New Zealand Fisheries Assessment Research Document 93/15

MIAEL estimation of biomass and fishery indicators for the 1993 hoki stock assessment

P.L. Cordue

MAF Fisheries Greta Point
P.O. Box 297
Wellington

November 1993

MAF Fisheries, N.Z. Ministry of Agriculture and Fisheries

MIAEL estimation of biomass and fishery indicators for the 1993 hoki stock assessment

P.L. Cordue

1. Executive summary

The method used by MAF Fisheries to estimate virgin biomass and fishery indicators for the western hoki stock is described in detail. The estimates obtained are presented and briefly discussed, but for a full discussion of the implications of the results the reader should see Sullivan & Cordue (1993).

A new estimation method is used in a stock reduction setting: estimators are constructed which have minimum integrated average expected loss. Associated with each estimator is a measure of its reliability, termed an information index.

Two future catch policies are considered: the status quo policy (a 200 000 t TACC), and an increased TACC policy (a 250 000 t TACC). The estimates of stock and fishery risk are shown to be unreliable for both policies; the estimators have very low information indices. However, a consideration of the maximum risks associated with each policy gives clear conclusions: the status quo policy has very low associated risks over the next 4 – 5 years and the increased TACC policy entails low risks up to and including 1995.

2. Introduction

This paper describes in detail the derivation of MAF Fisheries' biomass and fishery indicator estimates, for the western hoki stock, given in Sullivan & Cordue (1993) and Annala (1993). There has been a major change, from previous years, in the estimation procedure used by MAF Fisheries.

There are three main components in the estimation procedure: the population model; the unknown parameters; and the estimators. The most significant change is that to MIAEL estimation. The acronym "MIAEL" stands for "minimum integrated average expected loss" and is a property of the estimators used in this year's assessment. The population model has been altered slightly from previous years to incorporate a fishing mortality outside the spawning season, and a number of (fixed) parameters have been updated. The unknown parameter set of the model has been extended from virgin biomass (B_0) to include a number of year class strengths.

The parameters have been estimated using relative biomass indices from a number of sources (catch per unit effort analysis, acoustic surveys and trawl surveys), and quasi age frequency data obtained from modal analysis of length frequencies of the west coast commercial catch. The estimation procedure operates in a stock reduction setting using, to a certain extent, the goodness of fit between observed indices and predicted model outputs to estimate the unknown parameters. The degree of influence the fit between observed and predicted values has on the estimates depends on the number and precision of the available indices. MIAEL estimators are constructed to have the "best" average performance given the quantity and quality of the data. An important advantage of MIAEL estimation is that a measure of the information content of the data, for each fishery indicator, can be determined. This measure, termed an information index, varies from indicator to indicator depending on the amount of improvement in the MIAEL estimator when the observed indices are used in the estimation procedure.

3. Methods

3.1 Population model

The structure of the population model remains very similar to that used in the previous two years (*see* Cordue *et al.* 1992). The previous model has a year structure which includes an instantaneous spawning season (during which the whole annual catch is removed) and a full year over which natural mortality is applied. Significant reference points in the model year are as follows.

"Beginning of year"

> when the fish spawned the previous year enter the population as one-year old fish, and natural mortality is first applied for the year.

"Beginning of spawning season"

> when the natural mortality has been applied but no catch removed.

"Mid-season"

> the middle of the spawning season, when in addition to deaths through natural mortality, half the catch has also been removed.

"End of season"

> the end of the spawning season when the whole of the catch has been removed.

The only structural change is to allow the annual catch to be split into pre-spawning season and spawning season components. That is, before the spawning season the pre-spawning season catch and deaths due to natural mortality are removed from the population, and the spawning season catch is removed as a "point mass". The pre-spawning season catch is removed assuming a constant instantaneous total mortality $Z = F + M$, where F is the fishing mortality and M is the natural mortality. F is determined by solving the Baranov catch equation for the given pre-spawning season catch.

A number of model parameters have been revised. The growth parameters were adjusted by Sullivan & Cordue (1993) who used MULTIFAN to analyse length frequency data. The recruitment variability was rounded from 0.66 to 0.65 to better reflect the approximate nature of the estimate. The proportion spawning was lowered from 100% to 70% due to new information presented by Livingston & Schofield (1993) and a new maturity ogive was derived (Sullivan & Cordue 1993). The full set of biological parameters is given in Table 1.

3.2 Estimation procedure

A generalisation of the approach of Cordue (1993) has been used. Virgin biomass, year class strengths, and fishery indicators are estimated. In addition, an information index is calculated for each estimate, which measures the reliability with which the estimate is determined within its known range.

The method requires an age structured population model with all parameters known except

B_0, the virgin biomass

and

$S = \{ S_i \mid i \in I \}$, a number of year class strengths.

To estimate $\theta = (B_0, S)$ and functions of θ there are available a known catch history, relative biomass indices, and "age frequency" data (derived by modal analysis from length frequency data). For each indicator that will be estimated there is a corresponding indicator function

$$f : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}$$

where n is the number of unknown year class strengths.

There are four steps in estimating each indicator $f(\theta)$ and its information index.

(i) Determine the range of f and define estimation loss over that range (i.e., the loss associated with choosing a particular estimate t when the true value is $f(\theta)$. e.g., squared error $(f(\theta) - t)^2$).

(ii) Search a general class of estimators (of $f(\theta)$) to determine the estimator which has minimum integrated average expected loss (MIAEL) under the definition of estimation loss in (i). Note, when estimation loss is squared error and B_0 is the only unknown parameter then the MIAEL estimator is that with minimum integrated mean squared error.

(iii) Estimate the information content of the data (*indices*) re $f(\theta)$ with

Info(*indices*, f) = 1 - IAEL(G,f, *indices*)/IAEL(H,f,ϕ)

where

G is the MIAEL estimator of $f(\theta)$ determined in (ii)

H is the MIAEL estimator of $f(\theta)$ when *indices* = ϕ = the empty set

IAEL() determines the integrated average expected loss

Note that when f maps onto an interval $[a, b]$ and estimation loss is squared error, then $H = (a + b)/2$.

