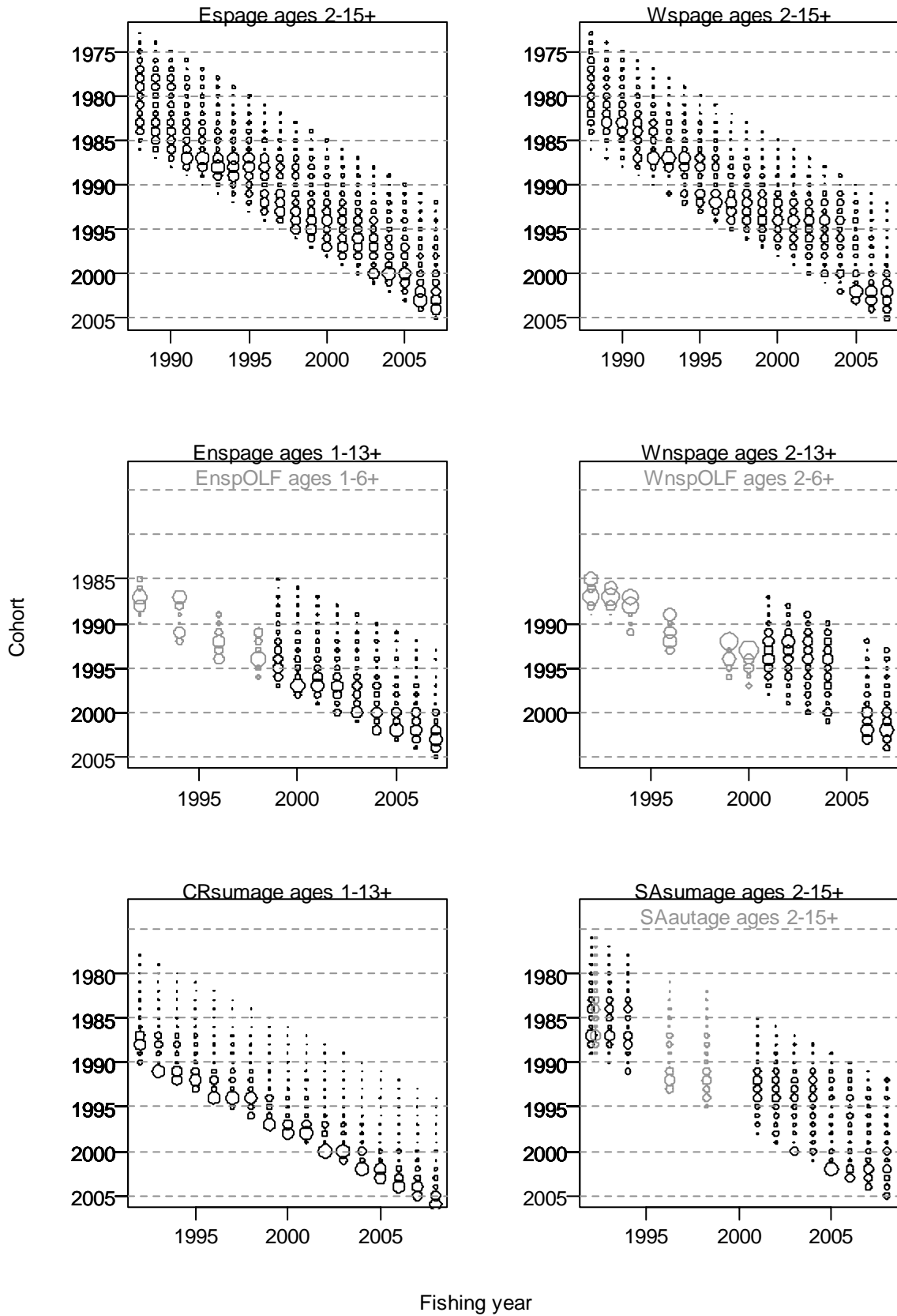
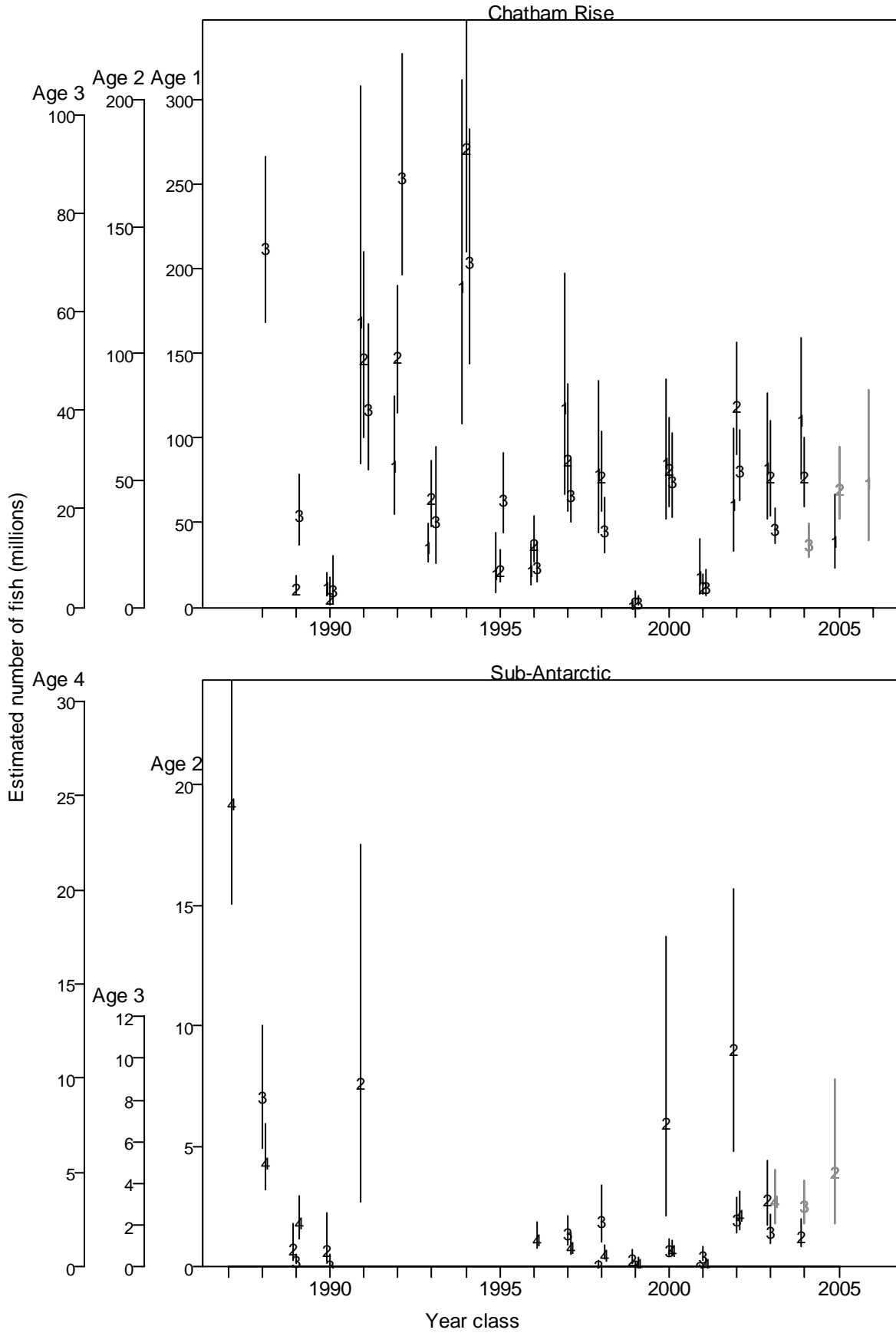
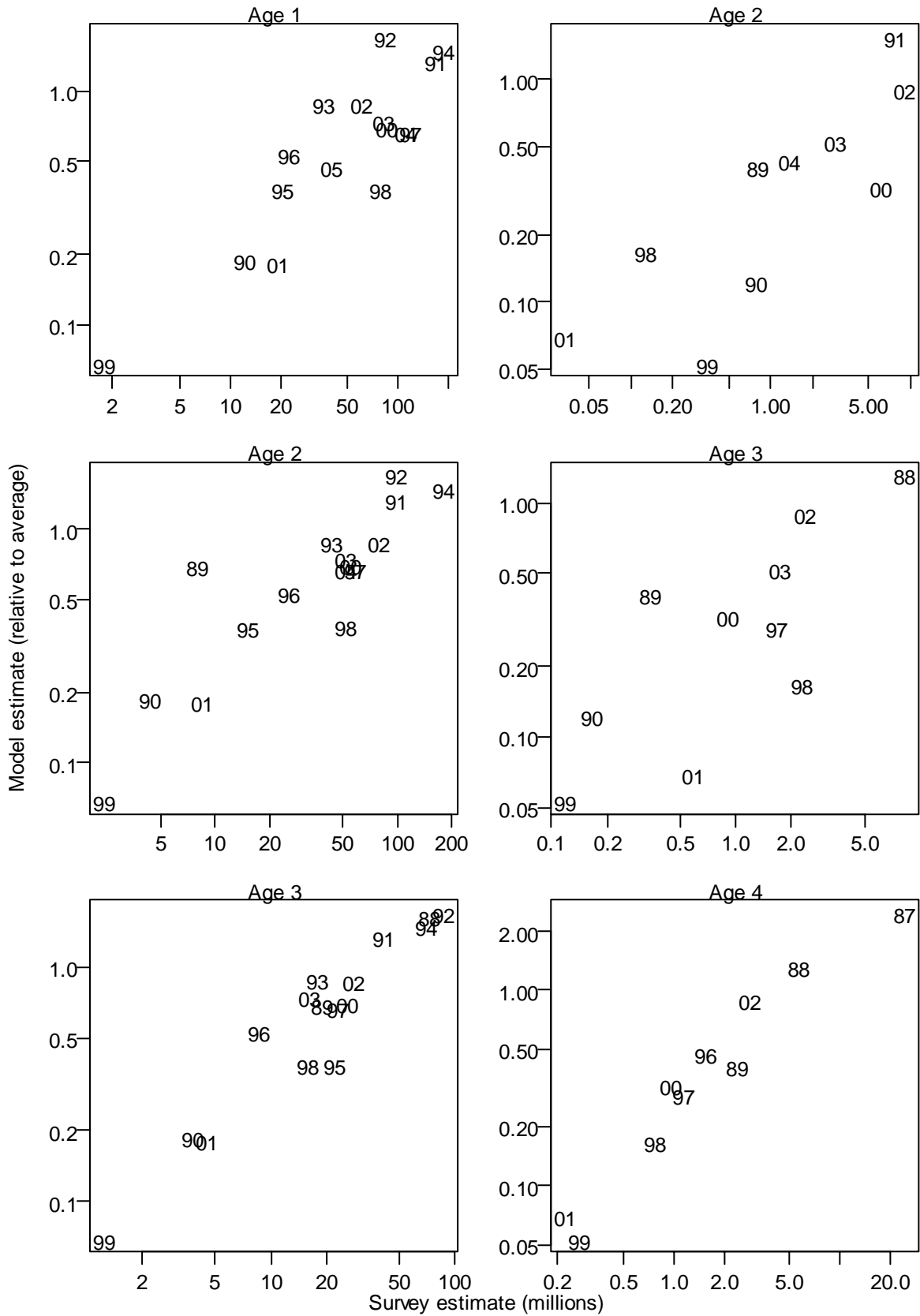


Figure 30: Proportions-at-age data, plotted by cohort and fishing year, with both sexes combined. The area of each circle is proportional to the associated proportion at age. Circle positions for the SAautage data have been offset horizontally to allow them to be plotted on the same panel as the SASumage data. In all panels the right-most column of circles is new to this assessment.



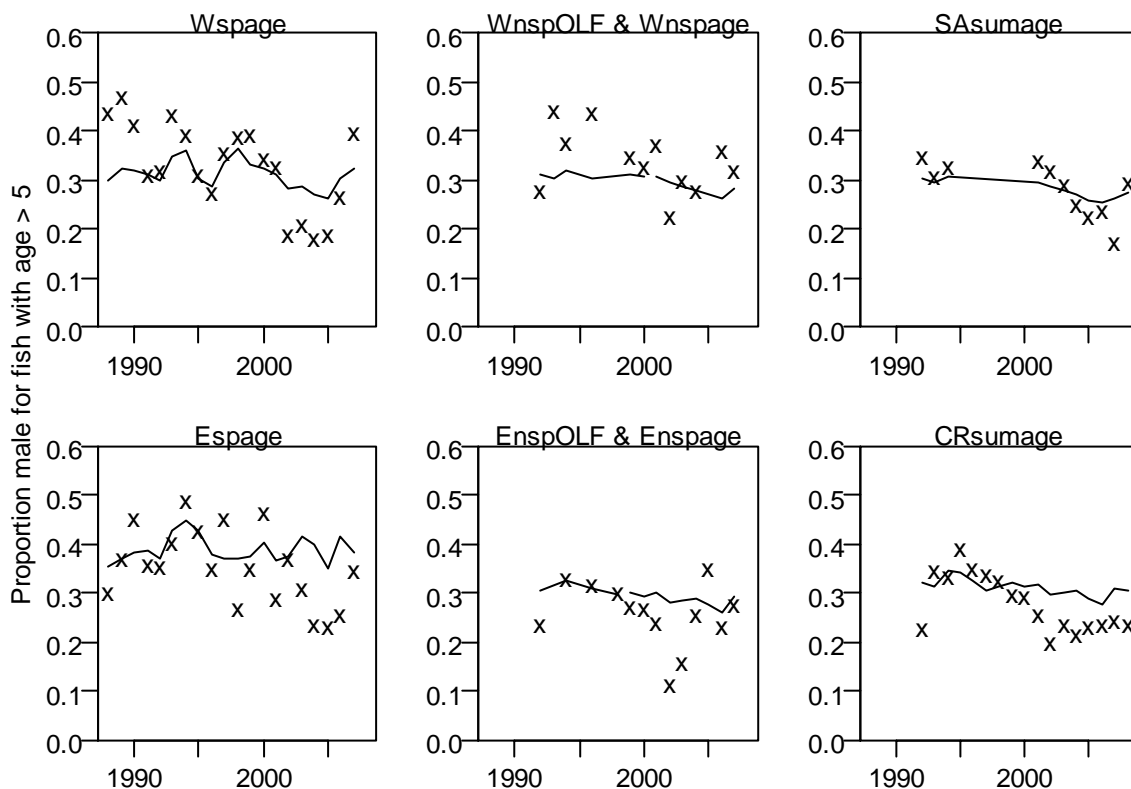
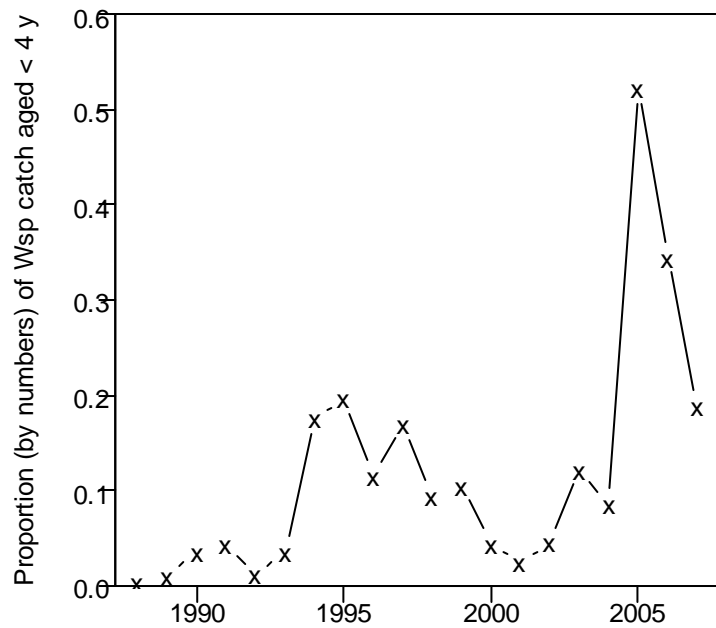
**Figure 31: Estimated numbers at age for each of the three youngest ages in the two major trawl survey series. Vertical bars indicate 95% confidence intervals; data points new to this assessment are plotted in grey.**



**Figure 32: Comparison of survey and model estimates of YCSs (each point represents one year class and the plotting symbol identifies the year). Left panels compare CRsumage estimates at ages 1, 2, and 3 with model estimates of total (E + W) YCS; right panel compare SAsumage estimates at ages 2, 3, and 4 with model estimates of W YCS. The model estimates are from run 4.4 in the last assessment; only survey data available for that assessment are included in the plot.**

The observations show two reassuring trends. The proportion of young fish in the catch from the west spawning fishery has continued to fall from its very high value in 2005 (Figure 33). This is evidence of improved recruitment to the western stock. Also, the trend, amongst older fish in the Wspace and SASumage data sets, towards an increasing dominance of females has reversed (Figure 34).

**Figure 33: Annual proportion of young fish (aged less than 4 y) in the catch from the west spawning fishery. Note that all plotted proportions are based on numbers of fish; proportions based on weight would be much smaller.**



**Figure 34: Observed ('x') and predicted (lines) proportions male, by year and data set, for older fish (ages > 5 y). Predicted values are from run 2.3.**

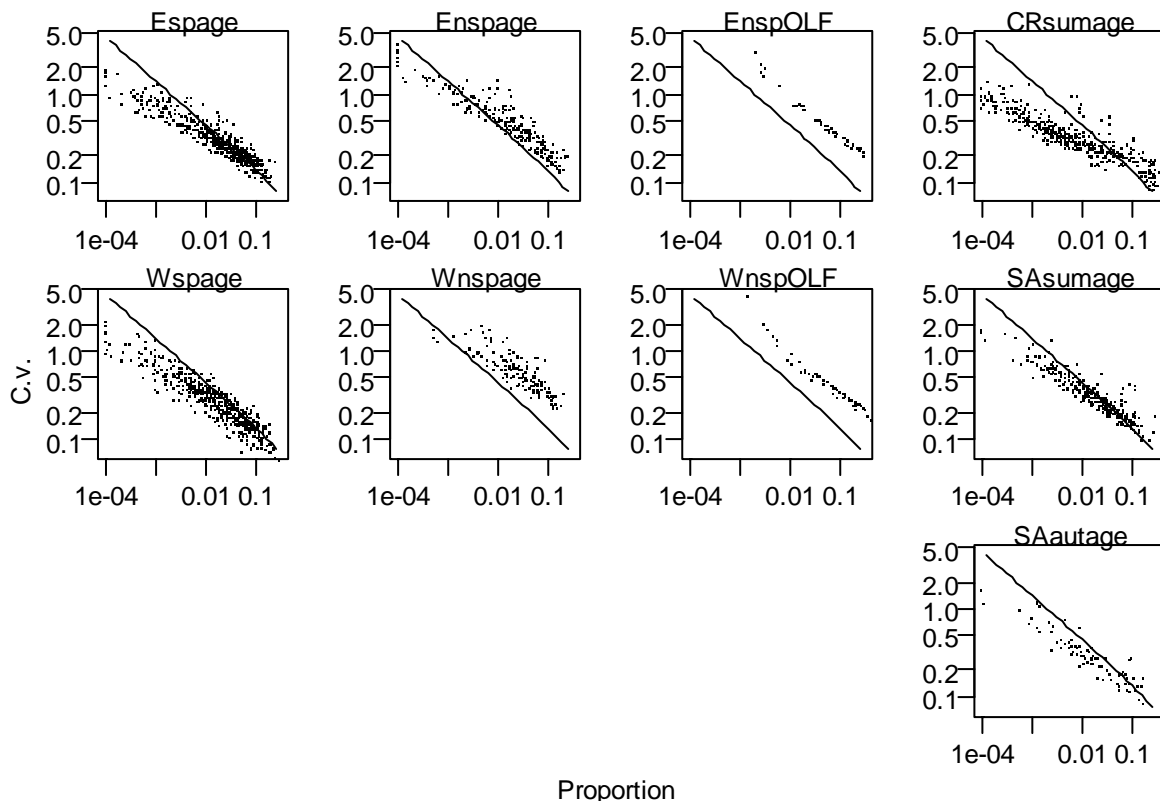
### 3.5 Error assumptions

The error distributions assumed were robust lognormal (Bull et al. 2008) for the proportions-at-age data, and lognormal for all other data. This means that the weight assigned to each datum was controlled by an error c.v. In this section I describe how these c.v.s were assigned.

For the biomass indices, two alternative sets of c.v.s were available (see Table 24). The total c.v.s represent the best estimates of the uncertainty associated with these data, and were used in all initial model runs. For the acoustic indices, these were calculated using a simulation procedure intended to include all sources of uncertainty (O'Driscoll 2002), and the observation-error c.v.s were calculated in a similar way but including only the uncertainty associated with between-transect (and within-stratum) variation in total backscatter. For the trawl indices, the total c.v.s were calculated as the sum of an observation-error c.v. (using the standard formulae for stratified random surveys, e.g., Livingston & Stevens (2002)) and a process-error c.v., which was set at 0.2, (following Francis et al. 2001) (note that c.v.s add as squares:  $c.v._{total}^2 = c.v._{process}^2 + c.v._{observation}^2$ ). In some model runs (see below) it was decided to upweight some biomass indices by using their observation, rather than total, c.v.s.

For almost all of the proportions-at-age observations, total c.v.s were treated as the sum of a process-error c.v. and an observation-error c.v. (the only exception was pspawn, for which an arbitrary c.v. of 0.25 was assumed, following Cordue (2001)). Observation-error c.v.s for the remaining otolith-based data were calculated by bootstrapping. For the OLF-based data the c.v.s used were the same as in 2004 (Francis 2005). As is typical with proportions, estimated c.v.s decreased as proportions increase (Figure 35).

Process-error c.v.s for the at-age data were estimated within the model (one c.v. for each data set) for all point estimates, as in previous years. For full Bayesian estimates, these c.v.s were fixed. Although there is some evidence that these process-error c.v.s should decrease with increasing age, there does not appear to be a strong need to implement such a relationship (Francis 2004).