(iv) Calculate the MIAEL estimate of $f(\theta)$, using the optimum estimator determined in (ii).

The definition of estimation loss in (i) will determine the nature of the MIAEL estimator. For risk functions I have used squared error, while for B_0 and the like I have used proportional squared error because the proximity in percentage terms is more important than that in absolute terms. I have defined proportional squared error, when estimating $f(\theta)$ with $t(X)$, as follows:

$$\left[\frac{f(\theta) - t(X)}{f(\theta)}\right]^2 = \left[1 - \frac{t(X)}{f(\theta)}\right]^2$$

The class of estimators to search is important, as it must be general enough to include an estimator which is nearly optimum, in a global sense, when estimators are ranked by integrated (average) expected loss. I have used two classes of estimators: C for functions which have 0 in their range and C′ for functions that do not.

$$C = \{E(Y) \mid Y \sim LN(f_{min}, f_{max}, a + f(B_0^*,S^*),b),\ a \in \mathbb{R},\ b \in \mathbb{R}^+\}$$

and

$$C' = \{E(Y) \mid Y \sim LN(f_{min}, f_{max}, a * f(B_0^*,S^*), b * f(B_0^*,S^*)),\ a \in \mathbb{R}^+,\ b \in \mathbb{R}^+\}$$

where B_0^* and S^* are least squares point estimators of B_0 and S, f maps onto $[f_{min}, f_{max}]$ and *LN(low,high,mn,sd)* is a lumped normal distribution (weight below *low* lumped into a point mass at *low*, and weight above *high* lumped into a point mass at *high*) formed from a normal distribution with mean equal to *mn* and standard deviation equal to *sd*.

The parameter a is included to allow for any bias in the point estimator, while the spread b puts into the distribution has a stabilising effect (the larger b is, the more robust the estimator is to fluctuations in the point estimate). These estimators are general enough to mimic the behaviour of a wide range of estimators, in particular all the estimators considered in the 1991 hoki stock assessment (Cordue *et al.* 1992, Cordue 1993).

The information index measures the performance of the MIAEL estimator with the given data, relative to the MIAEL estimator when there is no data (other than the catch history). The index varies between 0 and 1.

The determination of the MIAEL estimator uses only the number and precision of the available biomass and/or age frequency indices (and lots of computer time). The observed indices are not used until the final step when the MIAEL estimate is calculated.

Some important theoretical results are derived in Appendix 1. In particular, the best constant estimators (termed *best k* estimators, which are real numbers within the given range) of $f(\theta)$ are derived for squared error and proportional squared error when no biomass or age frequency indices are available. If f maps onto an interval $[a,b]$ then they are:

squared error: $(a + b)/2$

proportional squared error: $a * b * log(b/a) / (b - a)$

3.3 Model input data

Data from a range of sources are used in the model to determine the least squares estimates of virgin biomass and year class strength. The quantity and quality of the data is used in simulations to determine the MIAEL estimators.

3.3.1 Catch history

The catch history for the western stock (Table 2) is from Sullivan & Cordue (1993). The figures for 1993 are an estimate agreed upon by the Hoki Working Group. The increasing proportion of pre-spawning season catch prompted this year's structural change to the model.

3.3.2 Biomass indices

Biomass indices have been used from three sources: acoustics surveys, catch per unit effort analysis, and trawl surveys (*see* Sullivan & Cordue 1993 for details). The acoustic surveys supply relative indices of mid-season recruited biomass and estimates of the average proportion of hoki outside the 25 nautical mile closed area (Table 3). The latter estimates are used as an availability index to adjust the catch per unit effort indices (the raw index is divided by the availability index). The adjusted CPUE indices are used as a relative index of mid-season recruited biomass. The trawl survey index is used as a relative index of end of season recruited biomass.

The Hoki Working Group assigned each index a median "coefficient of variation" to reflect the relative reliability of the indices as measures of abundance: in calculating the least squares estimates, each index was given a relative weighting inversely proportional to the square of its coefficient of variation. The assignments (Table 4) were the same as previous years with the acoustics index being assigned 0.25 (with an adjustment between years for the number of snapshots as in Cordue *et al.* 1992), the CPUE index 0.35, and the new trawl index 0.25.

3.3.3 Age frequency data

Age frequency data obtained from length frequencies using the software package "MIX" (Sullivan & Cordue 1993) are given in Table 5. When using the data in MIAEL estimation, the associated random variables are assumed to have a multinomial distribution with the sample sizes as listed in Table 5. The sample sizes were calculated so as to ensure that the variances of the random variables are compatible with the standard errors of the MIX estimates (Appendix 2).

The observed age frequencies were used as relative indices of 4 and 5 year old recruited cohort strength and were assigned a median *c.v.* of 0.40 (adjusted between years by the number of tows sampled, in an manner analogous to the *c.v.s* of the acoustic indices re the number of snapshots). The indices were used in a relative rather than absolute sense, so that only the relative size of annual proportions could affect the parameter estimates. (This meant that no assumptions about catchability at age had to be made - if used as absolute indices, a catchability ogive would have had to have been assumed.)

4. Results and discussion

The method described in section 3.2 was applied in a modified form. Preliminary estimates of year class strength were determined and for the remainder of the analysis these were assumed known. This approach was adopted because it was very unlikely that the full methods would yield results very different from the modified methods. There were two considerations. First, the information indices of the preliminary estimates were very high, indicating that the year class strengths were accurately estimated. Second, it would have been difficult to apply either full method (MAF Fisheries or the FIB consultants) in the available time, given the available computer resources.

4.1 Model ogive

The maturity and recruitment ogive used in the model (*see* Table 1) was derived by Sullivan & Cordue (1993). The procedure used estimates of the number of 4 and 5 year old males and females in the 1991 and 1992 west coast South Island (WCSI) catches (Sullivan & Cordue 1993) and an estimate of the maximum proportion of any age class which spawn (Livingston & Schofield 1993). The method is not ideal as the estimates were assumed to have no observation error and only a subset of the catch data was used.