**Figure 35: Observation-error c.v.s for the proportions-at-age data sets. Each point represents a proportion at a specific age and sex for a given year. The diagonal line, which is the same in each panel, is added to aid comparison between panels; it shows the relationship between proportion and c.v. that would hold with simple multinomial sampling with sample size 500.**

### 3.6 Parameters, priors, and penalties

The number of parameters estimated in the final model runs was 139 (for runs 4.4 and 4.7) or 117 (for run 4.5) (Table 28). Most of the associated prior distributions were intended to be uninformative; the main exceptions were those for the catchabilities (O'Driscoll et al. 2002), pE, and natural mortality (Smith 2004). For selectivity[Wspsl].shift\_a and migration[Whome].annual\_variation\_values, normal priors were used with standard deviations more or less arbitrarily chosen to discourage extreme values (see sections 7.1 and 7.3, respectively, of Francis (2006)).

As in previous assessments, the model estimated natural mortality separately by sex (when sex was included in the model) because of the trends with age in the sex ratio.

**Table 28: Parameters estimated in the final model runs, and their associated prior distributions. Where the number of parameters varied between model runs the two values given are for runs 2.3 and 2.4, respectively. Distribution parameters are: bounds for uniform and uniform-log; mean (in natural space) and c.v. for lognormal; and mean and s.d. for normal and beta.**

Parameter(s)	Description	Type	Distribution		No. of parameters
			Parameters		
log_Bmean_total	$\log(B_{\text{mean,E}} + B_{\text{mean,W}})$	uniform	12.6 <sup>a</sup>	16.2	1
Bmean_prop_stock1 (=pE)	$B_{\text{mean,E}}/(B_{\text{mean,E}} + B_{\text{mean,W}})$	beta[0.1,0.6] <sup>b</sup>	0.344	0.072	1
recruitment.YCS	year-class strengths	lognormal	1	0.95	64
q[CSacous].q	catchability, CSacous	lognormal	0.77	0.77	1
q[WCacous].q	catchability, WCacous	lognormal	0.57	0.68	1
q[CRsum].q	catchability, CRsumbio	lognormal	0.15	0.65	1
q[SASum].q	catchability, SASumbio	lognormal	0.17	0.61	1
q[SAAut].q	catchability, SAAutbio	lognormal	0.17	0.61	1
natural_mortality	$M_{\text{male}} \& M_{\text{female}}$ ages 5–9	lognormal	0.182	0.509	8,0
natural_mortality.all	$M$	lognormal	0.298	0.153	0,1
process error c.v.s		uniform	0.1	1	7
selectivity[Wspsl].shift_a	Wspsl shift	normal	0	0.25	1
migrations	Whome, Wspmng, Espmng	uniform	various		40,24
comm. selectivities	Epspl,Wspsl,Enspsl,Wnspsl	uniform	various		8,9
surv. selectivities	CRsl, SASl	uniform	various		<u>6</u>
					141,119

<sup>a</sup> A lower bound of 13 was used for run 4.5

<sup>b</sup> This is a beta distribution scaled to have its range from 0 to 0.6, rather than the usual 0 to 1

In addition to the priors, bounds were imposed for all parameters with non-uniform distributions. For the catchability parameters these were those calculated by O'Driscoll et al. (2002) (where they are called “overall bounds”); for other parameters they were usually set at the 0.001 and 0.999 quantiles of their distributions. Some bounds were adjusted in some runs to avoid poor model behaviour; these adjustments did not appear to have a significant effect on the model results.

Penalty functions were used for three purposes. First, any parameter combinations that caused any exploitation rate to exceed its assumed maximum (Section 3.3) were strongly penalised. Second, the most recent YCSs were forced to be the same for E and W (but this penalty was dropped in MCMC runs) (Section 3.3). The third use of penalty functions was to link the spawning migration ogives for the two stocks (as per the constraints in Table 22).



#### 4. INITIAL ASSESSMENT RESULTS

The aim of the analyses presented in this section was to apply the ideas obtained from the preliminary analyses (Section 2), using the data new to this assessment, to investigate any problems that arose, and then to decide which runs would be used in the formal assessment (whose results are presented in Section 5).

The first MPD runs using all the new observations were labelled 2.1 and 2.2, and were based on final runs 4.4 and 4.5 in the 2007 assessment (see Table 1) with the following changes:

- all runs used the new year’s data, and the modified pspawn data, as described in Section 3;
- all runs assumed natal fidelity for reasons given in Section 2.7;
- no biomass indices were upweighted; and
- the assumed value of steepness was changed to 0.75 from 0.9 for reasons given in Section 2.4.

Both new runs showed a poor fit to SAsumbio (Figure 36A). When the trawl-survey biomass indices were upweighted, in runs 2.3 and 2.4, the fits to SAsumbio improved, but were still not good (Table 29). In particular, these runs failed to replicate the large increase between the biomass estimates from the last two surveys (Figure 36B). The conflict between these two surveys seems to be confined to the biomass estimates, because the fits to the associated proportions at age data are no worse than those from other years (Figures 37 & 38).

The apparent conflict between the last two SAsumbio surveys was investigated in runs 2.5 and 2.6, which were the same as run 2.4, except that the one year’s estimate was omitted in each. The fit to the remaining part of SAsumbio was no better when the 2007 estimate was dropped (in run 2.5), but was adequate when the 2008 estimate was omitted (run 2.6) (Table 29, Figure 36C). This result supports the conclusion in Section 2.6 that the conflict between these two surveys was more likely to be caused by abnormally high catchability in the last survey, rather than abnormally low catchability in the previous one.

**Table 29: Goodness of fit to biomass indices, as measured by the SDNR (standard deviation of the normalised residuals), for some new model runs. For this table the normalised residuals were calculated using the original c.v.s (i.e., ignoring changes in c.v.s for upweighting trawl biomass data sets).**

Run	Description	CRsumbio	SAsumbio	SAautbio	CSacous	WCacous
2.1	As 4.4, no upweighting	0.84	1.64	0.83	1.01	1.14
2.2	As 4.5, no upweighting	0.90	1.82	0.83	1.03	0.95
2.3	As 4.4, trawl upweighted	0.77	1.47	0.82	1.01	1.16
2.4	As 4.5, trawl upweighted	0.78	1.44	0.94	1.06	1.05
2.5	As 2.3, drop 2007 SAsumbio	0.77	1.45	0.82	1.01	1.15
2.6	As 2.3, drop 2008 SAsumbio	0.76	1.02	0.81	1.01	1.17

The final two MPD runs show that the assessment is not very sensitive to the assumed value of steepness. A lower value (0.6, in run 2.7) slightly improved the overall fit and decreased the estimates of current biomass, and a higher value (0.9, in run 2.8) had the opposite effect (Table 30).

**Table 30: Comparison of key aspects of all initial MPD fits.**

Run	Description	Objective function	$B_{\text{current}}(\%B_0)$	
			E	W
2.1	As 4.4, no upweighting	-316.7	51.2	33.2
2.2	As 4.5, no upweighting	-388.4	46.3	45.0
2.3	As 4.4, trawl upweighted	-287.5	45.6	25.4
2.4	As 4.5, trawl upweighted	-353.9	38.6	28.5
2.5	As 2.3, drop 2007 SAsumbio	-290.4	44.9	28.5
2.6	As 2.3, drop 2008 SAsumbio	-301.5	47.2	20.6
2.7	As 2.3, steepness = 0.6	-288.2	44.7	24.2
2.8	As 2.3, steepness = 0.9	-286.9	46.2	26.3

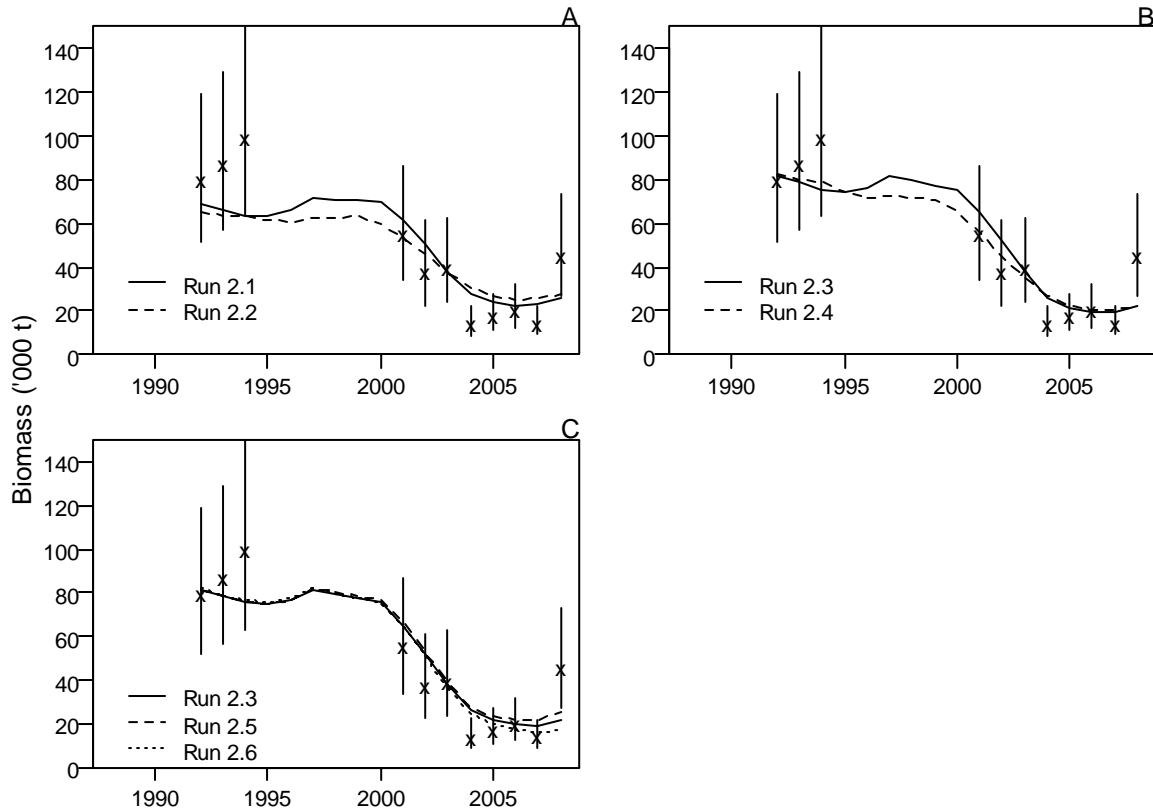


Figure 36: Fits to SASumbio for runs 2.1 to 2.6, showing observed ('x', with vertical lines showing 95% confidence intervals) and expected values (lines). Plotted years are as in the model (so the last survey is plotted at 2008).

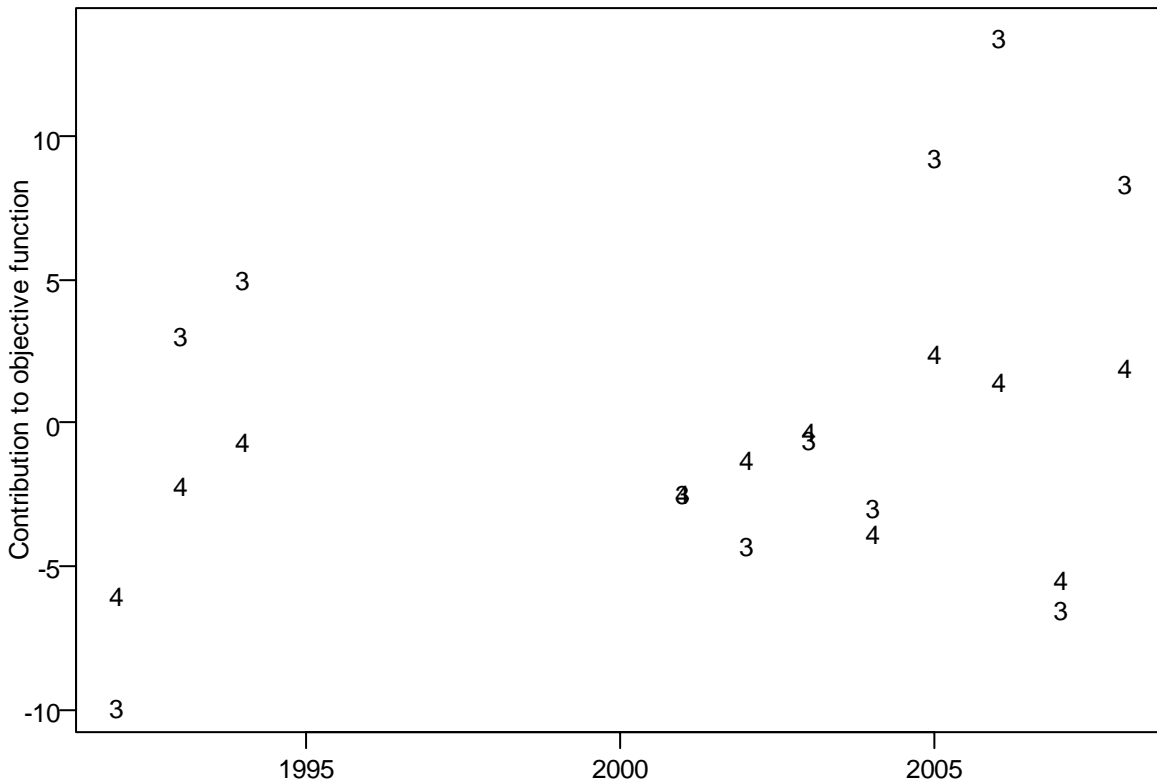
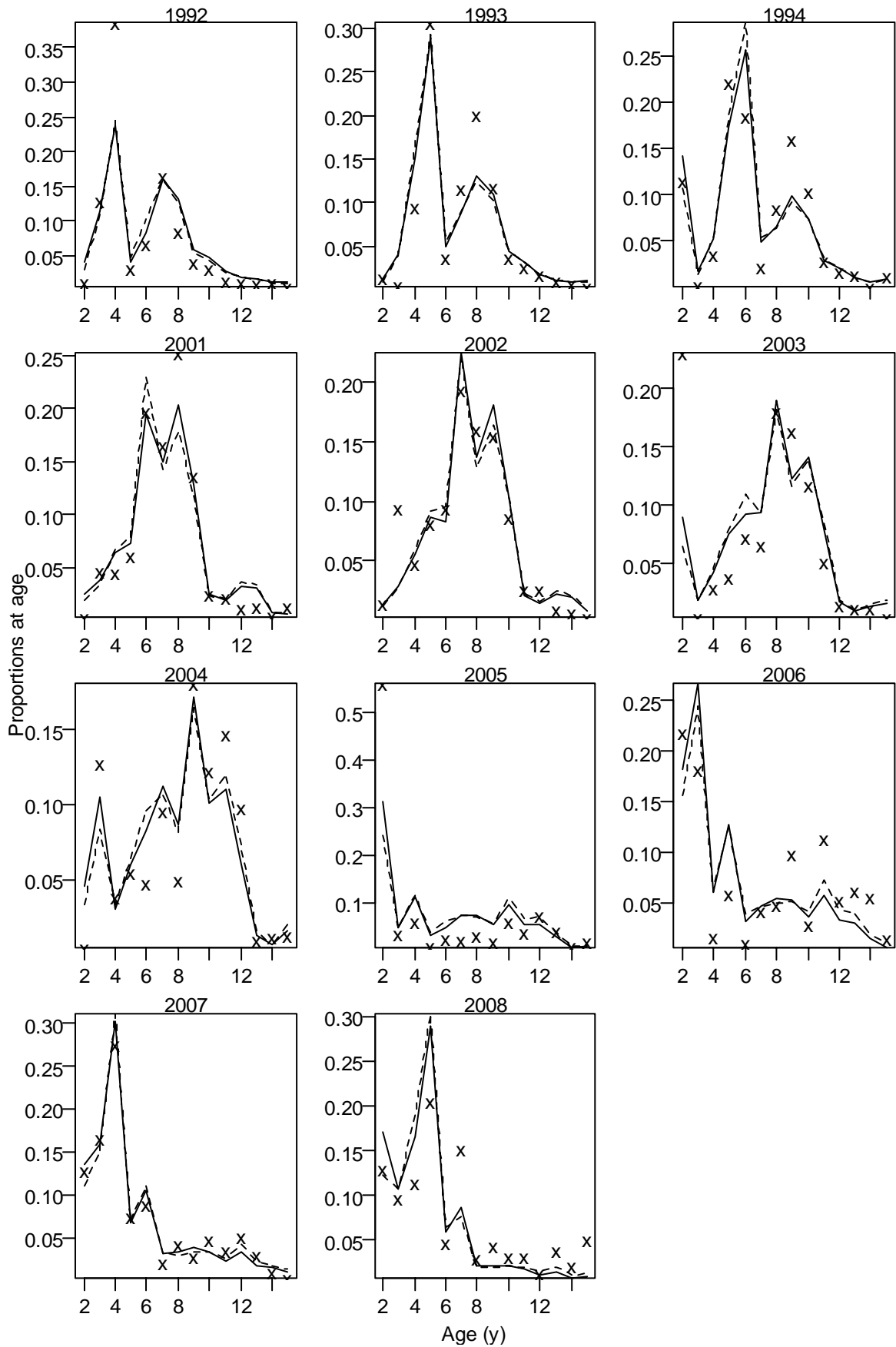


Figure 37: Goodness of fit to the SASumage data, by year, for runs 2.3 ('3') and 2.4 ('4'). The y-value for each point indicates the contribution to the objective function from one year's SASumage data; smaller values indicate a better fit. Plotted years are as in the model (so the last survey is plotted at 2008).



**Figure 38: Observed ('x') and expected (lines) proportions at age in the summer sub-Antarctic survey (data set SASumage) for runs 2.3 (solid line) and 2.4 (broken lines).**

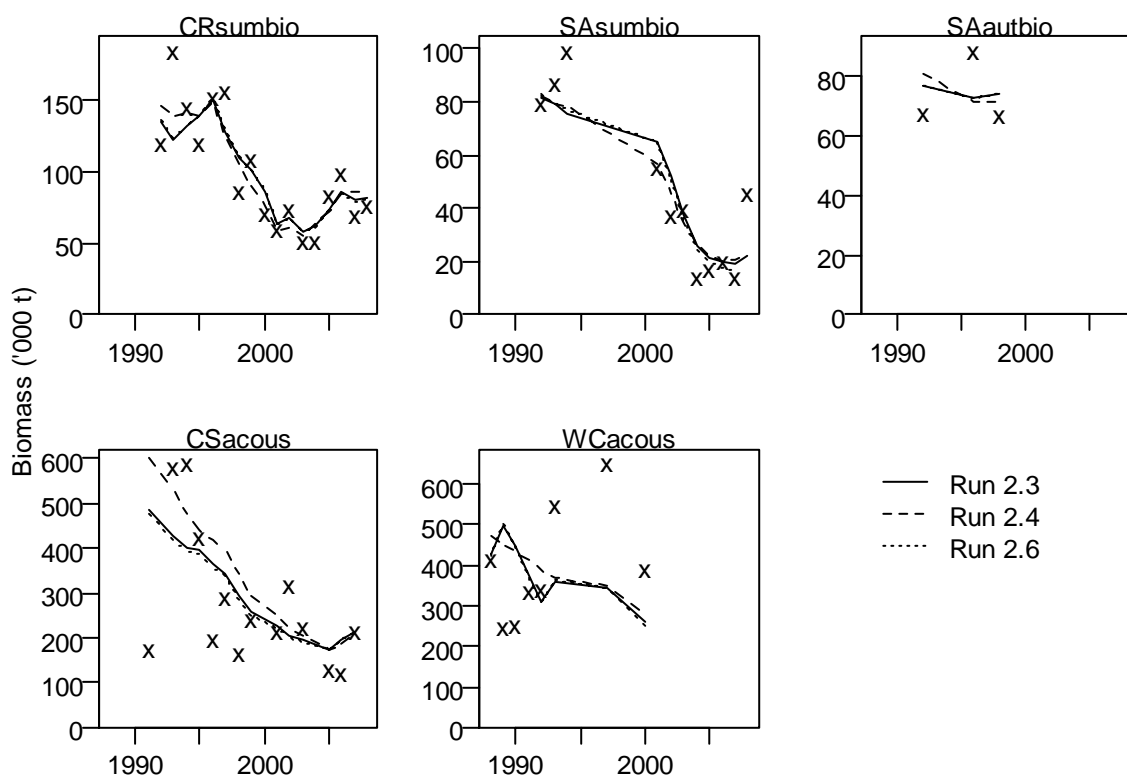
The remaining results in the section concern just runs 2.3, 2.4, and 2.6, which were selected as being the provisional main runs for the 2008 assessment. There were two reasons for choosing run 2.6, rather than 2.5. First, the 2008 SAsumbio estimate is more of an outlier than is that from 2007 (Figure 39). The assessment model allows (via process error) for year-to-year variation in survey catchability. What seems likely is that the catchability for the last survey was outside the range allowed for. The second reason for preferring run 2.6 is that it extends the range of estimates of  $B_{\text{current}}$  in runs 2.3 and 2.4, but run 2.5 does not (see Table 30).

In comparison to the 2007 runs, all the new runs produce more optimistic estimates of the status of the W stock in 2007, and estimates of unfished biomass ( $B_0$ ) were almost always slightly higher (Table 31). This is exactly the opposite of what was found when the 2007 runs were compared to those from 2006 (cf table 24 of Francis (2008b)).

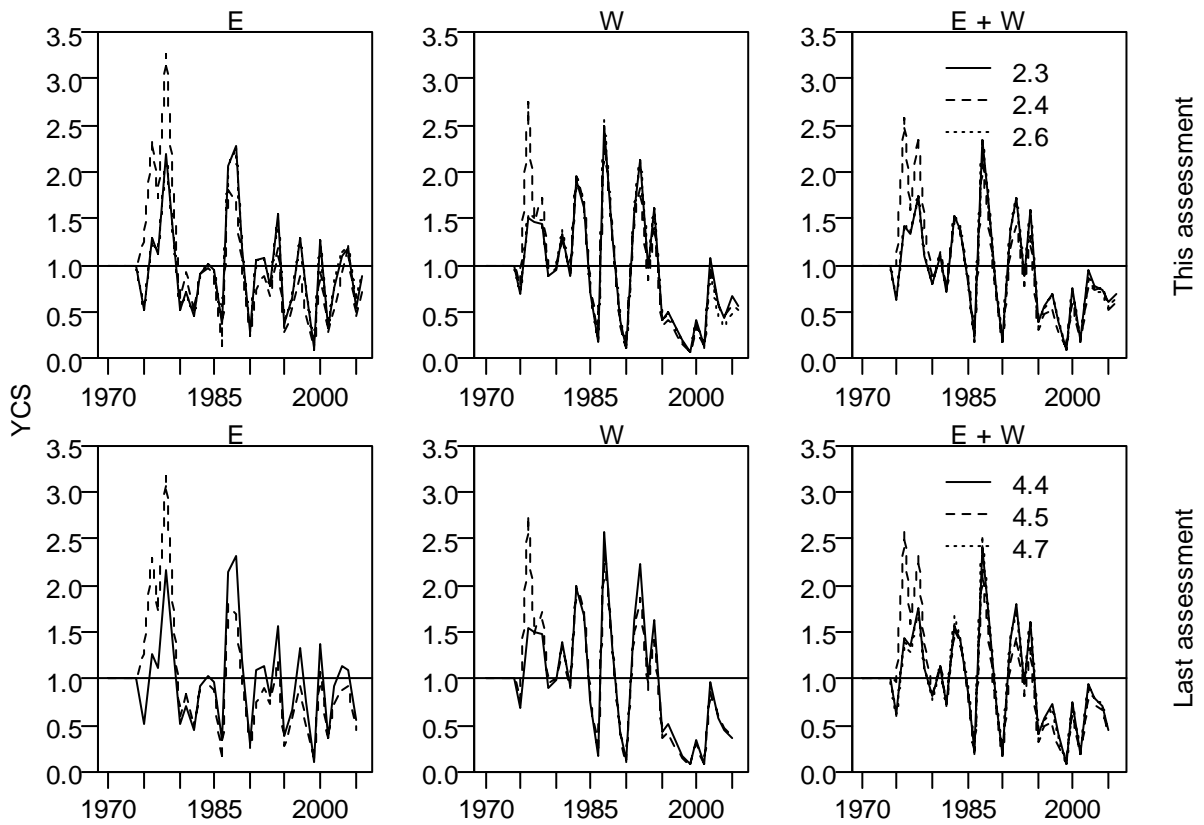
**Table 31: Comparison of old and new biomass estimates for the individual stocks, E and W, and the combined E + W stock. In each group of runs, the first is from 2007 and the other(s) is (are) analogous run(s) for 2008.**

Run	$B_0$ ('000 t)			$B_{2007}(\%B_0)$			$B_{2008}(\%B_0)$		
	E	W	E + W	E	W	E + W	E	W	E + W
4.4	501	856	1357	45	16	27	–	–	–
2.3	510	880	1391	45	23	31	46	25	33
2.6	508	873	1381	45	18	28	47	21	30
4.5	636	1144	1779	34	23	27	–	–	–
2.4	637	1062	1699	37	26	30	39	29	32

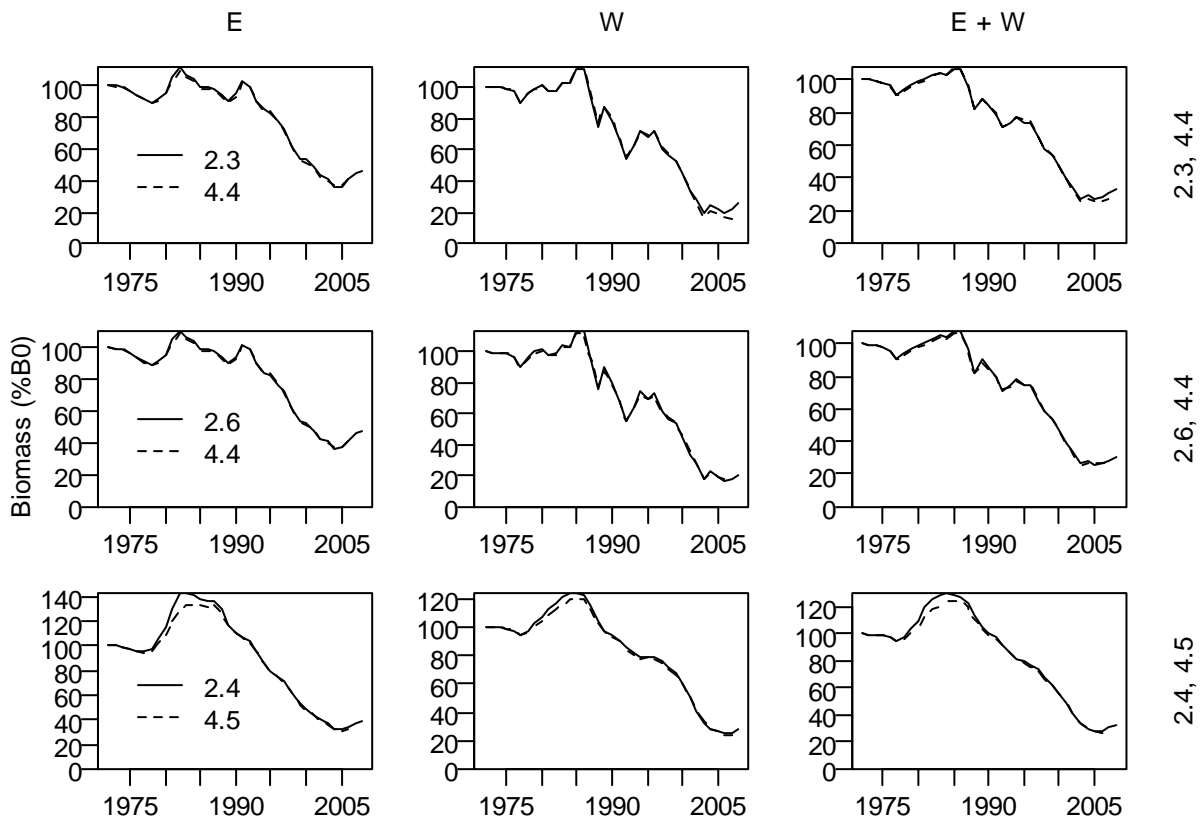
Figures 40 & 41 show that estimated YCSs and biomass trajectories differ little from those in the previous assessment. Other plots (not shown) showed that this was also true for exploitation rates, and ogives for selectivity, migration, and natural mortality.



**Figure 39: Fits to biomass indices for runs 2.3, 2.4 and 4.6, showing observed ('x') and expected values (lines).**



**Figure 40:** YCS estimates for new runs 2.3, 2.4, and 2.6 (upper panels) and the main runs from last year's assessment (lower panels). Note that for run 4.7, where there is no natal fidelity, YCSs are defined only for the E + W stocks combined.



**Figure 41:** Comparison of biomass trajectories from different runs: E stock (left column), W stock (middle column), and E + W stocks combined (right column). Each rows of panels compares a new run (solid lines) with the corresponding run from 2007 (broken lines).

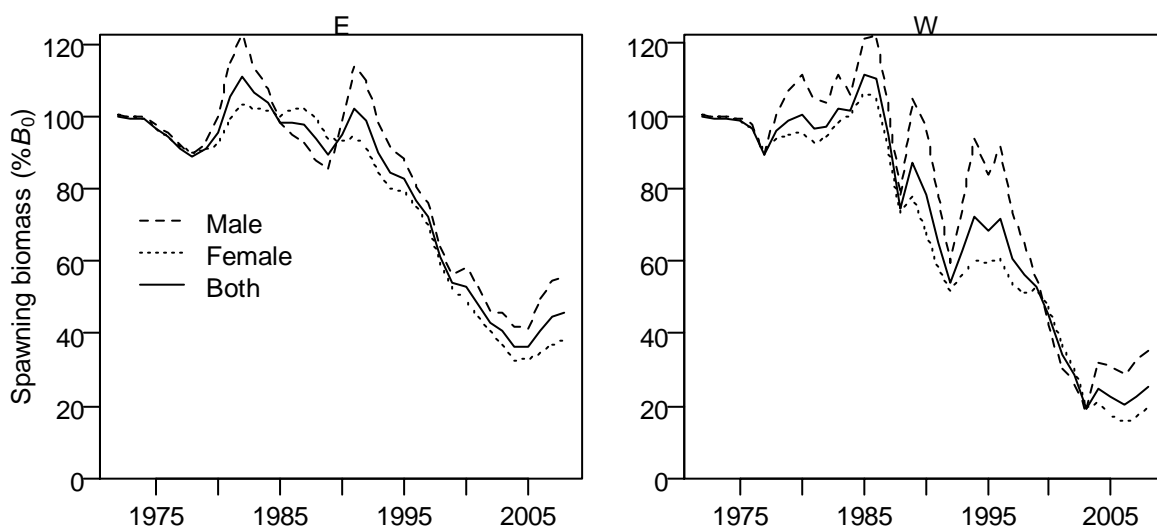
#### 4.1 Spawning biomass trajectories by sex

In a recent review of the use of MSY-based reference points it was recommended that spawning biomass trajectories should be presented for females only (as is routinely done in other countries), rather than for both sexes combined (as has been the practice in New Zealand). However, it was also noted that a case could be made for using male biomass for hoki, since males are typically in the minority in spawning populations of this species. The Hoki Working Group decided that the sensitivity of biomass trajectories to sex should be investigated, but that no change would be made to the presentation of results in the Working Group Report until a decision has been made, for all Working Groups, about this matter. The investigation was carried out for run 2.3 only, because sex is ignored in run 2.4.

For run 2.3, the spawning biomass of both stocks is more depleted for females than for males (Figure 42, Table 32). This difference by sex may be explained in terms of changes in the age structure of the population. It is well known that, for hoki, males and females occur in roughly equal numbers at age 1, but that there is an increasing predominance of females with increasing age (see figure 22 in Francis 2007) which is presumably caused by higher natural mortality rates for males. Thus we should expect the proportion of females to be positively correlated with mean age in the population. This correlation is apparent both in the long-term decrease in mean age that occurs as the stocks are fished down, and in the year-to-year variations associated with the arrival of strong or weak year classes (Figure 43). The effect of the long-term trend in proportion female is that the biomass of females is more depleted than that of males, as seen in Table 32.

**Table 32: Current status (MPD estimates of spawning biomass in 2008, expressed as % $B_0$ ) for each stock from run 2.3, for three definitions of spawning biomass: both sexes combined, males only, or females only.**

	Definition of spawning biomass		
	Both sexes	Male	Female
E	46	56	38
W	25	35	20



**Figure 42: Trajectories of spawning biomass (as a percentage of  $B_0$ ) from MPD of run 2.3, plotted by sex and for both sexes.**

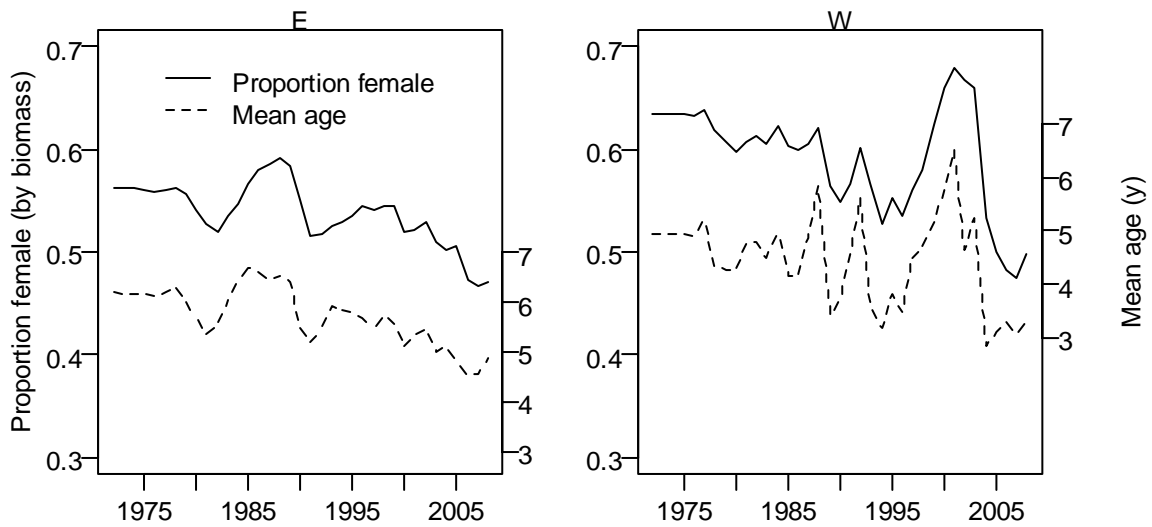


Figure 43: Plot showing how the proportion female in the spawning biomass (solid lines, left-hand scales) from MPD of run 2.3 is closely related to the mean age (broken lines, right-hand scales) of that biomass.

## 5. FINAL ASSESSMENT RESULTS

After much discussion, the Hoki Working Group decided that the final model runs presented in the Working Group Report should be 2.3 and 2.4. My view was that, because of the uncertainty about the last sub-Antarctic survey (see Section 2.6), it was important to include run 2.6, which omitted the biomass estimate from that survey. Thus I have chosen to present all three runs here. They are distinguished by three characteristics (Table 33).

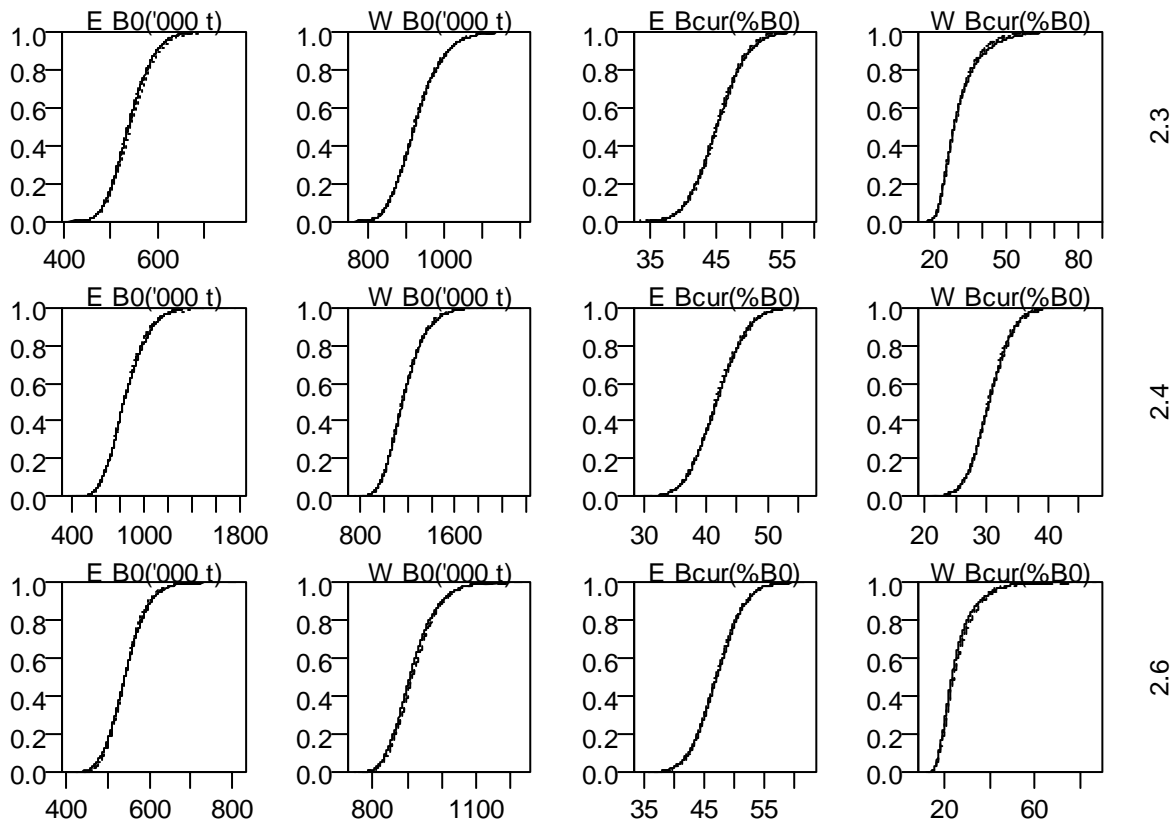
Table 33: Distinguishing characteristics for the two final model runs selected by the Hoki Working Group (2.3 and 2.4) and an additional run, 2.6, whose results are presented here.

Label	Response to lack of old fish in the observations	Sex in model and selectivities length-based?	Last sub-Antarctic survey biomass dropped?
2.3	$M$ dependent on age	Yes	No
2.4	Domed spawning selectivity	No	No
2.6	$M$ dependent on age	Yes	Yes

Three MCMC chains of length 2 million samples were created for each of these runs, each chain having a different starting point, which was generated by stepping randomly away from the MPD. As in 2007, those migration or selectivity parameters that were found to be at a bound in the MPD run (Table 34) were fixed for the MCMC runs in order to improve convergence. Also, for run 2.4, upper bounds for some selectivity parameters were reduced to improve MCMC performance, exactly as was done for run 4.5 in 2007 (see section 7.5 in Francis (2008b)). Diagnostic plots comparing the three chains for each run suggest reasonably good convergence for all runs (Figure 44). For all the remaining results, the first quarter of each chain was discarded, the three chains for each run were concatenated, and the resulting chain was thinned to produce a posterior sample of length 1000.

Table 34: Migration and selectivity parameters held fixed in MCMC runs (with fixed values in parentheses).

Run	Parameters (fixed values)
2.3	WspmgM2(1), EspmgF8[0.6], WspmgF8[0.6], Enspsl.sR(44), CRsl.a1(64), SAsl.a1(84), SAsl.sR(44)
2.4	Whome.6(1), CRsl.a1(1), CRsl.sL(1)
2.6	Whome.2(0.01), WspmgM2(1), EspmgF8[0.6], WspmgF8[0.6], Enspsl.sR(44), CRsl.a1(64), SAsl.a1(84), SAsl.sR(44)



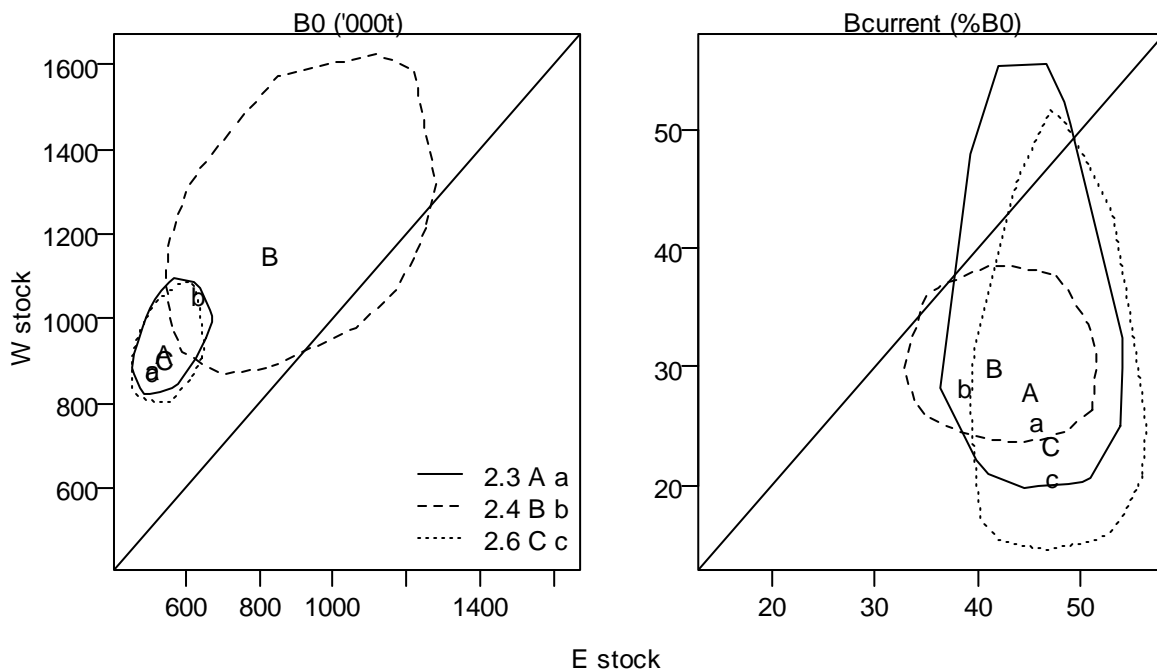
**Figure 44: Diagnostics for MCMC chains for the three runs: 2.3 (top row), 2.4 (middle row), and 2.6 (bottom row). Each panel contains cumulative probability distributions, for  $B_0$  or  $B_{current}$ , for three chains from the same model run.**

The MCMC results show, as in 2007, that the W spawning stock was originally much larger than the E, but is currently about the same size or smaller (between 210 000 t and 340 000 t) and so is more depleted (Table 35, Figure 45). Also, the three runs, taken together, indicate a wider range of uncertainty than any one run (Figure 45). In terms of estimated biomass in 2007, the new assessment is similar to that from last year (Figure 46). All runs suggest the W stock is rebuilding, both in absolute terms (Figure 47) and relative to  $B_0$  (Figure 48). As in previous years, the selectivity and migration ogives for runs assuming age-dependent natural mortality (runs 2.3 and 2.6 in the current assessment) are very different from those in runs without this assumption (Figures 49, 50). As suggested by the MPD runs, recent W YCSs are estimated to be below the long-term average, but higher than in the seven-year period of very weak recruitment, 1995–2001 (Figure 51).