It is better to use all the data and fit the ogive by least squares within the model. It is assumed that 5 year old males are fully recruited, and that the male and female ogives are identical except that females recruit 1 year later than males. Therefore, the ogives are fully determined if the proportions of recruited 3 and 4 year males are established. The least squares method was used assuming different values for the maximum spawning proportion (60%, 70%, and 80%) and in each case the estimates, to one decimal place, were 0.1 and 0.4, which are very close to the figures used.

4.2 Year class strength

A least squares stepwise procedure was used to determine which years the year class strength would be estimated for. The number of free years was increased in turn from 1 to 4, including as a free parameter at each stage the single year which gave the greatest reduction in the total sum of squares (Table 6). After 3 years no useful improvement was seen in the sum of squares (only 2% from 3 to 4 free years). The Hoki Working Group decided that the available data were of a quality that the fitting of only 1 or 2 year class strengths was warranted, so only two year class strengths were estimated.

Preliminary MIAEL estimates of year class strength were 0.45 and 3.55 for 1986 and 1987 respectively. Subsequent minor adjustments to the CPUE index and the age frequency sample sizes gave revised estimates of 0.44 and 3.35 (Table 7) only slightly different from the previous figures. The preliminary figures were used in later analysis (as they were the only ones available at the time) which assumed the year class strengths to be known (as detailed in Table 8). Table 8 also gives a set of year class strengths, for later years, which were qualitatively estimated by the Hoki Working Group from Chatham Rise length frequency data and used in a sensitivity test of estimated risks.

4.3 Virgin and current biomass

The MIAEL estimates of virgin biomass (B_0) and "current" biomass (B_{beg94}) are given in Table 9 with their information indices. These figures, when multiplied by a constant (different for each), provide estimates of MCY and CAY respectively. The resulting estimates are also MIAEL estimates (*see* Appendix 1) and have the same information indices.

The information indices are about 20% which suggests that the estimates are of some use in pinpointing the unknown values within their known bounds. The least squares estimates are well to the left of the *best k* estimates. The MIAEL estimates, by necessity, lie between the two. However, the MIAEL estimates are still determined mainly by the *best k* estimates, which also have some sensitivity to the arbitrary value of B_{max} (3 million t). Longer time series of biomass indices will be needed before the estimates can be used with confidence.

4.4 Fishery indicators

For two fixed target catch policies (Table 10) a number of fishery indicators have been estimated. The fishery indicators considered were:

(1) $P(B_{low} < 0.2B_0)$ where B_{low} = minimum $\{ B_{93}, \ldots ,B_{98} \}$

The probability of the mid-season biomass falling below 20% of the virgin level in any year from 1993 to 1998. Termed the *stock risk.*

(2) $P((C/T)_{low} < 0.8)$ where $(C/T)_{low}$ = minimum $\{ C_{93}/T_{93}, \ldots ,C_{98}/T_{98} \}$

The probability that in any year from 1993 to 1998 the catch will be less than 80% of the target catch. Termed the *fishery risk.*

(3) $P(B_{last} \geq B_{msy})$ where $B_{last} = B_{98}$ and $B_{msy} = 25.9\%B_0$

The probability that the mid-season biomass in 1998 will be at or above the biomass level corresponding to MSY (*see* Sullivan & Cordue 1993).

(4) $E(B_{last}/B_{msy})$

The expected mid-season biomass in 1998 as a proportion of B_{msy}.

(5) $E(B_{last}/B_0)$

The expected mid-season biomass in 1998 as a proportion of virgin biomass.

(6) $E(B_{low}/B_0 \mid B_{low} < 0.2B_0)$

The expected minimum mid-season biomass given that the mid-season biomass, in some year from 1993 to 1998, is less than $20\%B_0$.

(7) $E(B_{beg94}/B_0)$

The expected beginning of year biomass in 1994 as a proportion of B_0.

(8) $E(\Sigma C_i/\Sigma T_i)$ where $\Sigma C_i = C_{93} + ... + C_{98}$ and $\Sigma T_i = T_{93} + ... + T_{98}$

The expected total catch as a proportion of the total target catch.

(9) $E(Av(\Delta C_i))$ where $Av(\Delta C_i) = (|C_{94}-C_{93}| + ... + |C_{98}-C_{97}|)/5$

The expected annual variation in catches.

MIAEL estimates of stock and fishery risk for the two policies are given in Table 11, together with some figures relating to their derivation (the least squares estimates and the bias and *s.d.* parameters of their lumped normal distributions). The MIAEL estimates and information indices for the first five fishery indicators are given in Table 12. For the remaining indicators only the range of possible values was calculated (Table 13): for many indicators the range is very restricted and estimation within the range is pointless.

Tables 14 and 15 contain ranges for the annual proportion of filletable fish within the catch by number and weight respectively. This proportion was estimated within the model as all fish aged 6 or greater plus 60% of 5 year old males and 85% of 5 year old females. This split approximates the proportion of fish 74 cm in length or longer.

The information indices (*see* Table 12) for the first three indicators (which include stock and fishery risk) are all less than or equal to 5% and the estimates differ little from the *best k* estimates which are at the midpoint of the possible ranges. The very low information indices mean that the range of each indicator (*see* Table 13) gives the only guide to their values.

For the status quo policy the maximum stock risk is 16%, the maximum fishery risk 3%, and there is at worst a 70% chance that B_{98} will be above B_{msy}. The policy appears to be a low risk option, particularly as the maximum annual stock risk does not reach double figures until 1998 (Table 16). The increased TACC policy has much higher maximum stock and fishery risks (about 50%); however, the maximum annual risks still do not reach double figures until 1996 and become large only in 1997.

4.5 Sensitivity analysis

Many assumptions have been made to derive biomass and risk estimates. In this section the consequences of alternative parameter values and modelling decisions are presented.

4.5.1 Maximum possible level of virgin biomass

To determine MIAEL estimators of virgin biomass (and some fishery indicators) an upper bound for virgin biomass, B_{max}, must be specified. The lower bound for virgin biomass, B_{min}, is a product of the model (being the smallest virgin biomass which enables the historical catches to be taken), but B_{max} is a subjective estimate. A bounded range is required to enable integrated average expected loss to be calculated for estimators, so that they can be ranked by it and the MIAEL estimator determined.