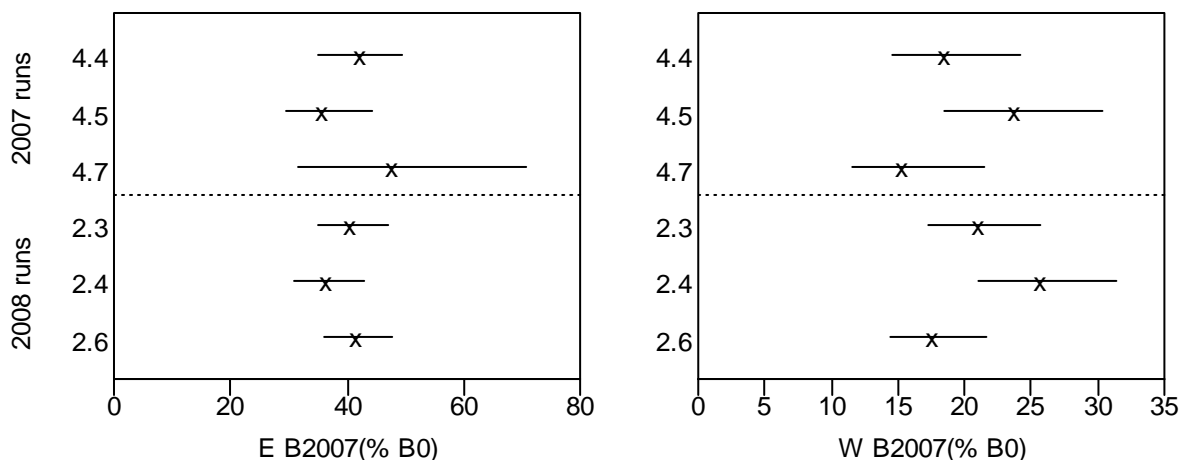
**Table 35: Estimates of spawning biomass (medians of marginal posterior, with 95% confidence intervals in parentheses) for the three final runs.  $B_{current}$  is the biomass in mid-season 2007.**

Run	$B_0$ ('000 t)		$B_{current}$ ('000 t)		$B_{current}$ (% $B_0$ )		
	E	W	E	W	E	W	E+W
2.3	542(468,638)	925(832,1058)	245(192,310)	263(179,464)	45(38,52)	28(20,48)	34(29,47)
2.4	828(588,1220)	1155(915,1555)	346(232,532)	349(242,528)	42(34,50)	30(25,37)	35(30,40)
2.6	540(466,637)	909(813,1057)	254(203,320)	212(131,446)	47(40,55)	24(15,45)	32(26,45)

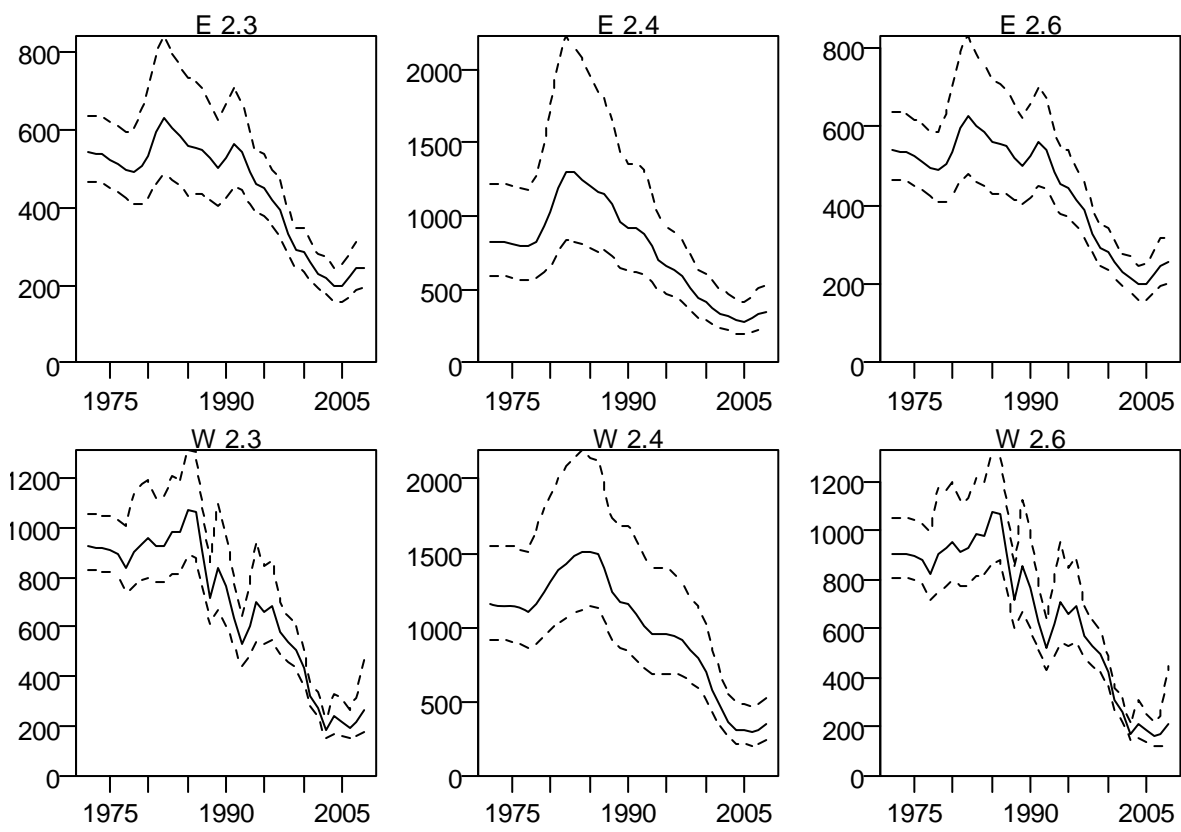




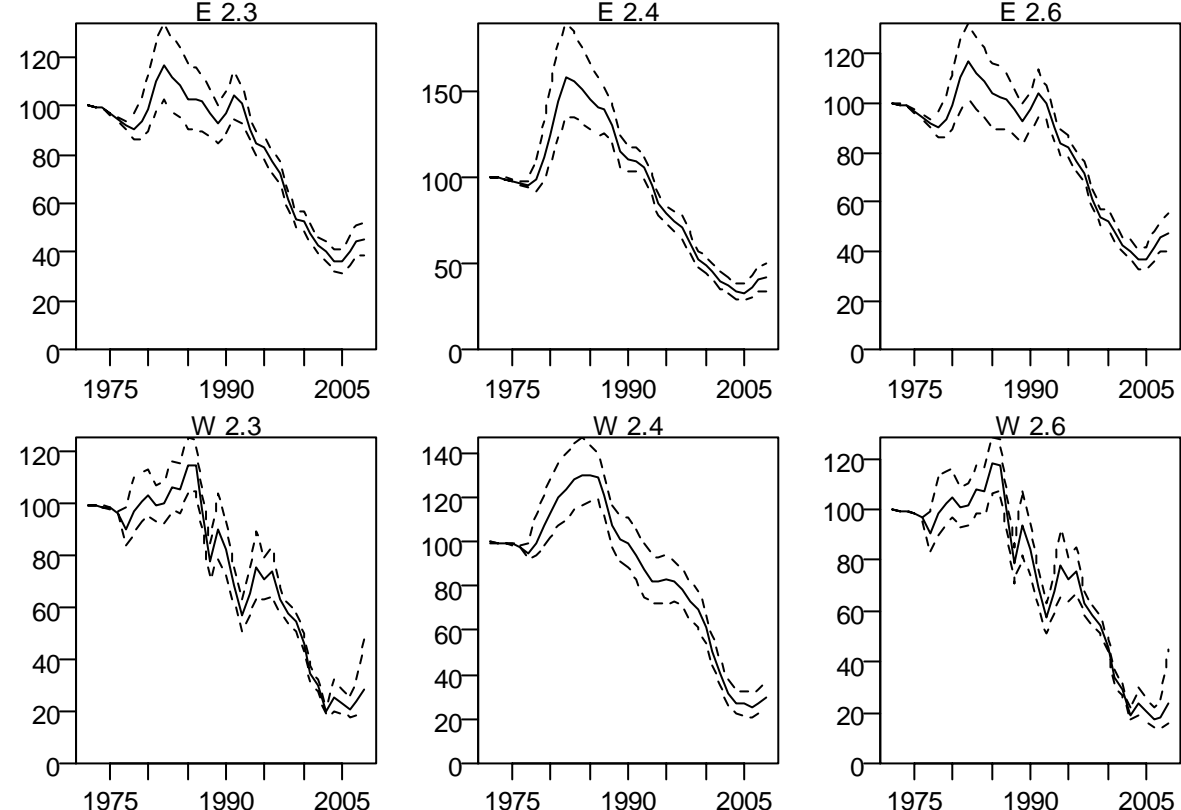
**Figure 45: Estimates and approximate 95% confidence intervals for virgin ( $B_0$ ) and current ( $B_{\text{current}}$  as  $\%B_0$ ) biomass by stock for the three final runs, 2.3, 2.4, and 2.6. In each panel the points 'A', 'B', 'C' indicate best estimates (median of the posterior distribution) for these three runs, 'a', 'b', 'c' are the MPD estimates, and the polygons (with solid, broken and dotted lines, respectively) enclose approximate 95% confidence intervals. Diagonal lines indicate equality ( $y = x$ ).**



**Figure 46: Comparison of 2008 runs (2.3, 2.4, 2.6) with those from 2007 (4.4, 4.5, 4.7): estimates of stock status in 2007 ( $B_{2007}$  as  $\%B_0$ ), with 95% confidence intervals shown as horizontal lines.**



**Figure 47:** Estimated spawning-biomass trajectories from the three MCMC runs, showing medians (solid lines) and 95% confidence intervals (broken lines) by run for E (upper panels) and W (lower panels).



**Figure 48:** As for Figure 47, but plotted as  $\%B_0$ .

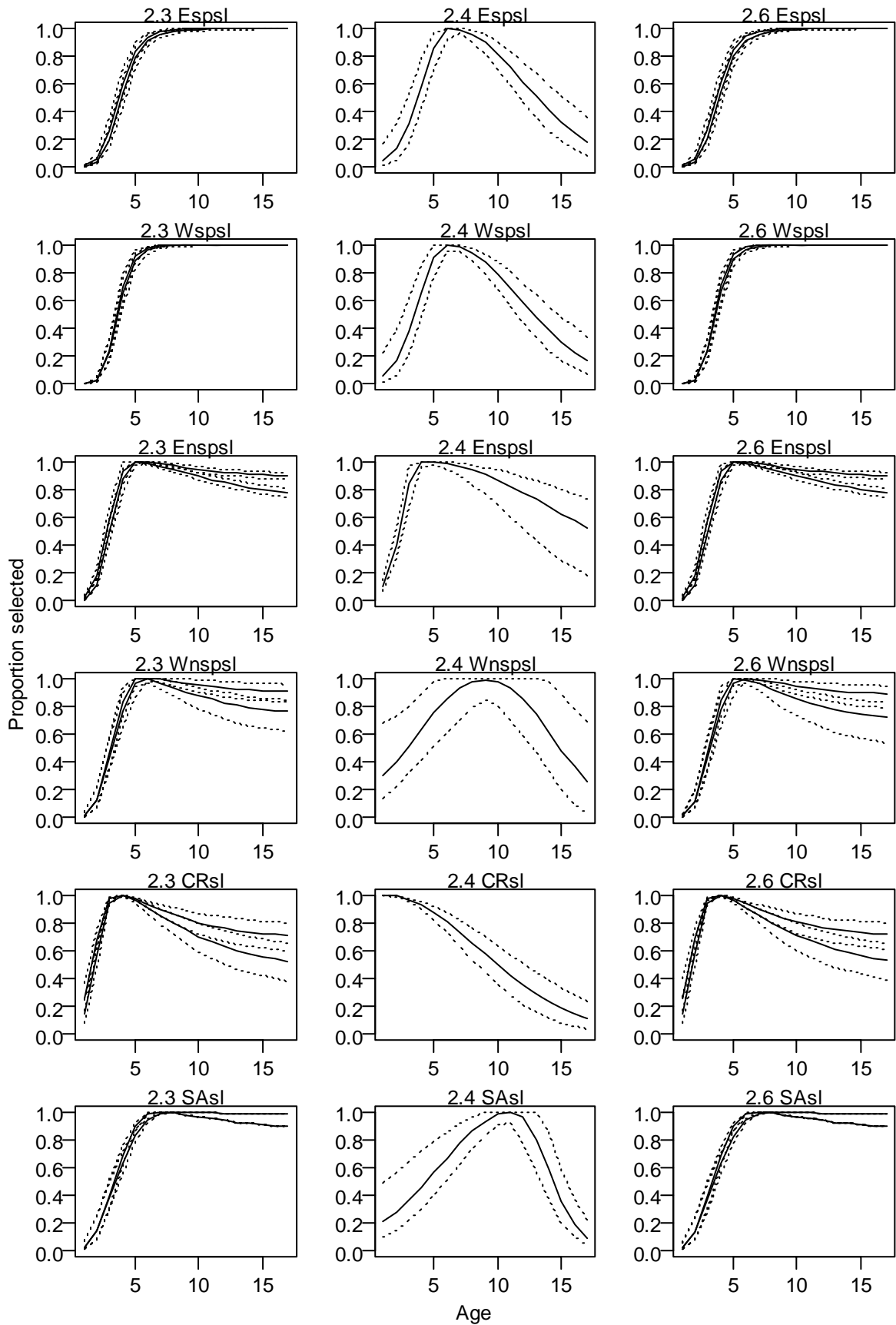
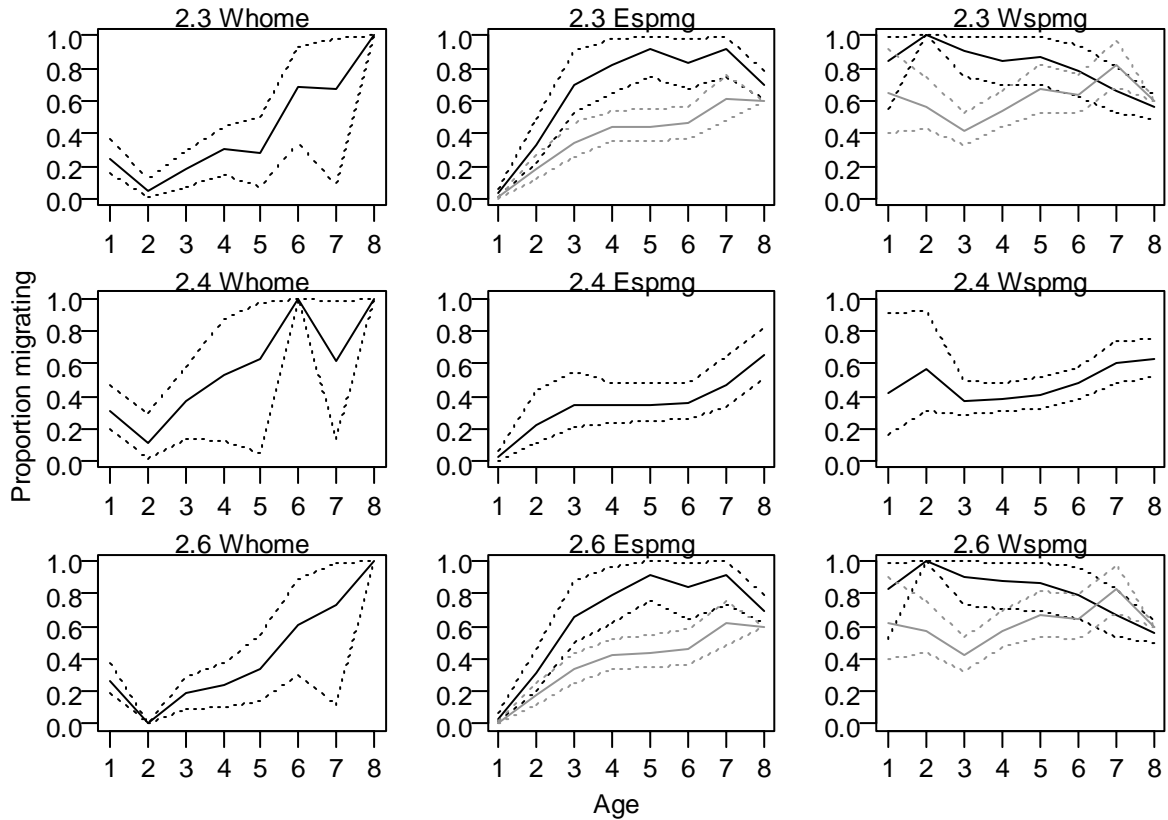
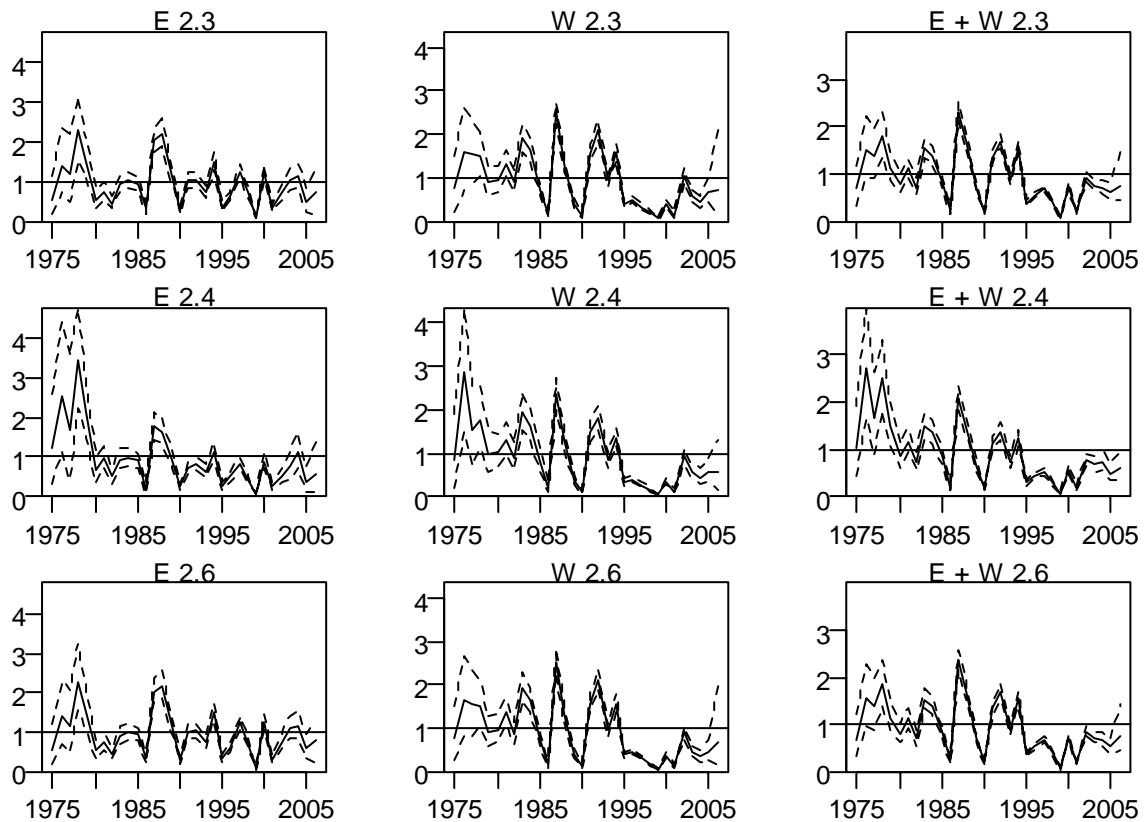


Figure 49: Posterior estimates of selectivity ogives for each of the three MCMC runs. Solid lines are medians, broken lines show 95% confidence intervals. Where ogives differ by sex they are plotted as black for males and grey for females. Where they differ by stock or time step the plotted curves are for one selected combination (E step 2 for Enspsl and CRsl, W step 2 for Wnspsl and SAsl).

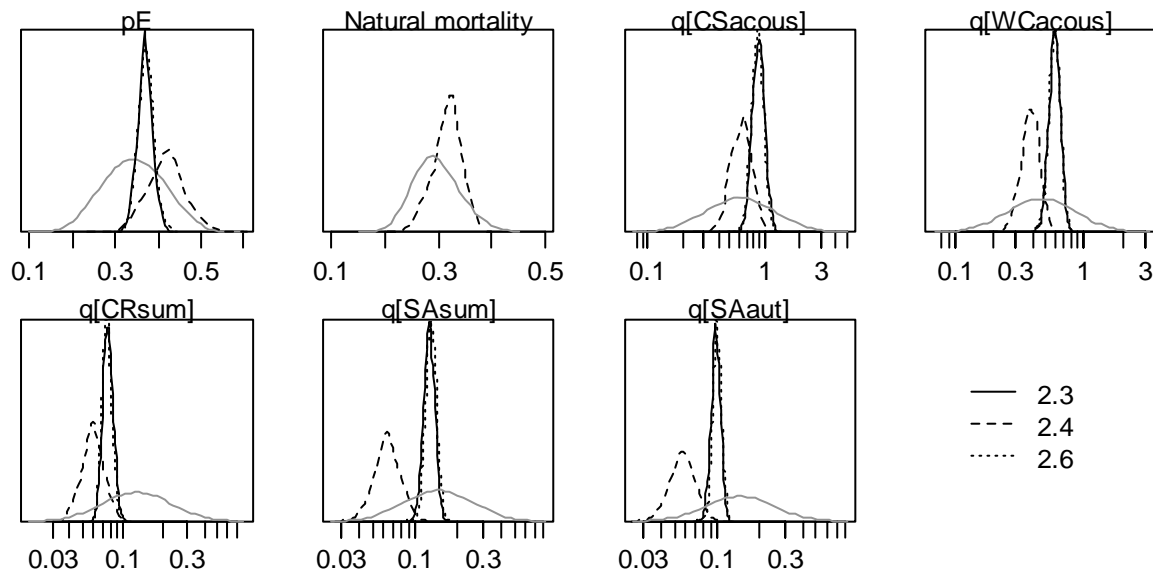


**Figure 50: Migration ogives estimated in each of the three MCMC runs. Solid lines are medians, broken lines show 95% confidence intervals. Where ogives differ by sex they are plotted as black for males and grey for females.**

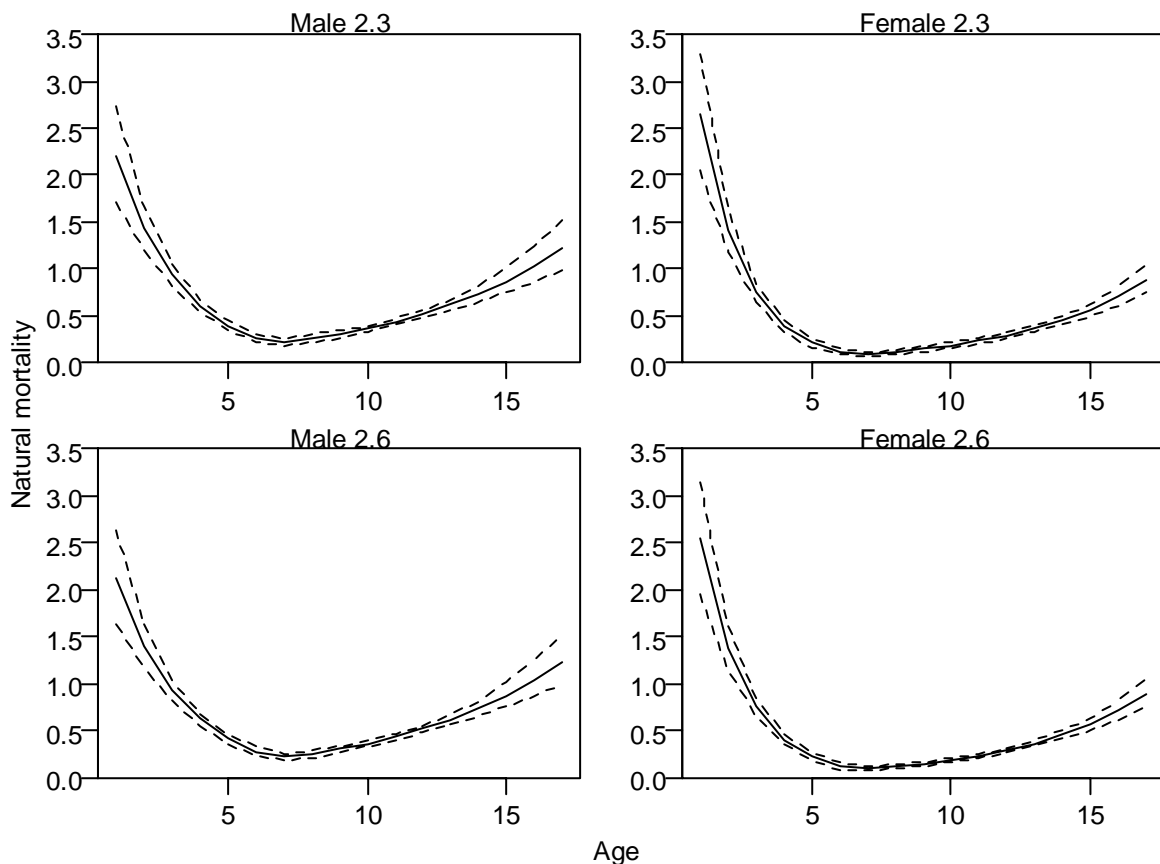


**Figure 51: Estimated year-class strengths (YCSs) from the MCMC runs, showing medians (solid lines) and 95% confidence intervals (broken lines) by run for E (left panels), W (middle panels) and E + W (right panels).**

A comparison of priors and posteriors for various parameters (Figure 52) showed no substantial changes from last year except that the natural mortality posterior for the run with domed selectivities has moved to the right (median 0.32 in run 2.4, compared to 0.27 in run 4.5 last year). Estimates of natural mortality, for the two runs in which this is age-dependent, are quite similar (Figure 53).