The sensitivity of *best k* estimates of B_0 to the value of B_{max} is shown in Table 17. The present value of B_{max} is 3 million t, and a 33% reduction in B_{max} results in a 16% reduction in the *best k* estimate of B_0, while a 33% increase gives a 12% increase in the estimate. These sensitivities are for MIAEL estimates with a 0% information index. The information index of the current B_0 estimate is 18% and so it will be less sensitive to changes in B_{max}. Indeed, the larger the information index the less the sensitivity and when the information index gives the user confidence in the estimate the sensitivity will be negligible.

4.5.2 Number of estimated year class strengths

A relatively arbitrary decision was made to estimate two year class strengths rather than one or three. The effect of this decision on the range of the first five fishery indicators for the increased TACC policy is shown in Table 18. Had only one year class strength been estimated, a somewhat more optimistic picture would have resulted with stock and fishery risk both being reduced by about 0.1 (from approximately 0.5 to 0.4). If three years had been estimated the view would have been slightly less optimistic with a small increase in stock risk and a similar fishery risk. Note, that if no year class strengths had been estimated the increased TACC policy would appear extremely risky.

4.5.3 Qualitative estimates of year class strength

Length frequency data from Chatham Rise trawl surveys are available for a number of years. The Hoki Working Group used these data and their own intuition to reach a qualitative consensus on the year class strengths for 1988 to 1991 inclusive (*see* Table 8). Quantitative estimates are problematic as the fish have not yet entered the commercial catches in large numbers, and different vessels carried out the trawl surveys (which is important because the young age classes, in particular, could have had different vulnerability to different trawl gear).

The ranges of the nine indicator functions, for the increased TACC policy, using the qualitative years class strengths, are given in Table 13. The annual expected proportion of fillets are given by number (Table 14) and weight (Table 15). The maximum stock and fishery risks increase (from approximately 0.5 to 0.6), but the range of the expected beginning of year biomass in 1994 also increases. This means that if the qualitative estimates were correct, then the short term outlook (1993 to 1995) is better than estimated by the baseline, but the medium term outlook is worse. This is because of the qualitatively strong 1988 year class and the weak 1989 and 1990 year classes.

4.5.4 Average year class strength

In the main assessment, two year class strengths were estimated assuming that the remainder were all of average strength (i.e., 1). An alternative assumption is that the average year class strength, over all years within the model, must equal 1. With this constraint on the year class strengths, the three parameters S_{86}, S_{87}, and c (the year class strength of every other year) were estimated by least squares to be 0.4, 3.2, and 0.9 respectively. For the increased TACC policy, this has little impact on the ranges of the first five fishery indicators (Table 19). The alternative estimates show a slightly improved picture, over the baseline, with the greatest improvement being in fishery risk (from 0.5 to 0.4).

4.5.5 Class of estimators

To determine MIAEL estimators a specific class of estimators must be searched. The estimators used in this paper have come from two parameter classes of lumped normal distributions. To test the sensitivity of estimates to the choice of class, another class of estimators has been considered.

Suppose the fishery indicator function is f and let

$$P = \{\, p * f(B_0^*, S^*) + (1-p) * bestk(f_{min}, f_{max}) \mid p \in [0,1] \,\}$$

where f maps onto $[f_{min}, f_{max}]$ and *bestk*( ) returns the best constant estimator (for a given definition of loss) for the given range. This is a one parameter class of estimators which includes the least squares estimator (p = 1) and the *best k* estimator (p = 0). It includes the two MIAEL estimators at the extremes of information content (*best k* when there is none, and least squares when it is complete) and a continuous set of estimators in between. It is therefore a good class of estimators to consider when searching for one which is near optimum in a global sense. I shall denote the best estimator within this class as the *best p* estimator.

In Table 20 the *best p* estimates of B_0 and B_{beg94} are given with their alternative information indices and the values of p used. The actual estimates are marginally larger than those from the lumped normal distribution estimators (*see* Table 9) and the information indices are fractionally lower. The latter is simply a reflection of the lower generality of a one parameter class of estimators, i.e., the *best p* estimator is not quite as good as the best lumped normal distribution estimator.

Table 21 gives the *best p* estimates for the first five fishery indicators for the two target catch policies. The estimates of probability (the first three indicators) are virtually identical to the baseline estimates (*see* Table 12) as are the information indices. This is because the information indices are so low (≤ 5%) that both estimators are hardly different from the *best k* estimator. For the estimates of biomass, the *best p* estimates are somewhat bigger and their information indices fractionally smaller.

The parameter value (*p*) associated with each *best p* estimator can be considered as an alternative information index. It is by definition the amount of weight given to the least squares estimate, and varies between 0 and 1 depending on the usefulness of the point estimator (which in turn depends on the information content of the available indices). For each quantity estimated the values of *p* are substantially higher than the given information index. It is hard to judge which value best summarises the information content of the indices. However, the advantage of the current information index is that its definition is independent of any class of estimators.

4.5.6 Biological parameters

A number of biological parameters used in the model are not well estimated so it is important to examine the effect of alternative values within plausible ranges. This was done for five different parameters: natural mortality; percentage spawning (the maximum for any age class); steepness (a stock recruitment relationship parameter); recruitment variability; and future availability (mean percentage outside the closed area).

Results are presented in Figures 1 and 2 for maximum annual stock risk and maximum annual fishery risk respectively. Changes in natural mortality have the largest effect, followed by changes in the percentage spawning. Lower natural mortality and increased percentage spawning both lead to substantially increased maximum stock and fishery risks from 1996 onwards. Movements in the other direction significantly reduce the risks. The only other notable effect is that of an increase in future availability. If all the west coast spawning stock were available to the fishing fleet, there would be a major reduction in fishery risk but only a small increase in stock risk. (However, this is not to advocate the removal of the closed area, which is in place for other reasons.)

5. Conclusion

Except for the estimated year class strengths, which were assumed known in the later analysis, the information indices of the estimators are low. In particular, the information indices of the stock and fishery risk estimates are less than or equal to 5%. This means that the MIAEL estimators which use the abundance indices perform only fractionally better than those that do not. That is, the data are of very little use in determining where, within their known bounds, the true risks lie.