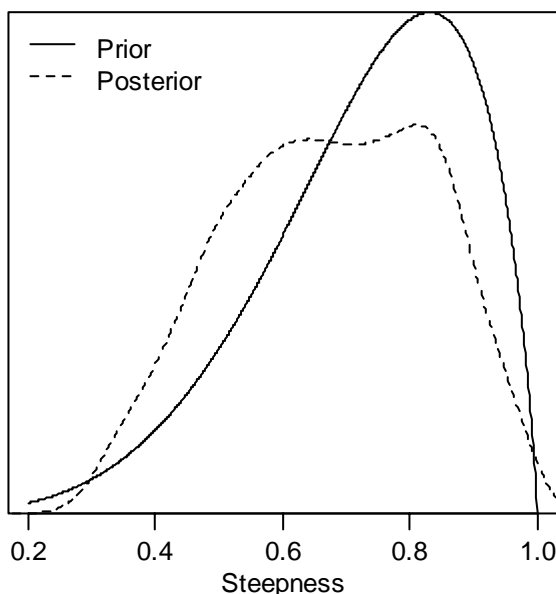
**Figure 52:** Prior (grey lines) and estimated posterior (black lines) distributions from the three MCMC runs for the following parameters: pE (proportion of  $B_0$  in E stock), natural mortality (independent of age, run 2.4 only), and survey catchabilities (acoustic and trawl).



**Figure 53:** Estimates of age-dependent natural mortality ogives for runs 2.3 and 2.6, showing median estimates (solid lines) and 95% confidence intervals (broken lines) for each sex.

## 5.1 Other analyses

Two other analyses were requested by the Hoki Working Group. The first was to investigate the effect of estimating stock-recruit steepness, using the ‘Domain 2: Not anadromous’ prior of Myers et al. (2002) (see Figure 22 above). Run 2.9 was the same as 2.3, except that steepness (assumed the same for both stocks) was estimated using a prior (a 4-parameter beta distribution with parameters  $A = -0.5$ ,  $B = 1.0$ ,  $\mu = 0.741$ ,  $\text{stdev} = 0.156$ , and bounds 0.2 and 1) that was chosen to approximate this prior [Myers et al. (2002) did not provide a parametric form for their prior]. For this run, the posterior for steepness differed only slightly from the prior, being shifted somewhat to the left of it (Figure 54).



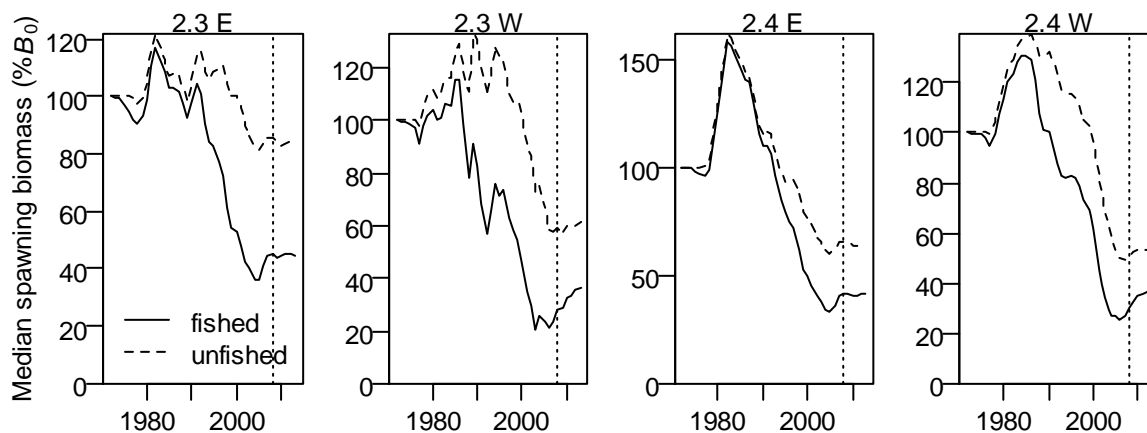
**Figure 54: Prior and posterior distributions for steepness in run 2.9.**

The effect of estimating steepness was small (Table 36; see also figures 6–8 in Francis (2008a) for plots of YCSs, biomass trajectories, and projections).

**Table 36: Effect on estimates of spawning biomass of estimating steepness; results (medians of marginal posteriors, with 95% confidence intervals in parentheses) for runs 2.3 (steepness fixed) and 2.9 (steepness estimated).  $B_{\text{current}}$  is the biomass in mid-season 2008.**

Run	$B_0$ ('000 t)		$B_{\text{current}}$ ('000 t)		$B_{\text{current}}$ (% $B_0$ )		
	E	W	E	W	E	W	E+W
2.3	542(468,638)	925(832,1058)	245(192,310)	263(179,464)	45(38,52)	28(20,48)	34(29,47)
2.9	552(465,671)	949(818,1180)	242(194,314)	259(173,491)	44(36,52)	27(18,49)	34(26,47)

The other requested analysis was way to compare estimated biomass trajectories from runs 2.3 and 2.4 with would have occurred had there been no fishing. This provides another way of illustrating the effect of both fishing and YCS on the population (Figure 55).

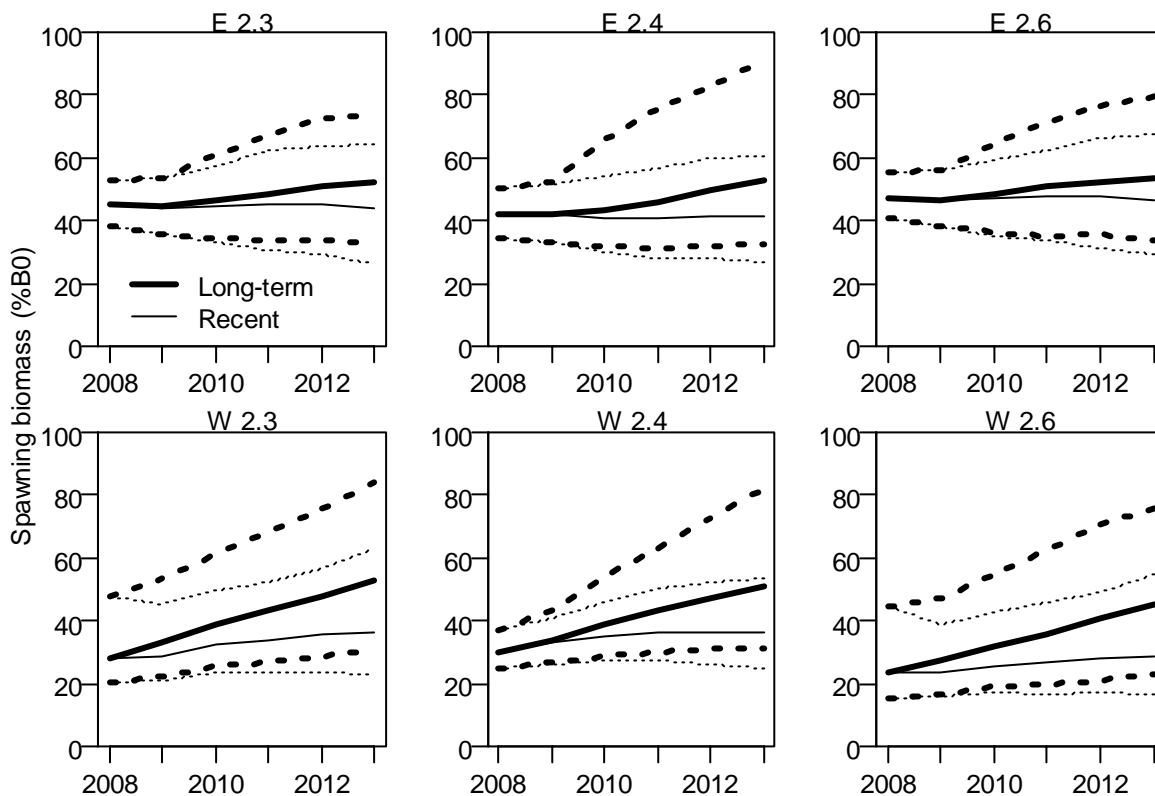


**Figure 55: Comparison between the historical spawning biomass trajectory (‘fished’, solid line) and that which would have occurred had there been no fishing (‘unfished’, broken line). Trajectories in the future (to the right of the vertical broken line) assume ‘recent’ recruitment and, for the ‘fished’ trajectory, the same catches as in 2008.**

## 5.2 Projections

Five-year projections were carried out for each run with two alternative recruitment scenarios: ‘long-term’ (future recruitment selected from estimated levels in 1975–2006) and ‘recent’ (recruitment selected from 1995–2006). Future catches for each fishery were assumed equal to those assumed for 2008 (see Table 19).

With long-term recruitment, median spawning biomass increased for both stocks in all years; with recent recruitment, it changed very little for the E stock but increased slowly for the W stock (Figure 56). Exploitation rates for the W stock are expected to decrease under both recruitment scenarios; for the E stock they will decrease under the long-term recruitment scenario, but remain fairly constant with recent recruitment (Figure 57).



**Figure 56:** Trajectories of median spawning biomass (as  $%B_0$ ) from the projections (solid lines) together with lower and upper bounds of a 95% confidence interval (broken lines) assuming ‘long-term’ (heavy lines) or ‘recent’ recruitment (light lines). Each panel shows results for one stock (E or W) from one of the three MCMC runs.

In previous assessments, the two most recent year classes were not included in those that were sampled to generate future recruitments for projections. This year the Working Group suggested it would be better to include these year classes. Thus, for the ‘recent’ recruitment scenario the years resampled were 1995–2006, and not 1995–2004, as they would have been had previous practice been followed. This change in assumptions made very little difference (Figure 58).

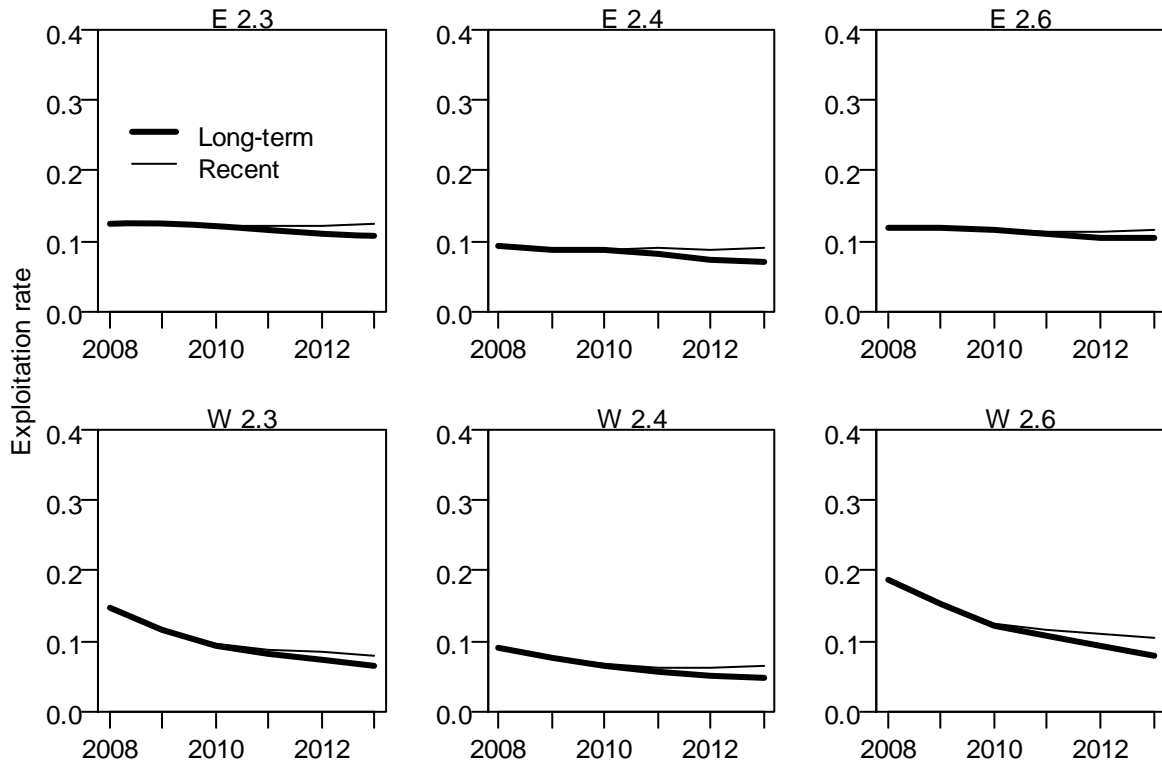


Figure 57: Estimated exploitation rates for 2008 and future years assuming long-term recruitment (heavy lines) or recent recruitment (light lines). Each panel shows results for one stock (E or W) from one of the three final runs (2.3, 2.4, 2.6).

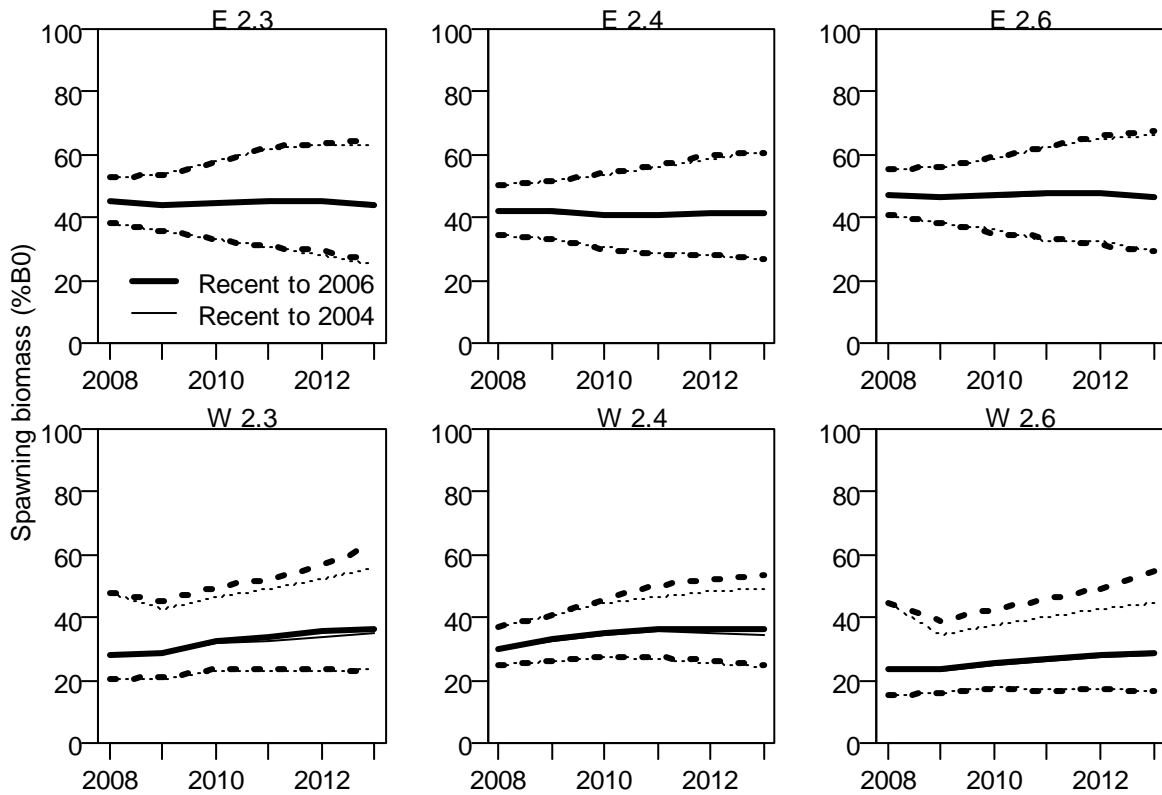


Figure 58: Effect on projections with 'recent' recruitment of defining the recent period as 1995–2006 (heavy lines) or 1995–2004 (light lines). Each panel shows results for one stock (E or W) from one of the three MCMC runs, with solid lines indicating median spawning biomass (as  $\%B_0$ ) and broken lines indicating 95% confidence intervals.



## 6. DISCUSSION

Both hoki stocks are estimated to be near their lowest level since the fishery began, with the W stock being much more depleted (24–30 % $B_0$ ), and closer to its lowest level, than the E stock (42–47 % $B_0$ ). The W stock experienced an extended period of poor recruitment from 1995 to 2001, but there is evidence of better (though still below average) recruitment in subsequent years (2002–06). Projections indicate that the current catch levels are likely to allow the W stock to rebuild; the E stock is likely to remain stable if future recruitment is similar to that in recent years, and to increase if that recruitment is comparable to the long-term average.

The uncertainty in this assessment is almost certainly greater than is implied by the confidence limits presented above. We may think of this uncertainty as having three types. The first of these – random error in the observations – is reasonably well dealt with in the assessment by the c.v.s that are assigned to individual observations. The second, which arises from annual variability in population processes (e.g., growth and migration – but not recruitment, which is modelled explicitly) and fleet behaviour (which affects selectivities), is more problematic. We deal with this, rather simplistically, by adding process error. This assumes that the structure of our model is correct “on average”, but that the real world fluctuates about that average. The problem is that we cannot be at all sure about this assumption. This leads to the third type of uncertainty: we cannot be sure that our model assumptions are correct on average. This is often called model uncertainty, and it is the type of uncertainty that is least well covered in this assessment. We have dealt with it by using three alternative model structures (expressed in our three final runs).

## 7. SOME POST-ASSESSMENT RESULTS

In this section I describe some additional analyses that were carried out after the assessment. These shed further light on the assessment, and may be of use in the 2009 assessment, but are presented separately because they were completed too late to affect the Working Group Report for 2008.

Three topics were investigated: models without natal fidelity, calculations for the RAM legacy database, and a reanalysis of the proportion spawning data from 1992 and 1993.

### 7.1. Fixing model runs without natal fidelity

This section fixes two problems with model runs that do not assume natal fidelity. I present results from a series of model runs to show the effect of these fixes. The first run, labelled 3.2, is an update of run 4.7 from the 2007 assessment. That is, it is the same as run 4.7, except that it uses the data that were new for the 2008 assessment and the new steepness value of 0.75.