However, it is clear from a consideration of the maximum annual risks that the status quo policy has very low associated risks over the next 4 to 5 years. It is also clear that the increased TACC policy entails low risks up to and including 1995. From 1996 onwards the risks become increasingly large and may be very substantial if natural mortality has been over-estimated or the percentage spawning has been under-estimated.

The estimation of year class strengths is crucial to these conclusions. The two policies considered are shown to be low risk options over the next 2 years primarily because the large 1987 year class has been incorporated into the model.

The low information indices and the importance of estimating year class strength highlight the need for continued data collection. There is also a need for further improvements to the population model and estimation procedures. In particular, the use of juvenile hoki length frequency data, collected by *Tangaroa* on the Chatham Rise, will be important for accurate forecasting of future recruited biomass. Done correctly, this could enable relative changes in biomass to be predicted some years in advance.

6. Acknowledgments

I thank Chris Francis for his original hoki model program which forms the nucleus of the current software. Also, a general thank you to the many MAF Fisheries staff who have collected data, or derived biomass indices, or estimated biological parameters which are used in the model.

7. References

Annala, J.H. (Comp.) 1993: Report from the Fishery Assessment Plenary, May 1993: stock assessments and yield estimates. 241 p.

Cordue, P.L. 1993: A Minimised Integrated Average Mean Squared Error approach to biomass and risk estimation. N.Z. Fisheries Assessment Research Document 93/1.

Cordue, P.L., McAllister, M.K., Pikitch, E.K., & Sullivan, K.J. 1992: Stock assessment of hoki 1991. N.Z. Fisheries Assessment Research Document 92/10.

Livingston, M.E. & Schofield, K.A. 1993: Seasonal estimates of cohort strength and spawning condition of hoki on the Southern Plateau 1992. Draft N.Z. Fisheries Assessment Research Document.

Sullivan, K.J. & Cordue, P.L. 1993: A review of the 1991-92 hoki fishery and assessment of hoki stocks for 1993. Draft N.Z. Fisheries Assessment Research Document.

Table 1: Biological parameters

	Male	Female
Natural mortality	0.30	0.25
Growth parameters		
L_{inf}	102.8	112.0
K	0.199	0.172
t_0	-1.01	-1.13
Length-weight relationship		
a	0.006	0.006
b	2.85	2.85

Steepness SRR = 0.75
Recruitment variability (rsd) = 0.65
Plus group: 20 years and older
Proportion spawning = 70%

Proportion mature at each age (recruitment is assumed to coincide with maturity)

Age	Male	Female
1	0.0000	0.0000
2	0.0000	0.0000
3	0.0714	0.0000
4	0.4290	0.0714
5	1.0000	0.4290
6	1.0000	1.0000
7+	1.0000	1.0000

Table 2: Catch history (figures for 1993 assumed, '000 t)

	Catch			Catch	
Year	Pre-spawn	Spawn	Year	Pre-spawn	Spawn
1972	0	5	1983	0	30
1973	0	5	1984	0	40
1974	0	10	1985	0	40
1975	0	10	1986	0	82
1976	0	30	1987	0	158
1977	0	60	1988	0	240
1978	0	5	1989	5	192
1979	0	18	1990	8	175
1980	0	20	1991	12	164
1981	0	25	1992	31	113
1982	0	25	1993	40	100

Table 3: Proportion of fish available to the WCSI fleet (mean proportion outside the 25 nautical mile closed area)

Year	Mean % outside
1972	57
.	.
.	.
1987	57
1988	73
1989	44
1990	49
1991	57
1992	63
1993+	57

Table 4: Biomass indices. The acoustics index is adjusted for species mix. The CPUE index is adjusted for mean proportion outside the closed area. The trawl index was obtained using fish >= 55 cm. The *c.v.s* are those used in the model.

Year	Acoustics index	*c.v.*	CPUE index	*c.v.*	Trawl index	*c.v.*
1987	-	-	1.71	0.35	-	-
1988	274	0.28	1.00	0.35	-	-
1989	171	0.22	1.00	0.35	-	-
1990	160	0.21	0.73	0.35	-	-
1991	259	0.25	0.62	0.35	79.5	0.25
1992	218	0.28	0.74	0.35	87.0	0.25
Season timing	mid season		mid season		end of season	

Table 5: Age frequency data (obtained from length mode analysis using MIX). The *c.v.s* were obtained by weighting a median *c.v.* of 0.40 according to the number of tows sampled (in the same manner as the median acoustic *c.v.* of 0.25 is weighted by number of snapshots). The "equivalent *n*" values were obtained by multiplying the median equivalent *n* (determined for each sex from the MIX standard errors of estimated proportion) by the number of tows sampled in each year.

Year	Equivalent *n* male	Equivalent *n* female	*c.v.*	sampled	Proportion at age male 4	male 5	female 4	female 5
1986	36	57	0.79	55	0.068	0.103	0.076	0.030
1987	68	108	0.61	104	0.115	0.188	0.014	0.086
1988	307	489	0.38	470	0.110	0.245	0.026	0.087
1989	242	386	0.40	371	0.000	0.108	0.000	0.080
1990	318	506	0.38	487	0.019	0.000	0.001	0.000
1991	375	597	0.36	574	0.274	0.071	0.073	0.027
1992	245	390	0.40	375	0.058	0.453	0.056	0.324

Table 6: Least squares estimates of virgin biomass (B_0) and year class strength (S) for 0, 1, 2, 3, or 4 free years. The percentage improvement is that in the total sum of squares from the previous step.

No. of free years	Estimated virgin biomass (million t)	Years	Estimated year class strength	Total sum of squares	Percent improvement
0	1.37	-	-	37.05	-
1	1.09	1987	3.78	16.33	56%
2	1.21	1986	0.44		
		1987	3.66	14.46	11%
3	1.92	1985	0.12		
		1986	0.33		
		1987	3.04	9.93	31%
4	1.76	1983	1.28		
		1985	0.12		
		1986	0.35		
		1987	3.20	9.73	2%

Table 7: MIAEL estimates of virgin biomass (B_0), 1986, and 1987 year class strengths (S_{86} and S_{87}) using proportional mean squared error and multiplicative classes of lumped normal distribution estimators. The parameters, a and b, of the multiplicative class C′ are defined in section 3.2.

	Least squares estimate	*best k*	a	b	MIAEL estimate	Information index
B_0	1.21e6	1.70e6	0.6	1.5	1.53e6	14%
S_{86}	0.44	0.53	1.0	0.3	0.44	92%
S_{87}	3.66	0.53	1.0	0.4	3.35	89%

Table 8: Year class strengths (used in the model to multiply the mean recruitment, as predicted by the stock recruitment relationship)

Assessment (proper)

Year	Strength	
1971	1.00	
.	.	
.	.	
1985	1.00	
1986	0.45	estimated
1987	3.55	estimated
1988	1.00	
1989+	unknown	

Assessment (sensitivity test)

Year	Strength	
1971	1.00	
.	.	
.	.	
1985	1.00	
1986	0.45	estimated
1987	3.55	estimated
1988	2.00	qualitative estimate
1989	0.20	qualitative estimate
1990	0.50	qualitative estimate
1991	2.00	qualitative estimate
1992+	unknown	

Table 9: MIAEL estimates of B_0 and B_{beg94} (when age class strengths are assumed known as in Table 8) using proportional mean squared error and multiplicative classes of lumped normal distribution estimators. The parameters, *a* and *b*, of the multiplicative class C' are defined in section 3.2.

	Least squares estimate	*best k*	*a*	*b*	MIAEL estimate	Information index
B_0	1.16e6	1.72e6	0.6	1.4	1.49e6	18%
B_{beg94}	0.98e6	1.66e6	0.1	2.1	1.32e6	19%

Table 10: Management policies considered. Target catches are '000 t.

Status quo policy (TACC unchanged)			Increased TACC policy		
	Catch			Catch	
Year	Pre-spawn	Spawn	Year	Pre-spawn	Spawn
1994	50	90	1994	50	140
1995	50	90	1995	50	140
1996	50	90	1996	50	140
1997	50	90	1997	50	140
1998	50	90	1998	50	140

Table 11: MIAEL estimates of stock and fishery risk, for the status quo policy and increased TACC policy (Table 10), using mean squared error and additive classes of lumped normal distribution estimators. The parameters, *a* and *b*, of the additive class C are defined in section 3.2.

	Status quo policy		Increased TACC policy	
	Stock risk	Fishery risk	Stock risk	Fishery risk
Least squares estimate	0.06	0.00	0.31	0.24
best k	0.08	0.02	0.25	0.25
Maximum risk	0.16	0.03	0.51	0.50
a	0.0	0.0	0.0	0.0
b	0.5	0.2	1.0	1.3
MIAEL estimate	0.08	0.01	0.26	0.25
Information index	2%	0%	5%	2%

Table 12: MIAEL estimates of five performance indicators for the status quo policy and the increased TACC policy (Table 10).

	Status quo policy		Increased TACC policy	
	MIAEL estimate	Information index	MIAEL estimate	Information index
Stock risk	0.08	2%	0.26	5%
Fishery risk	0.01	0%	0.25	2%
$P(B_{last}>=B_{msy})$	0.83	3%	0.50	5%
$E(B_{last}/B_{msy})$	1.89	23%	1.39	20%
$E(B_{last}/B_0)$	0.49	23%	0.36	20%

Table 13: Minimum and maximum values of nine performance indicator functions for the status quo policy and the increased TACC policy (with and without the qualitative estimates of year class strength (YCS)).

	Status quo (unknown YCS)		Increased TACC (unknown YCS)		Increased TACC (qualitative YCS)	
Stock risk	0.00	0.16	0.00	0.51	0.00	0.62
Fishery risk	0.00	0.03	0.00	0.50	0.00	0.61
$P(B_{last}>=B_{msy})$	0.70	1.00	0.30	1.00	0.21	1.00
$E(B_{last}/B_{msy})$	1.36	3.43	0.92	3.22	0.79	3.01
$E(B_{last}/B_0)$	0.35	0.89	0.24	0.83	0.21	0.78
$E(B_{low}/B_0 \mid B_{low}<.2B_0)$	0.15	0.17	0.14	0.18	0.15	0.18
$E(B_{beg94}/B_0)$	0.77	1.36	0.77	1.36	0.83	1.41
$E(\Sigma C_i/\Sigma T_i)$	1.00	1.00	0.94	1.00	0.92	1.00
$E(Av(\Delta C_i))$ ('0 000 t)	0.00	0.04	1.00	1.79	1.00	1.93

Table 14: Minimum and maximum values for the proportion of filletable fish by number for 1993 to 1998 for the status quo policy and the increased TACC policy (with and without the qualitative estimates of year class strength (YCS)).

Year	Status quo (unknown YCS)		Increased TACC (unknown YCS)		Increased TACC (qualitative YCS)	
1993	0.85	0.88	0.85	0.88	0.88	0.90
1994	0.84	0.88	0.84	0.88	0.92	0.94
1995	0.82	0.87	0.81	0.87	0.82	0.88
1996	0.80	0.86	0.77	0.86	0.75	0.85
1997	0.77	0.85	0.74	0.85	0.71	0.84
1998	0.76	0.85	0.72	0.85	0.70	0.84

Table 15: Minimum and maximum values for the proportion of filletable fish by weight for 1993 to 1998 for the status quo policy and the increased TACC policy (with and without the qualitative estimates of year class strength (YCS)).

Year	Status quo (unknown YCS)		Increased TACC (unknown YCS)		Increased TACC (qualitative YCS)	
1993	0.90	0.93	0.90	0.93	0.91	0.93
1994	0.90	0.93	0.90	0.93	0.96	0.97
1995	0.89	0.93	0.89	0.93	0.90	0.94
1996	0.88	0.93	0.86	0.92	0.85	0.92
1997	0.86	0.92	0.83	0.92	0.82	0.92
1998	0.85	0.92	0.81	0.92	0.80	0.91

Table 16: Maximum values of annual stock risk ($P(B_i<0.2B_0)$) and annual fishery risk ($P(C_i<0.8T_i)$) in each year from 1993 to 1998 under the status quo and increased TACC policies.