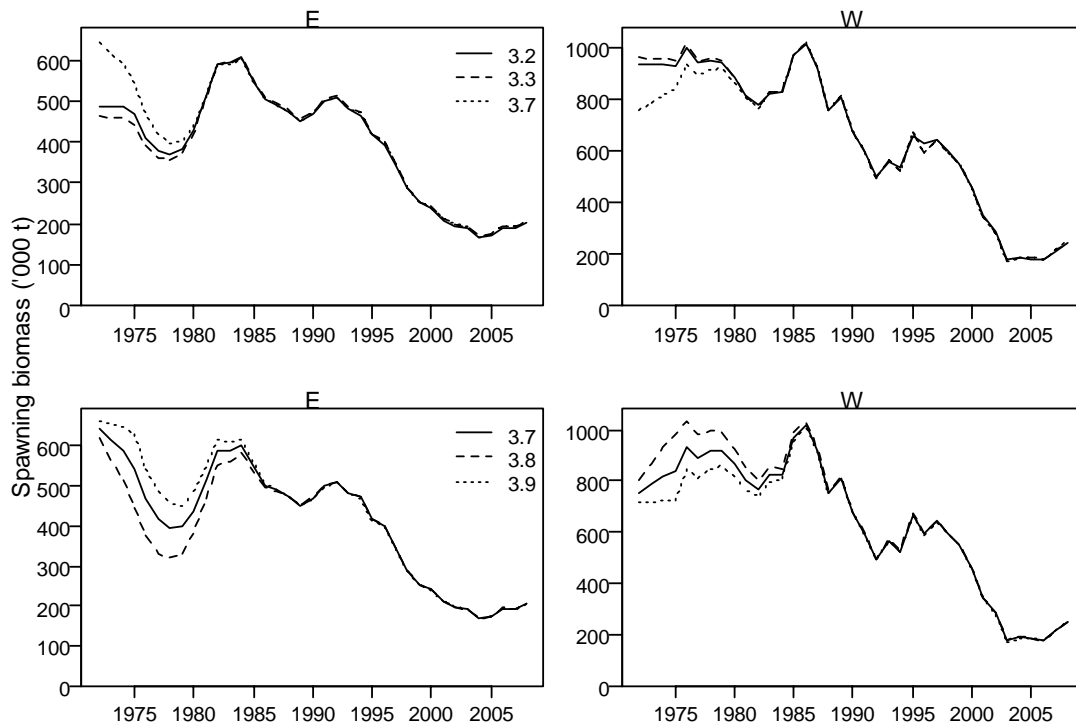
The first problem arose because CASAL did not allow the SSB (spawning-stock biomass) that is used in the stock-recruit relationship to be defined as occurring in multiple areas. Because of this, all models without natal fidelity have defined this SSB as the WC biomass, rather than that in WC and CS combined. After CASAL was modified to allow a multi-area SSB, run 3.3 was constructed to be the same as 3.2, but with the SSB defined for WC + CS. This made very little difference either to the estimated biomass trajectories for the two spawning areas (upper panels, Figure 59) or the fit to the observations (Table 37).

The second problem is that discussed above in Section 2.7:  $B_0$  for each spawning ground has not been well defined in existing models without natal fidelity. In results from these models,  $B_0$  for each stock has been assumed to be the same as  $B_{init}$ , and this assumption led to two others: the prior distribution on  $p_E [= B_{0E}/(B_{0E} + B_{0W})]$  that is used in models without natal fidelity was replaced by the same prior on  $p'_E [= B_{initE}/(B_{initE} + B_{initW})]$ ; and the prior distributions for the annual multipliers for the Whome

migration,  $m_y$ , had no mean (for full details of these assumptions see Francis (2006)). It now seems sensible, as discussed in Section 2.7, to drop the first assumption ( $B_{0E} = B_{initE}$  and  $B_{0W} = B_{initW}$ ) and define  $B_{0E}$  and  $B_{0W}$  to be median spawning biomasses after a 20-year projection without fishing. Having dropped this assumption, we should also drop the two associated assumptions (i.e., have no prior on either  $p_E$  or  $p_E'$ , and reinstate the mean,  $\mu$ , for the priors on the  $m_y$ ). It also seemed a good opportunity to correct a small anomaly in the initialisation of the population. Previously, this has been done with  $m_y = 0$ , but it makes more sense to use  $m_y = \bar{m}$ , where  $\bar{m}$  is the mean of the  $m_y$  values that are resampled in projections (i.e., the values for years 1977–2008). These changes, which were applied in run 3.7 (which was otherwise the same as 3.3), made little difference to the fit to the observations (Table 37), and affected the spawning biomass trajectories only in the years before the first observation in 1988 (upper panels, Figure 59).

For run 3.7 it was necessary to provide a value for  $\mu$ , the mean of the prior on the  $m_y$ . This was set to -0.5, which was chosen as being close to the mean of the  $m_y$  in both the 2007 run 4.7 (-0.46) and the new run 3.3 (-0.47). To illustrate how insensitive the key model results are to this somewhat arbitrary choice, two further runs were constructed that were identical to 3.7, except for the value of  $\mu$ , which was set to -1.0 for run 3.8, and 0.0 for run 3.9. These three models are very similar, both in terms of overall fit to the observations (Table 37) and in the post-1988 biomass trajectories (lower panels, Figure 59).

Further, 20-year projections without fishing produced very similar biomass trajectories from each of these runs (Figure 60), and thus very similar estimates of  $B_0$  by stock (Table 38).



**Figure 59:** Comparison of estimated spawning biomass trajectories from runs 3.2, 3.3, and 3.7 (upper panels) and 3.7–3.9 (lower panels).

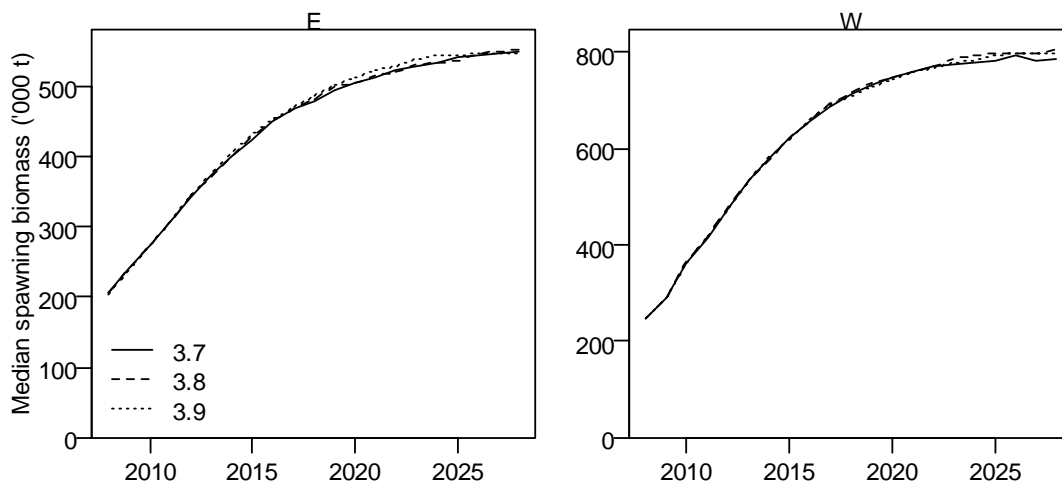
**Table 37:** Comparison of how well each of several model runs fit the observations, relative to run 3.2. For each run, the tabulated value is the total contribution of the observations to the objective function subtracted from that for run 3.2; positive (or negative) values indicate the run fitted the observations better (or worse) than run 3.2.

Run	3.2	3.3	3.7	3.8	3.9
Goodness of fit	0.0	1.4	1.3	1.5	0.9

MPD biomass estimates from run 3.7 are not dissimilar to those from the final runs 2.3 and 2.4 (Table 39).

Note that the projections shown here are point-based. That is, they include uncertainty about future year-class strengths (including the year-to-year variation in the Whome migration that causes the strength of year classes to differ between the two spawning grounds) but not about the current (2008) biomass. These are the projections that will be used to estimate  $B_0$  by stock for MPD runs. Full sample-based projections will be used to estimate these  $B_0$ s for MCMC runs.

I conclude that run 3.7 resolves both of the problems described above and recommend that model runs without natal fidelity be used again in the 2009 assessment.



**Figure 60:** Comparison of median spawning biomass trajectories from point-based 20-year projections with no fishing for runs 3.7-3.9.

**Table 38:** Estimates of  $B_0$  by stock for runs 3.7–3.9 (i.e., the median biomass in 2028 in each of the projections of Figure 60).

Stock	$B_0$ ('000 t)		
	3.7	3.8	3.9
E	548	552	546
W	784	806	798

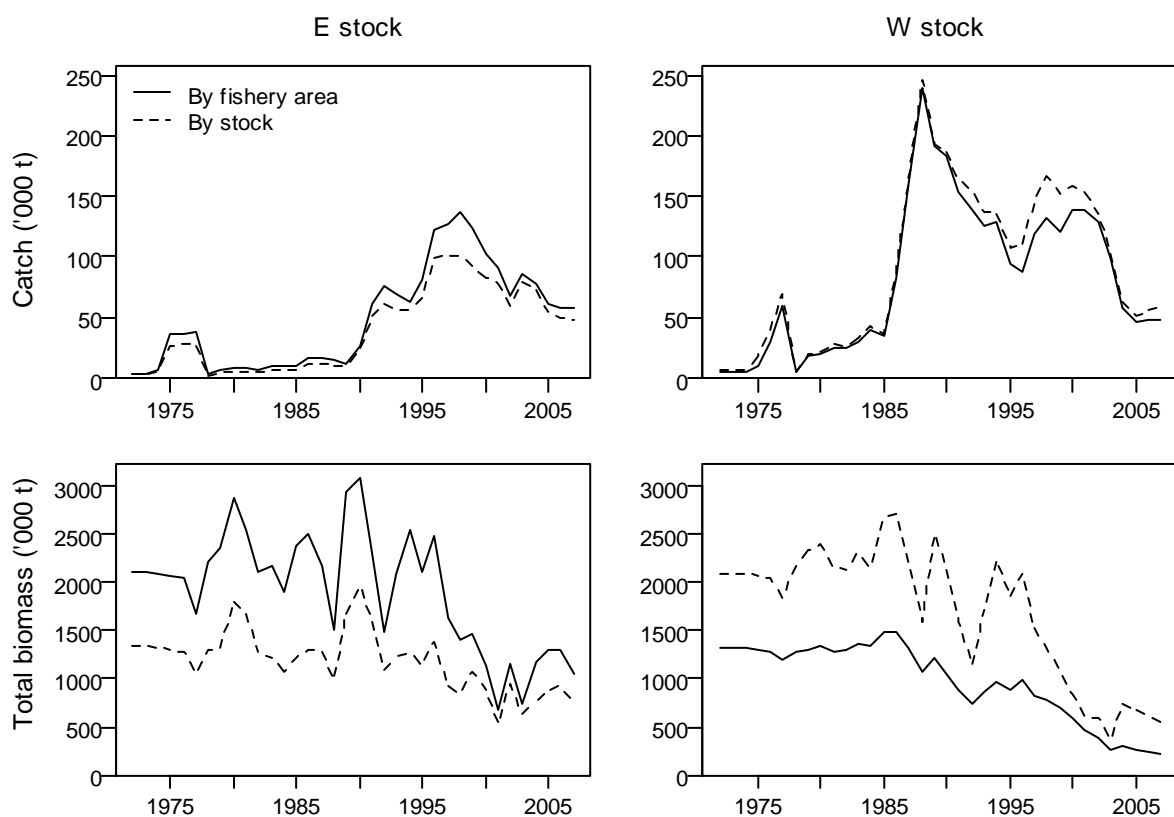
**Table 39:** Comparison of MPD biomass estimates for new run 3.7 with those from final runs 2.3 and 2.4.

Run	$B_0$ ('000 t)		$B_{current}(\%B_0)$	
	E	W	E	W
2.3	510	880	46	25
2.4	637	1062	39	29
3.7	548	784	37	32

## 7.2 Calculations for the RAMlegacy database

This section relates to the hoki time series that were recently provided to the new RAMlegacy database of stock-recruit data sets (<http://www.marinebiodiversity.ca/RAMlegacy/srdb/updated-srdb>). These were based on the point estimates (i.e., MPD) of run 4.4 from the 2007 assessment. There are two issues of interest that arose in the preparation of these time series.

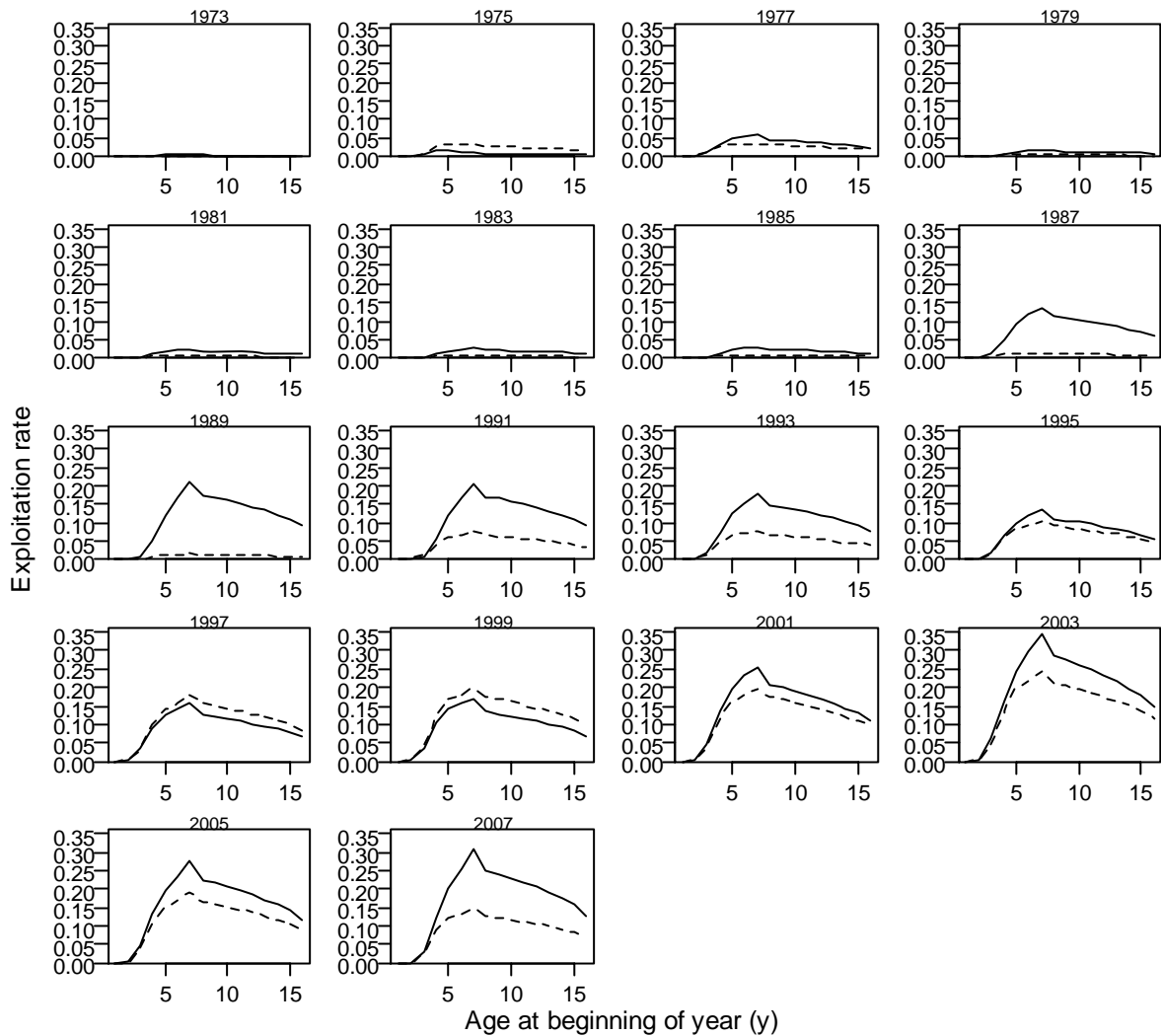
The first issue is relatively minor. It concerns the effect of shifting from a fishery focus to a stock focus. In the assessment it is useful, for management purposes, to sum catches by fishery area into E (catches from areas CR and CS) and W (areas SA and WC). For the RAMlegacy time series the catches were grouped strictly by stock. That is, the catches from CR were separated by stock (as estimated in the stock assessment model). Because hoki catches are dominated by the spawning fisheries the two methods of grouping do not produce very different time series (upper panels, Figure 61). However, there is a big difference when these methods are used to calculate total biomass at mid-spawning season (lower panels, Figure 61).



**Figure 61: Comparison of two methods of calculating catch (upper panels) and total biomass (mid-spawning season, lower panels): by fishery area (i.e., defining E as CS + CR, and W and WC + SA) and by stock (which involves separating CR catch and biomass by stock).**

The second issue, which concerned measures of fishing intensity, was much more substantial. In the stock assessment, fishing intensity is expressed in terms of exploitation rates, which are calculated by fishery, and usually presented only for the two spawning fisheries (e.g., Figure 57 above, and figure 3 in the Plenary Report (Ministry of Fisheries 2008)), because these dominate the catches. These exploitation rates are usually calculated by dividing the catch by the pre-fishery selected (= exploited) biomass. Another way to get them is to calculate, for each age (and sex), the number of fish caught divided by the number that were in the pre-fishery population, and then to find the maximum of these ratios across ages. For the purposes of the RAMlegacy database, there were three objections to our measures of fishing intensity: they apply to fisheries rather than the whole stock; they are maxima across ages, rather than weighted averages; and they are exploitation rates rather than instantaneous rates. The rest of this section describes how I addressed the first two objections.

The calculation of  $E_{sy}$ , the exploitation rate for stock  $s$  in year  $y$  was in two steps. First, age-specific exploitation rates,  $E_{asy}$ , were calculated by summing catches at age ( $C_{asfy}$ ) across fisheries and dividing by initial numbers at age ( $N_{asy}$ ):  $E_{asy} = \sum_f (C_{asfy}) / N_{asy}$ , where  $a$  indexes age and  $f$  the fisheries. The resulting age-specific rates seemed not unreasonable (Figure 62). Then, a biomass-weighted average across age was calculated. Initially, this was done using the equation  $E_{sy} = \sum_a (N_{asy} w_{as} E_{asy}) / \sum_a (N_{asy} w_{as})$ , where  $w_{as}$  is the mean weight of a fish of age  $a$  in stock  $s$ . However, this was unsatisfactory, because it gave extremely high weights to the youngest ages (Figure 63A). This high weighting was a consequence of the very high natural mortality that was estimated for young fish (see Figure 53 for estimates for the 2008 models). So, the last equation was changed to  $E_{sy} = \sum_a (C_{asy} w_{as} E_{asy}) / \sum_a (C_{asy} w_{as})$ , where  $C_{asy} = \sum_f C_{asfy}$ . In other words, the age-specific exploitation rates were weighted by the biomass in the catch, rather than that in the pre-fishery population, which resulted in more sensible weights (Figure 63B). The resultant exploitation rates are smaller than those usually presented in the stock assessment (Figure 64) but show similar trends. [For simplicity, I have ignored sex in the equations in this paragraph. Everywhere there is a subscript for age there should also be one for sex, and the summations in the second equation should be over both age and sex. Also, the age index is for beginning-of-year ages, which are one year less than those in the spawning fishery.]



**Figure 62:** Age-specific exploitation rates by stock for every second year, from run 4.4 (each curve is the average of male and female curves, which do not differ substantially). Because fish age increments during the spawning fishery the ages plotted are one year less than ages recorded in the spawning fishery

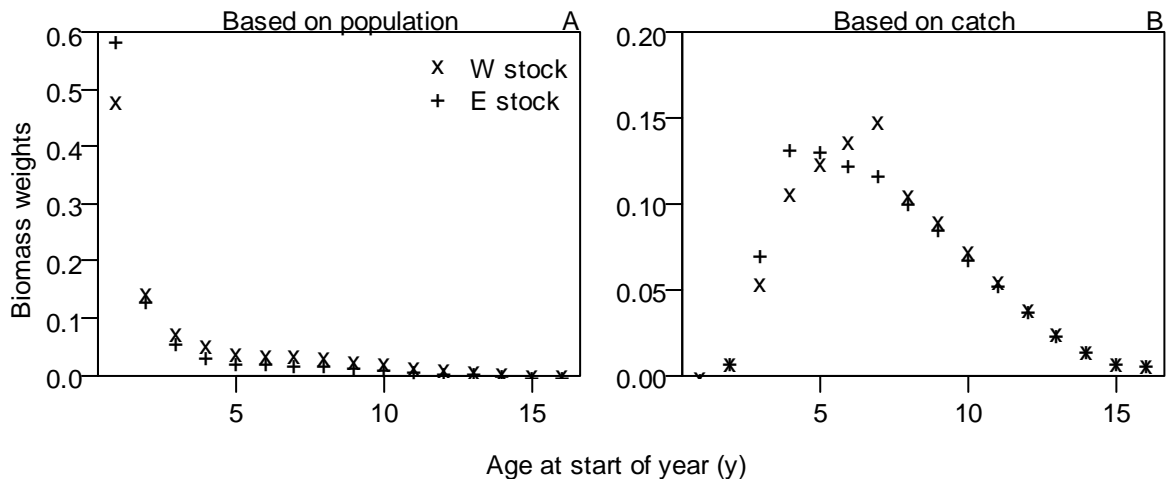


Figure 63: The average biomass weights, by stock, used to calculate annual exploitation rates ( $E_{sy}$ ) from age-specific rates ( $E_{asy}$ ): A, original weights, based on the pre-fishery population biomass; and B, final weights, based on catch biomass. In both panels the plotted weights were calculated by summing over sex and averaging over years.

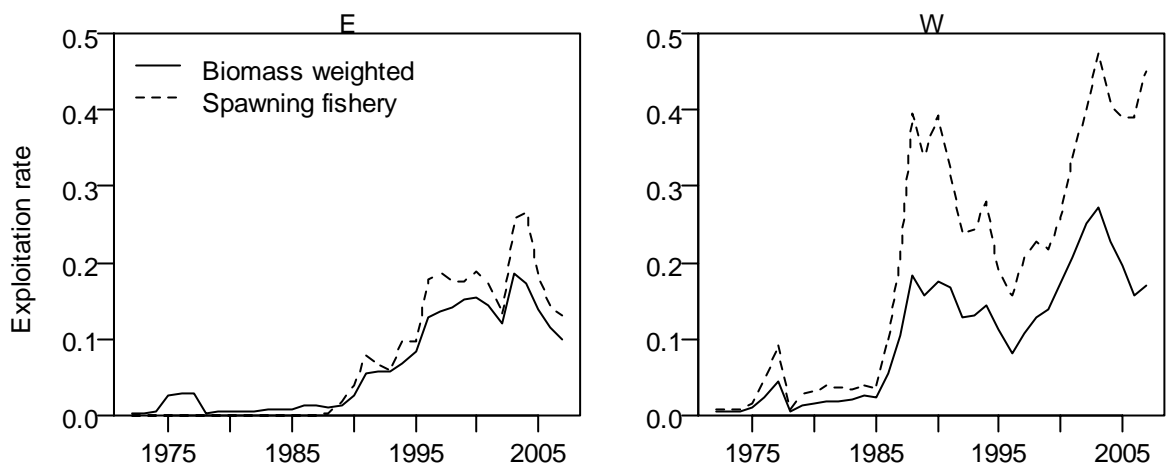


Figure 64: Comparison of biomass-weighted exploitation rates for each stock (solid lines) with those for the spawning fishery (broken lines), as presented in stock assessments.