	Status quo policy		Increased TACC policy	
Year	Stock risk	Fishery risk	Stock risk	Fishery risk
1993	0.00	0.00	0.00	0.00
1994	0.00	0.00	0.00	0.00
1995	0.00	0.00	0.00	0.00
1996	0.01	0.00	0.13	0.12
1997	0.07	0.00	0.34	0.33
1998	0.16	0.03	0.48	0.47

Table 17: The *best k* estimates ('000 000 t) of B_0 for B_{max} values of 2, 3, 4, and 5 million t. A value of 3 million t is assumed in the baseline case.

B_{max}	*best k*
2	1.44
3	1.72
4	1.93
5	2.10

Table 18: Minimum and maximum values of five performance indicator functions for the increased TACC policy with 0, 1, 2 (baseline) or 3 year class strengths estimated.

	None		1 year class		2 year classes		3 year classes	
Stock risk	0.00	0.86	0.00	0.39	0.00	0.51	0.00	0.57
Fishery risk	0.00	0.86	0.00	0.40	0.00	0.50	0.00	0.47
$P(B_{last}>=B_{msy})$	0.12	1.00	0.39	1.00	0.30	1.00	0.29	1.00
$E(B_{last}/B_{msy})$	0.63	2.68	1.03	3.40	0.92	3.22	0.88	2.89
$E(B_{last}/B_0)$	0.16	0.69	0.27	0.88	0.24	0.83	0.23	0.75

Table 19: Minimum and maximum values of five performance indicator functions for the increased TACC policy with average year class strength for non-estimated values (the baseline case) and average year class strength over all years.

	Average over all years		Baseline	
Stock risk	0.00	0.49	0.00	0.51
Fishery risk	0.00	0.39	0.00	0.50
$P(B_{last}>=B_{msy})$	0.34	1.00	0.30	1.00
$E(B_{last}/B_{msy})$	0.95	2.99	0.92	3.22
$E(B_{last}/B_0)$	0.25	0.78	0.24	0.83

Table 20: *Best p* estimates of B_0 and B_{beg94}.

	p	Information index	Estimate
B_0	0.27	16%	1.57e6
B_{beg94}	0.25	17%	1.49e6

Table 21: *Best p* estimates (Est.) of five performance indicators for the status quo policy and the increased TACC policy (*see* Table 9).

	Status quo policy			Increased TACC policy		
	p	Info. index	Est.	p	Info. index	Est.
Stock risk	0.13	2%	0.08	0.20	5%	0.26
Fishery risk	0.07	0%	0.02	0.16	2%	0.25
$P(B_{last}>=B_{msy})$	0.16	3%	0.83	0.16	5%	0.51
$E(B_{last}/B_{msy})$	0.34	22%	1.93	0.30	19%	1.47
$E(B_{last}/B_0)$	0.34	22%	0.50	0.30	19%	0.38

Appendix 1: Invariance properties of MIAEL estimators and derivation of best constant MIAEL estimators

In this appendix I give two sets of theoretical results for MIAEL estimators with some illustrative derivations.

Firstly, I shall derive the *best k* estimators for the two loss functions used in this assessment. These are MIAEL estimators of $f(\theta)$, of the form $k \in [a = f_{min}, b = f_{max}]$, and are:

squared error: $(a + b)/2$

proportional
squared error: $a * b * log(\, b/a \,) \, / \, (b - a)$

Secondly, I shall prove some invariance properties of MIAEL estimators which are as follows.

For a given class of estimators D, if $M(X)$ is the MIAEL estimator of $f(\theta)$ in D, EL_θ is the expected loss function, and $c \in \mathbf{R} - \{ 0 \}$ then,

Result 1: If $\forall \; t(X) \in D \;\; EL_\theta(c + f(\theta), c + t(X)) = EL_\theta(f(\theta), t(X))$ then

$c + M(X)$ is the MIAEL estimator of $c + f(\theta)$ in
$D + c = \{ \, t(X) + c \mid t(X) \in D \, \}$

and

$Info(indices, c + f) = Info(indices, f)$

Result 2: If $\forall \; t(X) \in D \;\; EL_\theta(c * f(\theta), c * t(X)) = EL_\theta(f(\theta), t(X))$ then

$c * M(X)$ is the MIAEL estimator of $c * f(\theta)$ in
$D * c = \{ \, t(X) * c \mid t(X) \in D \, \}$

and

$Info(indices, c * f) = Info(indices, f)$

Corollary 1: For squared error, result 1 is true for $c \in \mathbf{R}$, and result 2 is true for $c \in \{-1, 1\}$.

Corollary 2: For proportional squared error, result 2 is true for $c \in \mathbf{R} - \{ 0 \}$.

Derivation of *best k* estimator for proportional squared error

I am estimating $f(\theta)$ with k where f maps onto $[a, b]$. The expected loss function is proportional squared error:

$$EL_\theta(f(\theta), k) = E_\theta(1 - k/f(\theta))^2 = (1 - k/f(\theta))^2$$

I shall derive, for the given k, an expression for the integrated average expected loss and then minimise the expression, with respect to k, to determine the MIAEL estimator.

Now, $\forall\, z \,\varepsilon\, [a, b]$, the average expected loss is

$$AEL(z) = \frac{\int_{\phi \varepsilon f^{-1}(z)} EL_\phi(f(\phi), k)\, d\phi}{\int_{\phi \varepsilon f^{-1}(z)} d\phi}$$

But, $\forall\, \phi \,\varepsilon\, f^{-1}(z),\ f(\phi) = z$, hence

$$AEL(z) = (1 - k/z)^2$$

The integrated average expected loss is

$$IAEL(k) = \int_{z=a}^{z=b} (1 - k / z)^2\, dz$$

Integrating and simplifying this gives

$$IAEL(k) = (b - a) - 2 * k * \ln(b/a) + k^2 * (1/a - 1/b)$$

To find the MIAEL estimator, I shall differentiate with respect to k, set the expression equal to 0, and solve for k.

$$d(IAEL(k)) / dk = -2 * \ln(b/a) + 2 * k * (1/a - 1/b) = 0$$

$$\rightarrow \qquad k = a * b * \ln(b/a) / (b - a)$$

Proof of result 2

I am given that $M(X)$ is the MIAEL estimator of $f(\theta)$ in D, EL_θ is an expected loss function, $c \in \mathbb{R} - \{ 0 \}$, and $\forall\ t(X) \in D\ \ EL_\theta(c * f(\theta), c * t(X)) = EL_\theta(f(\theta), t(X))$.