### 7.3 Reanalysis of 1992 and 1993 proportion spawning data

The histological data associated with the 1992 and 1993 estimates of proportion spawning (Vignaux et al. 1995, Livingston et al. 1997) were rediscovered after the assessment. In this section I describe these data and the effect of reanalysing them using an approach similar to that in Section 2.4.1.

#### 7.3.1 Comparison of current and original data sets

For each fish, the rediscovered data sets contained identifying information (trip [tan9204 or tan9304], station, and fish number), biological information (length, weight, sex [always female], gonad weight, gonad stage), and an integer score labelled gon. The biological information, which was not needed for the reanalysis, agreed exactly with what is currently in the *trawl* database. A comparison of histograms of gon with table 1 of Livingston et al. (1997) produced a clear interpretation of this score

as describing the histology of the gonad (Table 40). Thus, each fish in the sample was classified as a pre-spawner if and only if gon was greater than 8.

**Table 40: Interpretation of the score gon.**

gon score	Interpretation
1	Chromatin nucleolar
6–8	Perinucleolar
9	Yolk vesicle
10–14	Vitellogenic
16	Ripe
>8	Pre-spawner

Two additional variables were added to each of the two histological data sets: stratum (from the *trawl* database) and fish age (from the *age* database). The resulting data sets will be referred to as the “current” histological data. From these, the following two tables were calculated:

$n_{sa}$  – the number of fish in the histological sample of age  $a$  in stratum  $s$ , and

$n'_{sa}$  – the number of pre-spawners in the histological sample of age  $a$  in stratum  $s$

Note that, as in Section 2.4, the ages used in this section are the ages at time of sampling, which is one year less than the ages used by Vignaux et al. (1995) and Livingston et al. (1997) (the latter ages are those that would be achieved during the spawning season following sampling). Age classes for which the histological sample contained no pre-spawners (Table 41) were ignored because their interpretation is straightforward, so there is no scope for reanalysis.

**Table 41: Age classes which were ignored in the reanalysis (because the histological sample included no pre-spawners) and their associated histological sample sizes.**

1992 survey		1993 survey		1998 survey	
Age	Sample size	Age	Sample size	Age	Sample size
1	3	1	0	1	11
2	0	2	9	2	12
		3	7		

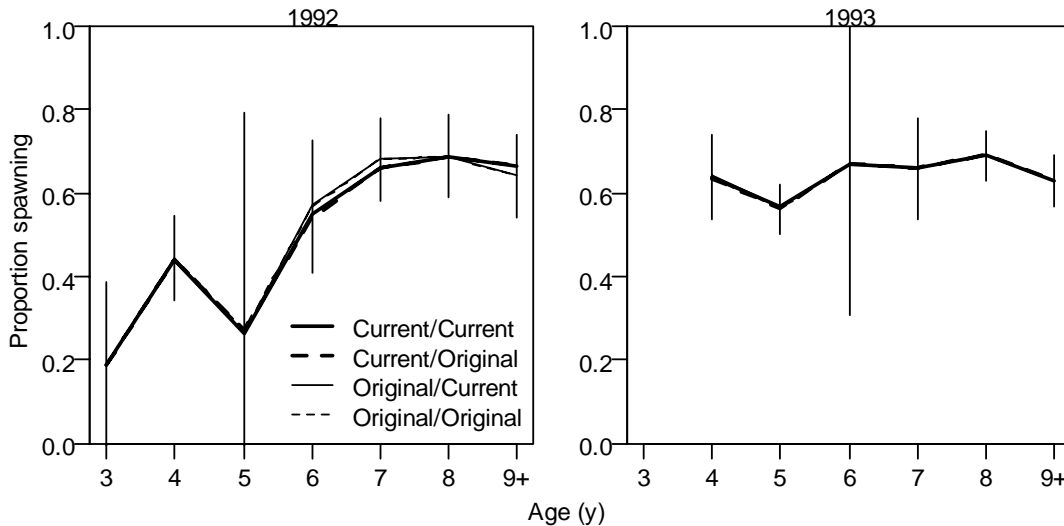
A comparison of the tables  $n_{sa}$  and  $n'_{sa}$  with similar tables calculated from the “original” histological data, and presented by Vignaux et al. (1995), showed that the two versions of the histological data are very similar, but not identical (see tables in Appendix 2). The slight differences between them involve both age and spawning status.

To calculate proportion spawning we need, as well as  $n_{sa}$  and  $n'_{sa}$ , one more table:

$N_{sa}$  – the estimated total number of fish age  $a$  in stratum  $s$ .

which is calculated from the estimated age frequency by stratum (AF) for each survey (tan9204 in 1992 and tan9304 in 1993). Again, we have two versions of this table: the “original” one (tables 2 and 3 of Vignaux et al. 1995), and the “current” one, as calculated using NIWA’s catch-at-age software (Bull & Dunn 2002). Differences between the two versions are mostly slight (see figures in Appendix 2).

To evaluate the effect of the above data differences, proportion spawning was calculated using the original method (Vignaux et al. 1995, Livingston et al. 1997) and all four combinations of current and original histological and AF data. The differences in the AF data had virtually no effect, and the effects from the histological differences were much smaller than the uncertainty in the original estimates (Figure 65). All further calculations in this section use only the current data sets.



**Figure 65: Estimates of proportion spawning for 1992 (left panel) and 1993 (right panel) using all four combinations of current and original histological and AF data (e.g., the label “Current/Original” means current histological data and original AF data). Vertical bars are approximate 95% confidence intervals for the original estimates (+/- 2 standard errors from table 4 of Livingston et al. (1997)).**

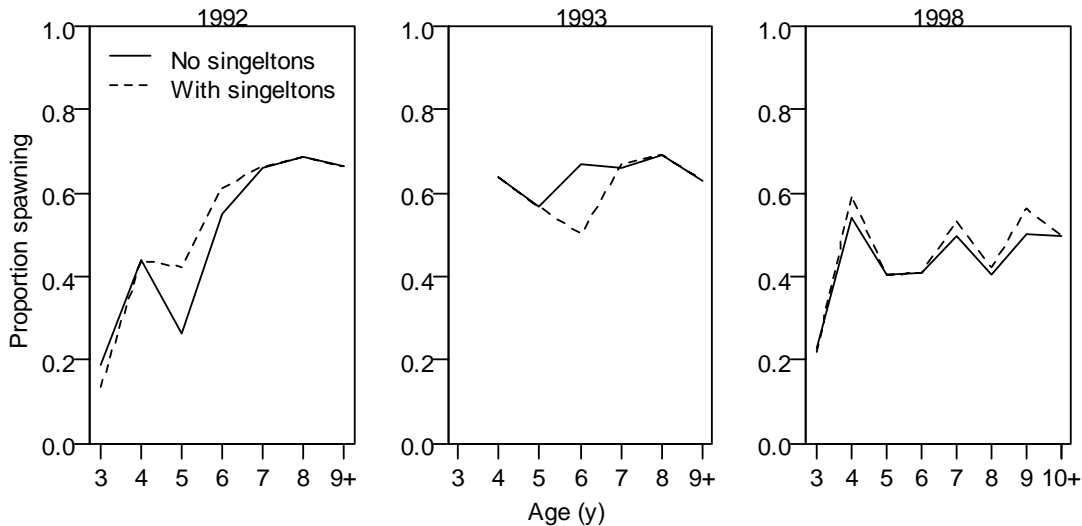
### 7.3.2 The treatment of singleton samples

The above calculations revealed an unsuspected, but significant, difference between the methods originally used in analysing the 1992 and 1993 data and those for the 1998 data. This involved the treatment of singleton samples: combinations of age and stratum with only one histological sample (i.e.,  $n_{sa} = 1$ ). Singletons were ignored for 1992 and 1993 (Vignaux et al. 1995, Livingston et al. 1997), but included for 1998 (Livingston & Bull 2000). Proportion spawning was always calculated as  $\sum_s (N_{sa} p_{sa}) / \sum_s N_{sa}$ , where  $p_{sa} = n'_{sa} / n_{sa}$ . However, for each age, those strata with singleton samples were omitted from both summations in the calculations for 1992 and 1993, but included for 1998.

The decision to include or exclude singletons has a big effect for age 5 in 1992 and age 6 in 1993, but only a relatively small effect for all ages in 1998 (Figure 66). Note that the two ages that are most affected both relate to the weak 1986 year class, for which sample sizes were small in both years. An examination of the data for this year class (Table 42) shows that it is not straightforward to decide whether to include singletons. For age 5 in 1992 it seems best to include the singletons (strata 4, 7, 10, 11, and 13) because they contribute almost half of the histological sample (5 of 11) and, without them, the included strata contain only 23% of the age class. However, it is of some concern that the combined weight given to the five fish in these singleton samples would be double that given to the six fish in the other samples. This problem of over-weighting is more severe for age 6 in 1993. Here, including singletons would have the unfortunate effect of giving more weight to the one fish in stratum 9 than the 7 fish in strata 12 and 8 combined.

The new methods of calculating proportion spawning (from Section 2.4.1), which are applied to the 1992 and 1993 data in the next section, use the singleton samples but avoid the problem of over-weighting them.





**Figure 66:** Effect, on estimates of proportion spawning, of including or excluding singletons (combinations of age and stratum for which there was only one histological sample). For 1992 and 1993 the estimates plotted with a solid line ('No singletons') are the same as those labelled "Current/Current" in Figure 65; for 1998, those plotted with a broken line ('With singletons') are those labelled "Corrected" in Section 2.4.

**Table 42:** Data used to calculate proportion spawning for the 1986 year class, which was aged 5 at the time of the 1992 survey and 6 for the 1993 survey.

Age 5 fish in 1992				Age 6 fish in 1993			
Stratum	Sample size	Number of pre-spawners	Percentage of age class in stratum	Stratum	Sample size	Number of pre-spawners	Percentage of age class in stratum
$s$	$n_{sa}$	$n'_{sa}$	$N_{sa}^{-1}$	$s$	$n_{sa}$	$n'_{sa}$	$N_{sa}^{-1}$
6	2	2	6	12	4	3	8
9	2	0	12	8	3	1	12
15	2	0	5	4	2	2	7
4	1	1	13	10	2	1	5
7	1	1	4	11	2	2	8
10	1	1	6	13	2	2	8
11	1	0	10	14	2	0	5
13	1	0	13	6	1	1	5
All others	0	0	30	9	1	0	23
Total	11	5	100	All others	0	0	19
				Total	19	12	100

<sup>1</sup>For ease of interpretation,  $N_{sa}$  is here converted to percentages

### 7.3.3 Reanalysing the current data

Results from a series of models, analogous to those developed for the 1998 data (see Section 2.4.1), fitted to the 1992 and 1993 data are shown in Table 43. As with the 1998 data, a significantly better fit was obtained (for both 1992 and 1993) when the strata were grouped geographically into six areas (i.e., the AICs for the 6areas models were lower than those for the original models). An even better fit resulted when the six areas were grouped into three groups (i.e., best3 was better than 6areas). However, the 6areas models seem preferable to the best3 models because the latter grouped non-contiguous areas (e.g., for both 1992 and 1993, SWC is grouped with Puys). The 2depths model, which grouped strata bathymetrically, was not as good as the 6areas model in either year, and was better than the original model only for the 1992 data (there were only two depths, compared to three with the 1998 data, because the earlier surveys did not use the 800–1000 m strata). The best estimates of proportion spawning (from the 6areas models) were quite different from the original estimates

(upper panels, Figure 67) but not very different from those from the best3 and 2depths models (lower panels, Figure 67).

**Table 43: Comparison of methods of estimating, by age, the proportion of females that will spawn in the following winter using the current data from 1992 and 1993 (cf Table 8, which shows similar results for the 1998 data). The best method is that with the lowest AIC.**

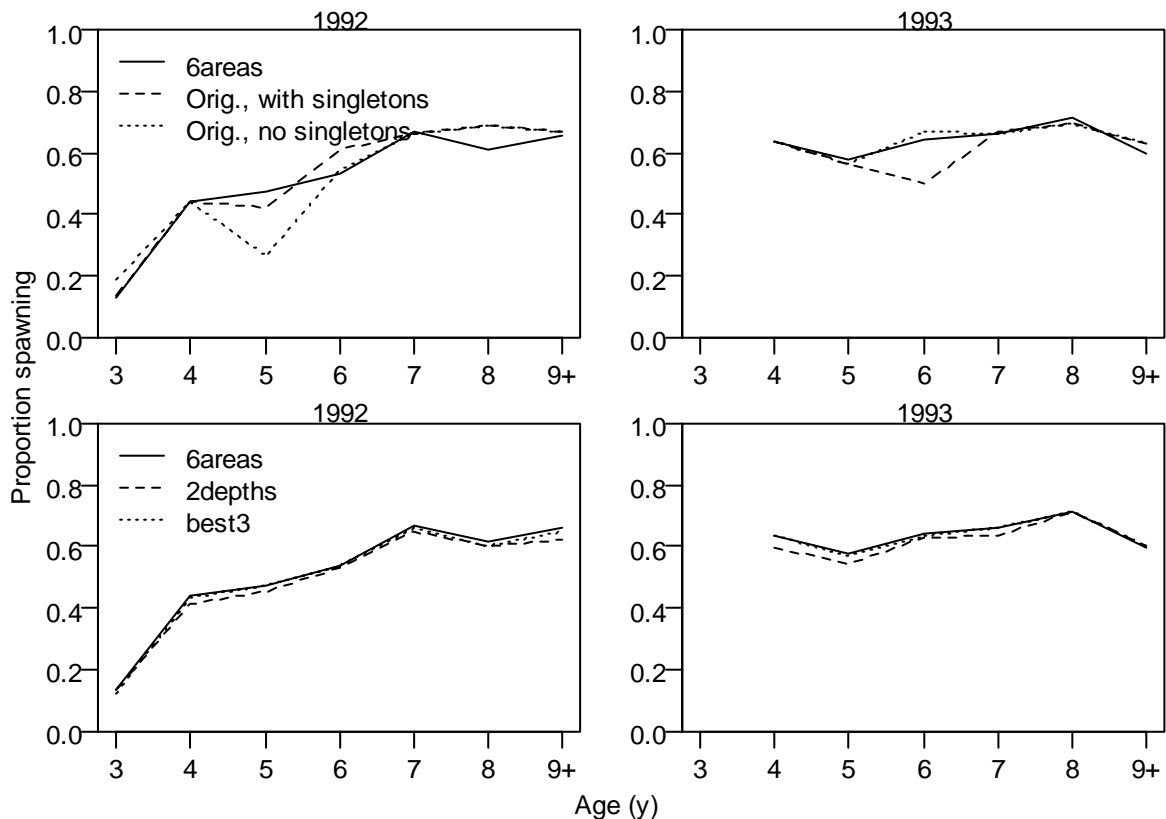
1992 data

Method	Description	Parameters estimated	AIC	Groups of areas
Original <sup>1</sup>	As in Livingston et al. (1997) <sup>1</sup>	83	748.1	
6areas	Strata grouped geographically into 6 areas	17	705.5	
2depths	Strata grouped bathymetrically into 2 depth ranges	9	711.2	
best2	6 geographical areas grouped into 2 groups	9	700.7	NWC.Puk.SSS Puys.SEC.SWC
best3	6 geographical areas grouped into 3 groups	11	698.6	NWC.Puk.SSS Puys.SWC SEC
best4	6 geographical areas grouped into 4 groups	13	699.0	NWC.Puk Puys.SWC SEC SSS
best5	6 geographical areas grouped into 5 groups	15	702.3	NWC Puk Puys.SWC SEC SSS

1993 data

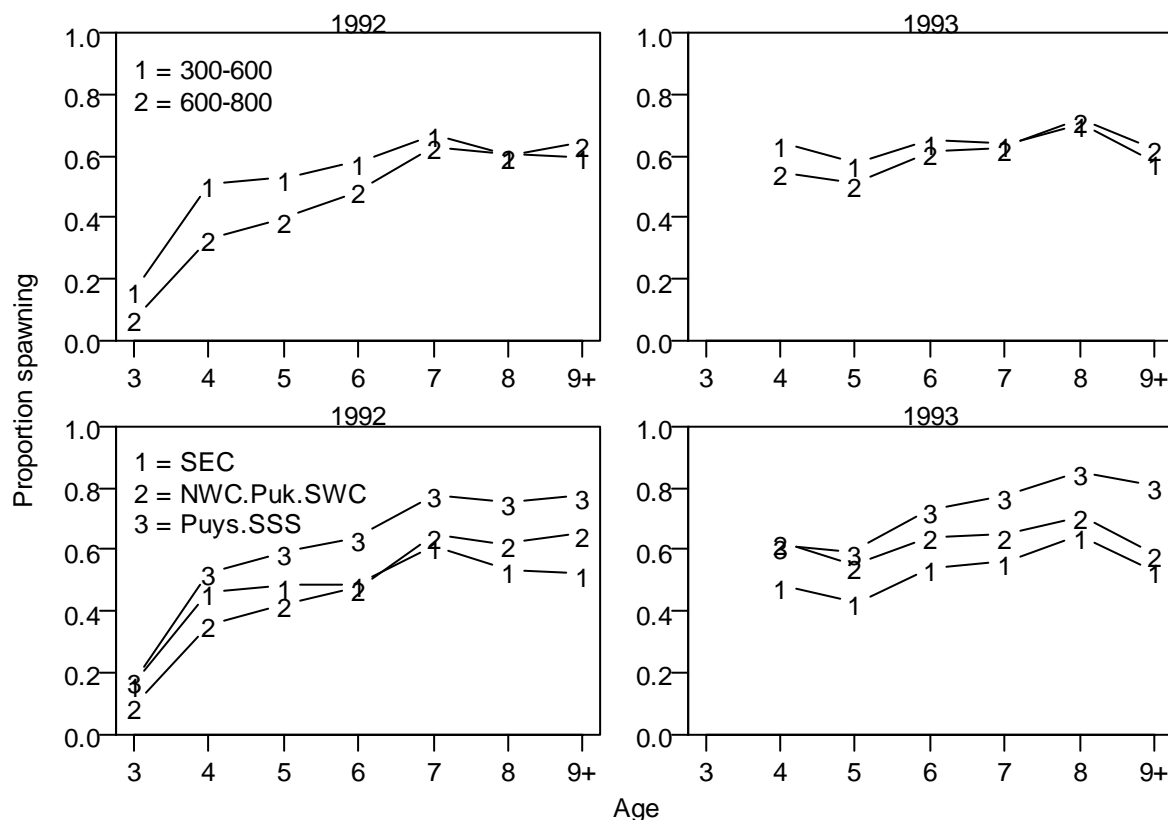
Method	Description	Parameters estimated	AIC	Groups of areas
Original <sup>1</sup>	As in Livingston et al. (1997) <sup>1</sup>	82	1466.1	
6areas	Strata grouped geographically into 6 areas	16	1458.9	
2depths	Strata grouped bathymetrically into 2 depth ranges	8	1491.1	
best2	6 geographical areas grouped into 2 groups	8	1460.5	NWC.Puk.SSS Puys.SEC.SWC
best3	6 geographical areas grouped into 3 groups	10	1449.8	NWC.Puk Puys.SEC.SWC SSS
best4	6 geographical areas grouped into 4 groups	12	1452.0	NWC Puk Puys.SEC.SWC SSS
best5	6 geographical areas grouped into 5 groups	14	1455.2	NWC Puk Puys SEC.SWC SSS

<sup>1</sup> Except that singleton samples were included to make comparisons with the other models valid



**Figure 67: Comparison of the best estimates of proportion spawning (from the 6areas model) for 1992 (left panels) and 1993 (right panels) with those from other models. In the upper panels the comparison is with two variants of the original model (with and without singletons); in the lower panels it is with the 2depths and best3 models.**

Estimates from the 1998 data suggested that proportion spawning decreases from the southeast to the northwest, and from shallow to deep strata (see Figure 15). Those from the 1992 and 1993 data supported the depth trend (upper panels, Figure 68), but not the geographical trend (lower panels, Figure 68).



**Figure 68: Investigation of bathymetric (upper panels) and geographical (lower panels) trends in proportion spawning from the 1992 and 1993 data. Estimates in the upper panels are from the 2depths models of Table 43; those in the lower panels are from 3-area models for 1992 and 1993 that used the same groups of areas as the best3 model for the 1998 data (see Table 8).**

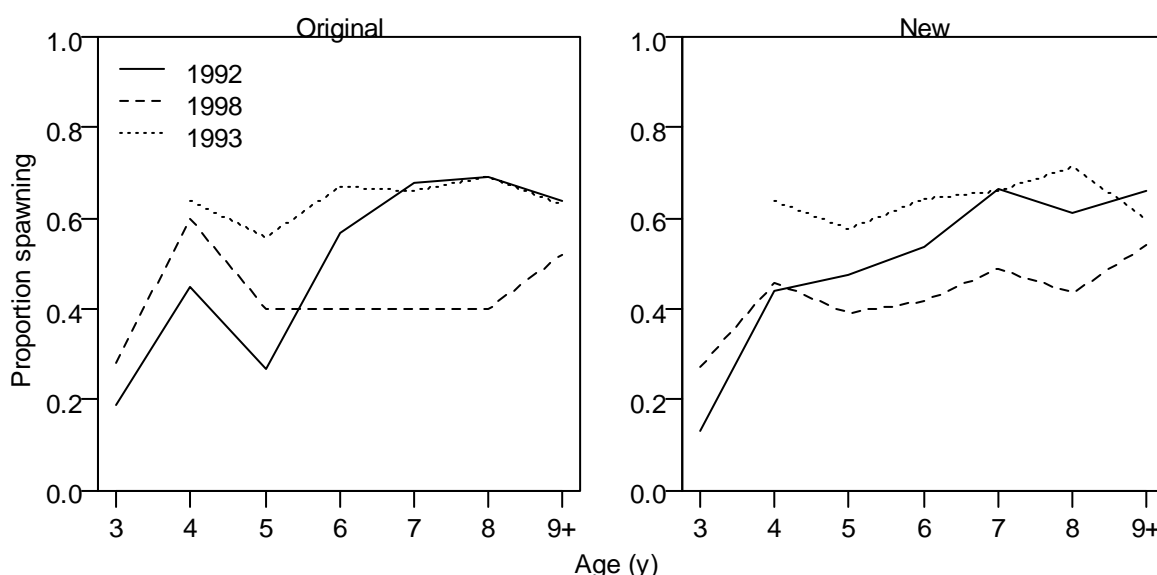
### 7.3.4 Further conclusions concerning the proportion spawning data

The analyses of Section 2.4 allowed us to address the first and last of the four points of concern listed at the beginning of the section (see Section 2.4.4). Having now reanalysed the 1992 and 1993 data we can address the other two points.

The second point of concern related to the decision to exclude the 1993 estimates of proportion spawning from the stock assessment. The logic of this exclusion was as follows. Because the 1993 survey was so late in the year, some females may have already left the survey area on their way to the spawning ground. Thus, estimates of proportion spawning from the 1993 survey could be too low, and including them in the stock assessment could negatively bias the assessment model's estimate of the spawning migration ogive,  $W_{spmg}$ . However, this argument is inconsistent with the fact that the estimates of proportion spawning for 1993 are almost always higher than those for 1992 and 1998, both for the original and new estimates (Figure 69). Thus, the assessment estimates of  $W_{spmg}$  are likely to be higher if the 1993 data are included, and not lower, as argued. I conclude that the 1993 proportion spawning estimates should be used in the 2009 assessment.

The third point of concern was that the methods used in analysing the 1998 data seemed to be slightly different from those for the earlier data. This has been shown to be true, although the methodological

differences related to the treatment of singleton samples (see Section 7.3.2) and not to the suspected issue (see section 4.3 of Francis 2007). These differences are removed in the present reanalyses.



**Figure 69: Original (left panel) and new (right panel) estimates of proportion spawning.** The original estimates are those given by Livingston et al. (1997) and Livingston & Bull (2000); the new estimates are from the 6areas models (Table 44).

For consistency, I recommend using the proportion spawning estimates from the 6areas model for all three years (Table 44). This is the preferred model for 1992 and 1993, and the 1998 estimates from this model are very similar to those from the best3 model (see Figure 11D), which were used in the 2008 assessment.

**Table 44: Estimates of proportion spawning recommended for use in the 2009 stock assessment.** These are from the 6areas models of Tables 8 and 43. ‘-’, not estimated.

Year	Age (y)						
	3	4	5	6	7	8	9+
1992	0.13	0.44	0.48	0.54	0.67	0.61	0.66
1993	-	0.64	0.58	0.65	0.66	0.71	0.60
1998	0.27	0.46	0.39	0.42	0.49	0.44	0.54

The recommended proportions in Table 44 contain no value for age 3 in 1993 for the same reason they contain no values for ages 1 and 2 in any year: small sample sizes (see Table 41). However, it might be unwise to withhold the limited information on age 3 fish in 1993 from the stock assessment if there was any indication that this information was inconsistent with that from 1992 and 1998. A tabulation of all the histological data for age 3 (Table 45) showed that there is no such inconsistency. When we restrict attention to the strata containing age-3 fish in 1993, it is not surprising that none of these fish were pre-spawners given the corresponding data from 1992 and 1998.

**Table 45: Comparison of histological data for 3-year old fish in 1993 with that in 1992 and 1998.**

Stratum	1993		1992		1998	
	Histological sample size	Number of spawners	Histological sample size	Number of spawners	Histological sample size	Number of spawners
1	4	0	0	0	4	0
2	1	0	3	0	10	0
4	1	0	1	0	3	1
9	1	0	7	0	21	8
All others	0	0	13	3	36	9
All	7	0	24	3	74	18

## 8. ACKNOWLEDGMENTS

I am grateful to Sira Ballara, Richard O'Driscoll, and Darren Stevens for providing data; and to members of the Hoki Working Group, in particular Rosie Hurst, for many suggestions during and after the assessment process. This work was funded under MFish project HOK2007/01.

## 9. REFERENCES

- Akaike, A. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control AC-19*: 716–723.
- Bagley, N.W.; O'Driscoll, R.L.; Francis, R.I.C.C.; Ballara, S.L. (2009). Trawl survey of middle depth species in the Southland and Sub-Antarctic areas, November-December 2007 (TAN0714). *New Zealand Fisheries Assessment Report 2009/xx*. xx p.
- Bull, B.; Dunn, A. (2002). Catch-at-age user manual v1.06.2002/09/12. NIWA Internal report 114. 23 p. (Unpublished report held in NIWA library, Wellington.)
- Bull, B.; Francis, R.I.C.C.; Dunn, A.; McKenzie, A.; Gilbert, D.J.; Smith, M.H.; Bian, R. (2008). CASAL (C++ algorithmic stock assessment laboratory): CASAL User Manual v2.20-2008/02/14. *NIWA Technical Report 130*. 275 p.
- Cohen, D.M.; Inada, T.; Iwamoto, T.; Scialabba, N. (1990). FAO species catalogue. vol. 10 Gadiform fishes of the world (order Gadiformes) : an annotated and illustrated catalogue of cods, hakes, grenadiers, and other gadiform fishes known to date. Food and Agriculture Organization of the United Nations, Rome.
- Cordue, P.L. (1994). MIAEL estimation of biomass and fishery indicators for the 1994 assessment of hoki stocks. New Zealand Fisheries Assessment Research Document 94/21. 51 p. (Unpublished report held in NIWA library, Wellington.)
- Cordue, P.L. (2001). MIAEL estimation of biomass and fishery indicators for the 2001 assessment of hoki stocks. *New Zealand Fisheries Assessment Report 2001/65*. 59 p.
- Francis, R.I.C.C. (1992). Recommendations concerning the calculation of maximum constant yield (MCY) and current annual yield (CAY). New Zealand Fisheries Assessment Research Document 92/8. 26 p. (Unpublished report held in NIWA library, Wellington.)
- Francis, R.I.C.C. (2003). Analyses supporting the 2002 stock assessment of hoki. *New Zealand Fisheries Assessment Report 2003/5*. 34 p.
- Francis, R.I.C.C. (2004). Assessment of hoki (*Macruronus novaezelandiae*) in 2003. *New Zealand Fisheries Assessment Report 2004/15*. 95 p.
- Francis, R.I.C.C. (2005). Assessment of hoki (*Macruronus novaezelandiae*) in 2004. *New Zealand Fisheries Assessment Report 2005/35*. 97 p.
- Francis, R.I.C.C. (2006). Assessment of hoki (*Macruronus novaezelandiae*) in 2005. *New Zealand Fisheries Assessment Report 2006/3*. 96 p.
- Francis, R.I.C.C. (2007). Assessment of hoki (*Macruronus novaezelandiae*) in 2006. *New Zealand Fisheries Assessment Report 2007/15*. 99 p.
- Francis, R.I.C.C. (2008a). Additional assessment results for hoki in 2008. WG-HOK-2008/19. 6 p. (Unpublished report held by Ministry of Fisheries, Wellington.)
- Francis, R.I.C.C. (2008b). Assessment of hoki (*Macruronus novaezelandiae*) in 2007. *New Zealand Fisheries Assessment Report 2008/4*. 109 p.
- Francis, R.I.C.C.; Haist, V.; Bull, B. (2003). Assessment of hoki (*Macruronus novaezelandiae*) in 2002 using a new model. *New Zealand Fisheries Assessment Report 2003/6*. 69 p.
- Francis, R.I.C.C.; Hurst, R.J.; Renwick, J.A. (2001). An evaluation of catchability assumptions in New Zealand stock assessments. *New Zealand Fisheries Assessment Report 2001/1*. 37 p.
- Grimes, P.J.; O'Driscoll, R.L. (2005). Estimating the proportion of female hoki on the Southern Plateau which spawn each year by microscopic examination of gonad samples. Final research report for MOF research project MDT200301 obj6

- Hicks, A.C.; Cordue, P.L.; Bull, B. (2002). Estimating proportions at age and sex in the commercial catch of hoki (*Macruronus novaezelandiae*) using length frequency data. *New Zealand Fisheries Assessment Report 2002/43*. 51 p.
- Horn, P.L.; Sullivan, K.J. (1996). Validated aging methodology using otoliths, and growth parameters for hoki (*Macruronus novaezelandiae*) in New Zealand waters. *New Zealand Journal of Marine and Freshwater Research* 30(2): 161–174.
- Livingston, M.E.; Bull, B. (2000). The proportion of female hoki developing to spawn on the Southern Plateau, April 1998. *New Zealand Fisheries Assessment Report 2000/13*. 20 p.
- Livingston, M.E.; Stevens, D.W. (2002). Review of trawl survey data inputs to hoki stock assessment 2002. *New Zealand Fisheries Assessment Report 2002/48*. 69 p.
- Livingston, M.E.; Vignaux, M.; Schofield, K.A. (1997). Estimating the annual proportion of nonspawning adults in New Zealand hoki, *Macruronus novaezelandiae*. *Fishery Bulletin* 95: 99–113.
- Ministry\_of\_Fisheries (2008). Report from the Fisheries Assessment Plenary, May 2008: stock assessments and yield estimates. 990 p. Ministry of Fisheries, Wellington, New Zealand
- Myers, R.A.; Barrowman, N.J.; Hilborn, R.; Kehler, D.G. (2002). Inferring Bayesian priors with limited direct data: applications to risk analysis. *North American Journal of Fisheries Management* 22: 351–364.
- Myers, R.A.; Bowen, K.G.; Barrowman, N.J. (1999). Maximum reproductive rate of fish at low population sizes. *Canadian Journal of Fisheries and Aquatic Sciences* 56: 2404–2419.
- Myers, R.A.; Bridson, J.; Barrowman, N.J. (1995). Summary of worldwide stock and recruitment data. *Canadian Technical Report of Fisheries and Aquatic Sciences* 2024. 327 p.
- O'Driscoll, R.L. (2002). Review of acoustic data inputs for the 2002 hoki stock assessment. *New Zealand Fisheries Assessment Report 2002/36*. 66 p.
- O'Driscoll, R.L.; Hurst, R.J.; Livingston, M.E.; Cordue, P.L.; Starr, P. (2002). Report of hoki working group technical meeting 8 March 2002. WG-HOK-2002/27 (Unpublished report held by Ministry of Fisheries, Wellington.)
- Pikitch, E.K.; Hilborn, R.; McAllister, M.K.; Punt, A.E. (1993). Stock assessment and decision analysis for the western stock of hoki in 1992. New Zealand Fisheries Assessment Research Document 93/10. 47 p. (Unpublished report held in NIWA library, Wellington.)
- Punt, A.E.; McAllister, M.K.; Pikitch, E.K.; Hilborn, R. (1994). Stock assessment and decision analysis for hoki (*Macruronus novaezelandiae*) for 1994. New Zealand Fisheries Assessment Research Document 94/13. 47 p. (Unpublished report held in NIWA library, Wellington.)
- Schofield, K.A.; Livingston, M.E. (1998). Ovarian development and the potential fecundity of western stock hoki (*Macruronus novaezelandiae*). *New Zealand Journal of Marine and Freshwater Research* 32(1): 147–157.
- Smith, M.H. (2004). Fitting priors for natural mortality and proportion of virgin hoki biomass in eastern stock. WG-HOK-2004/14. 7 p. (Unpublished report held by Ministry of Fisheries, Wellington.)
- Sullivan, K.J.; Coombs, R.F. (1989). Hoki stock assessment 1989. New Zealand Fisheries Assessment Research Document 89/4. 26 p. (Unpublished report held in NIWA library, Wellington.)
- Sullivan, K.J.; Cordue, P.L. (1990). Hoki stock assessment 1990. New Zealand Fisheries Assessment Research Document 90/5. 30 p. (Unpublished report held in NIWA library, Wellington.)
- Sullivan, K.J.; Cordue, P.L. (1992). Stock assessment of hoki for the 1992–93 fishing year. New Zealand Fisheries Assessment Research Document 1992/12. 43 p. (Unpublished report held in NIWA library, Wellington.)
- Vignaux, M.; Livingston, M.E.; Schofield, K.A. (1995). Spawning fraction and maturity ogive estimates of female hoki, western stock, 1992 and 1993. New Zealand Fisheries Assessment Research Document 95/1. 32 p. (Unpublished report held in NIWA library, Wellington.)

## Appendix 1: Files defining the final runs

Each of the final model runs is completely defined, in the context provided by the CASAL manual (Bull et al. 2008), by two input files — population.csl and estimation.csl — and, for run 2.3, a user.prior\_penalty.cpp file. These files are presented in this appendix, which may be obtained, as a pdf, from the Science Officer at MFish (science.officer@fish.govt.nz).

## Appendix 2: Comparison of current and original data sets for proportion spawning

The tables and figures in this appendix compare two versions of the data used to calculate proportion spawning for 1992 and 1993: the “current” data (as described in Section 7.3) and the “original” data, as tabulated in Vignaux et al. (1995).

**Table A2.1: Comparison of histological sample sizes, by stratum and age, of current and original versions of the histological data for A, 1992, and B, 1993. Where the two versions agree a single number is given; where they differ they are presented as ‘current/original’. Sample sizes for the original data are from tables 4 and 5 of Vignaux et al. (1995).**

A, 1992 data

Stratum	Age (y)							Total
	3	4	5	6	7	8	9+	
1	0	0	0	1/0	0	0	0	1/0
2	3	4	0	1	1	0	1	10
3	0	7	0	0	1	0	1	9
4	1	11	1	7	11	4	6	41
5	2	2	0	3	6	5	2	20
6	0	7	2	0	7	3	5	24
7	1	7	1	3	5	2	4	23
8	3	15	0	4	10	11	12	55
9	7	15	2	0	13	4	4	45
10	1	16	1	6	16/17	8	12	60/61
11	0	5/6	1	6	11	6	13/12	42
12	0	4	0	5	15/16	17	14/16	55/58
13	3	10	1	1	7	6	7	35
14	3	19	0	4	22	13	9	70
15	0	8	2	4	12	4	14/15	44/45
All	24	130/131	11	45/44	137/139	83	104/106	534/538

B, 1993 data

Stratum	Age (y)							Total
	4	5	6	7	8	9+		
1	8	11	0	3	0	1	23	
2	3	13	0	2	4	7	29	
3	7	13	0	1	9	5	35	
4	12/13	34	2	5	24	25	102/103	
5	7	18	0	3	8	6	42	
6	12	17	1	5	13	14	62	
7	4	25	0	7	12	8	56	
8	11	49	3	5	14	15	97	
9	19	30	1	8	34	23	115	
10	11	33	2	3	21	27	97	
11	0	19	2	11	15	40	87	
12	4	14	4	15	27	43	107	
13	3	7	2	9	30	21	72	
14	11	33	2	15	27	36	124	
15	3	30	0	9	18	12	72	
All	115/116	346	19	101	256	283	1120/1121	

**Table A2.2: Comparison of numbers of pre-spawners, by stratum and age, in current and original versions of the histological data for A, 1992, and B, 1993. Where the two versions agree a single number is given; where they differ they are presented as 'current/original'. '-' = unknown (this occurs only in the original data set where the sample size, for the given age and stratum, was 1). Numbers of pre-spawners for the original data were calculated from tables 4, 5, 8, and 9 of Vignaux et al. (1995).**

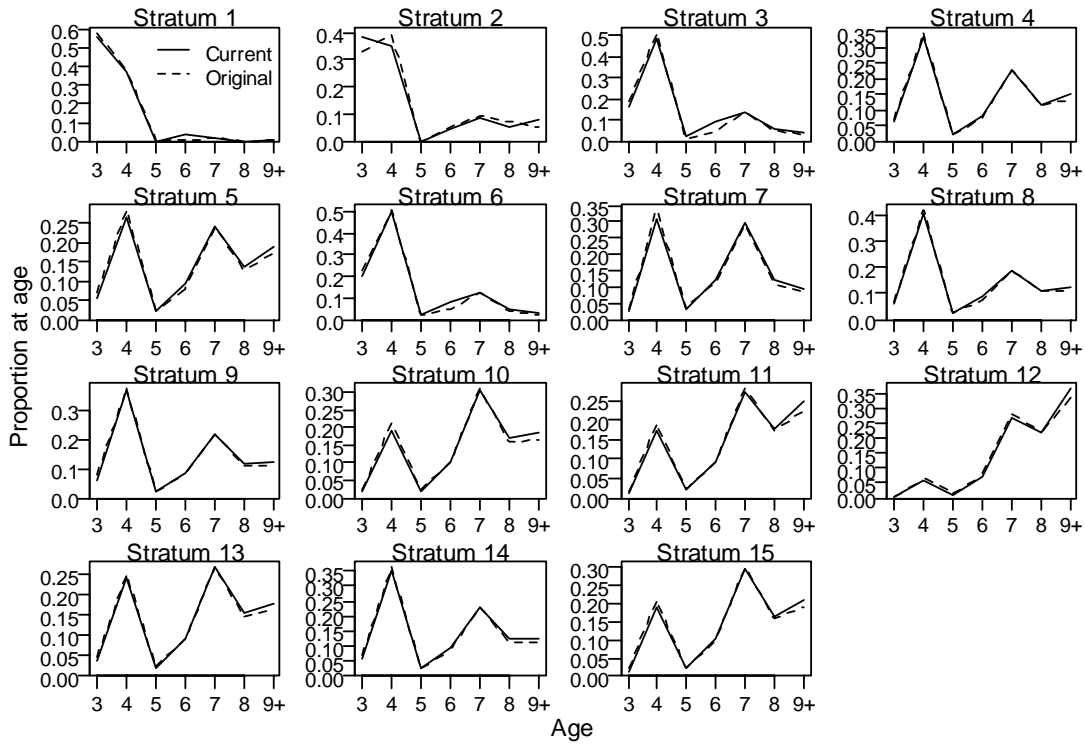
A, 1992 data

Stratum	Age (y)						
	3	4	5	6	7	8	9+
1	0	0	0	0	0	0	0
2	0	2	0	0/-	1/-	0	0/-
3	0	4	0	0	1/-	0	1/-
4	0/-	7	1/-	3	10	3	5
5	1	1	0	3	4	4	1
6	0	4	2	0	3	1	1
7	0/-	2	1/-	0	1	0	2
8	1	2	0	3	8	7	10
9	0	8	0	0	10	3	3
10	0/-	4/3	1/-	2	6/7	4	7/8
11	0	1/2	0/-	4/5	7/8	6	9/7
12	0	1	0	3	13/15	10	10
13	0	6	0/-	1/-	4	6	4
14	1	10	0	3	13	5	5
15	0	2	0	1	8	1	7

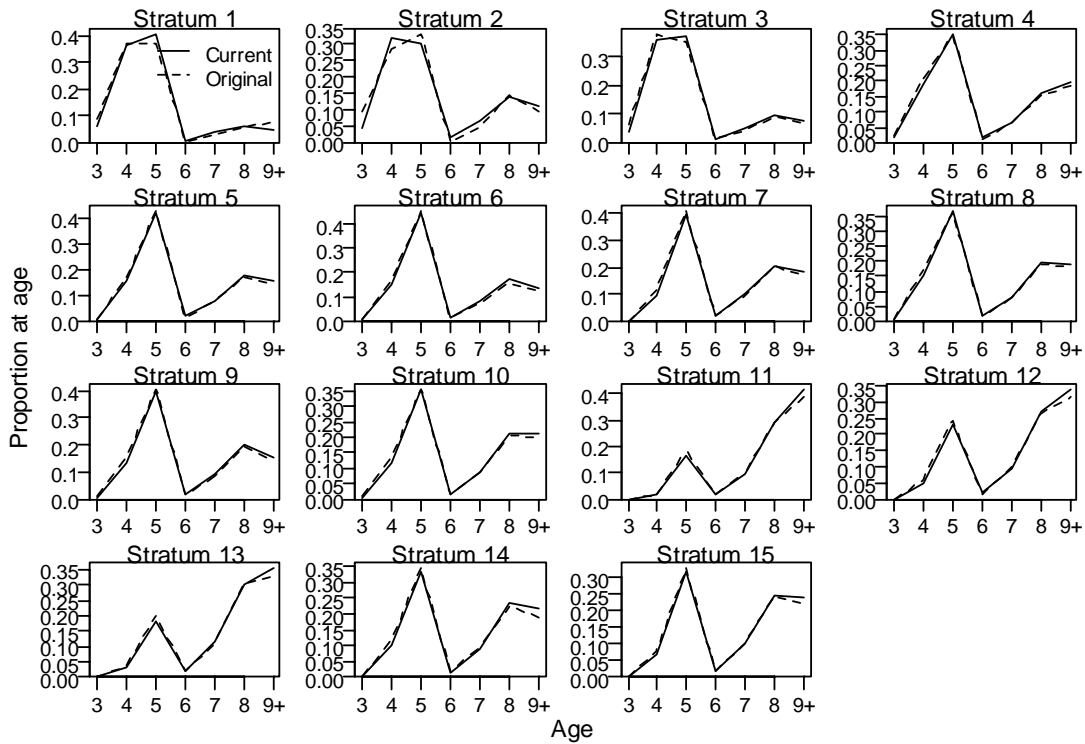
B, 1993 data

Stratum	Age (y)						
	4	5	6	7	8	9+	
1	4	3	0	1	0	1/NA	
2	1	5	0	2	2	5	
3	4	10	0	1/NA	8	5	
4	8/9	26	2	4	21	20	
5	4	15	0	2	4	4	
6	11	16	1/NA	5	13	7	
7	1	14	0	1	9	5	
8	8	25	1	4	9	7	
9	14	14	0/NA	5	17	20	
10	5	11	1	2	12	16	
11	0	8	2	7	13	22	
12	1	10	3	9	21	27	
13	1	5	2	5	22	11	
14	4	17	0	10	17	12	
15	3	6	0	6	14	9	





**Figure A2.1: Comparison of current and original versions of age-frequency data used in calculating proportions spawning for 1992.**



**Figure A2.2: Comparison of current and original versions of age-frequency data used in calculating proportions spawning for 1993.**