I shall prove that $c * M(X)$ is a MIAEL estimator with the same information index as $M(X)$. The main results are proven in (c) and (d) below; results (a) and (b) are preliminary results needed in (c) and (d).

(a) Take some $t(X) \in D$, to show $\forall\ z \in [a, b]$ that $AEL_f(z) = AEL_{c*f}(c * z)$.

$\forall\ c * z \in range(c * f)$, by definition

$$AEL_{c*f}(c * z) = \frac{\int_{\phi \in (c*f)^{-1}(c*z)} EL_\phi(c * z, c * t(X))\, d\phi}{\int_{\phi \in (c*f)^{-1}(c*z)} d\phi}$$

As $EL_\phi(c * z, c * t(X)) = EL_\phi(z, t(X))$, the result will follow if $(c * f)^{-1}(c * z) = f^{-1}(z)$. To show, $(c * f)^{-1}(c * z) = f^{-1}(z)$. Suppose $\phi \in (c * f)^{-1}(c * z)$, then $c * f(\phi) = c * z$, and $z = f(\phi)$, hence $\phi \in f^{-1}(z)$. Conversely, if $\phi \in f^{-1}(z)$, then $z = f(\phi)$ and $c * z = c * f(\phi)$, hence $\phi \in (c * f)^{-1}(c * z)$.

(b) To show $\forall\ t(X) \in D$ that $IAEL_f(t(X)) = |c| * IAEL_{c*f}(c * t(X))$.

Take any $t(X) \in D$, then by definition

$$IAEL_f(t(X)) = \int_{z=a}^{z=b} AEL_f(z)\, dz$$

and

$$IAEL_{c*f}(c * t(X)) = \int_{x \in range(c*f)} AEL_{c*f}(x)\, dx$$

The result follows by making the substitution $x = c * z$ in the previous equation. For example, if $c > 0$, then

$$IAEL_{c*f}(c * t(X)) = \int_{x=c*a}^{x=c*b} AEL_{c*f}(x)\, dx = \int_{z=a}^{z=b} AEL_{c*f}(c * z)\, dz * c$$

which, by (a), is $c * IAEL_f(t(X))$.

(c) To show that $c * M(X)$ is the MIAEL estimator of $c * f(\theta)$ in $c * D$.

Take any $c * t(X) \varepsilon\ c * D$, then

$$\begin{aligned} IAEL_{c*f}(c * M(X)) &= IAEL_f(M(X)) / |c| && \text{from (b)} \\ &\leq IAEL_f) / |c| && \text{by definition of } M(X) \\ &= IAEL_{c*f}(c * t(X)) && \text{from (b).} \end{aligned}$$

Hence, $c * M(X)$ is the MIAEL estimator of $c * f(\theta)$ in $c * D$.

(d) To show, for any set of indices I, that $Info(I, c * f) = Info(I, f)$.

By definition

$$Info(I, f) = 1 - IAEL_f(M(X)) / IAEL_f(k)$$

where k is the *best* k estimator of $f(\theta)$.

But, from (b)

$$IAEL_f(M(X)) = |c| * IAEL_{c*f}(c * M(X))$$

and $IAEL_f(k) = |c| * IAEL_{c*f}(c * k)$.

Hence

$$Info(I, f) = 1 - IAEL_{c*f}(c * M(X)) / IAEL_{c*f}(c * k)$$

which, by (c), is $Info(I, c * f)$ since both $M(X)$ and k are MIAEL estimators.

Appendix 2: Estimation of equivalent sample sizes for multinomial age frequency random variables

The age frequency data used in this assessment was derived from length frequency data from the commercial catch (using MIX). To apply the MIAEL estimation method I need, in addition to the observed age frequencies, the distribution of the associated random variables (so that simulated age frequency data can be generated). The random variables associated with the observed age frequencies have a complex distribution which is dependent on the method of sampling length frequencies, the sample sizes, and the procedure used to transform length frequencies into age frequencies.

I have assumed that the derived age frequency, for a given sex in a given year, has a multinomial distribution with a sample size proportional to the number of tows sampled in that year. Further, I have estimated the constants of proportionality, for each sex, by assuming that the sample size for the year with the median number of tows is equal to the statistic $n^{\#}_{sex}$ (which is derived below). That is, for year j, and for a given *sex*,

$$(\text{sample size})_{j,sex} = n^{\#}_{sex} * t_j \;/\; t_{med}$$

where t_j = the number of tows sampled in year j, and $t_{med} = median\ \{\ t_j\ \mid\ j \in J\ \}$ (where J is the set of years for which age frequency data is available).

The statistic $n^{\#}_{sex}$ is a median of medians, derived as follows.

For a given sex and year: for $i = minage...maxage$, let Y_i be the observed proportion of fish (in the catch) of age i, and let S_i be the standard error of Y_i. Suppose, for $i = minage...maxage$, that $Y_i \approx X_i/n$ where $(X_{minage},...,X_{maxage})$ has a multinomial distribution with parameters $(p_{minage},...,p_{maxage},n)$.

Now,

$$var(X_i/n) = p_i\,(1 - p_i)\;/\;n$$

If we equate Y_i with p_i and S_i^2 with $var(X_i/n)$ then an estimate of n can be derived (for each i):

$$S_i^2 = Y_i\,(1 - Y_i)\;/\;n_i^*$$

so

$$n_i^* = Y_i\,(1 - Y_i)\;/\;S_i^2$$

A single estimate of n is needed and I have used,

$$n^* = median\ \{\ n_i^*\ \mid\ i = minage...maxage\ \}$$

Such an estimate of n is available for each sex in each year, and $n^{\#}_{sex}$ is the median (across years) of the estimates:

$$n^{\#}_{sex} = median\ \{\ n_j^*\ \mid\ j \in J\ \}$$

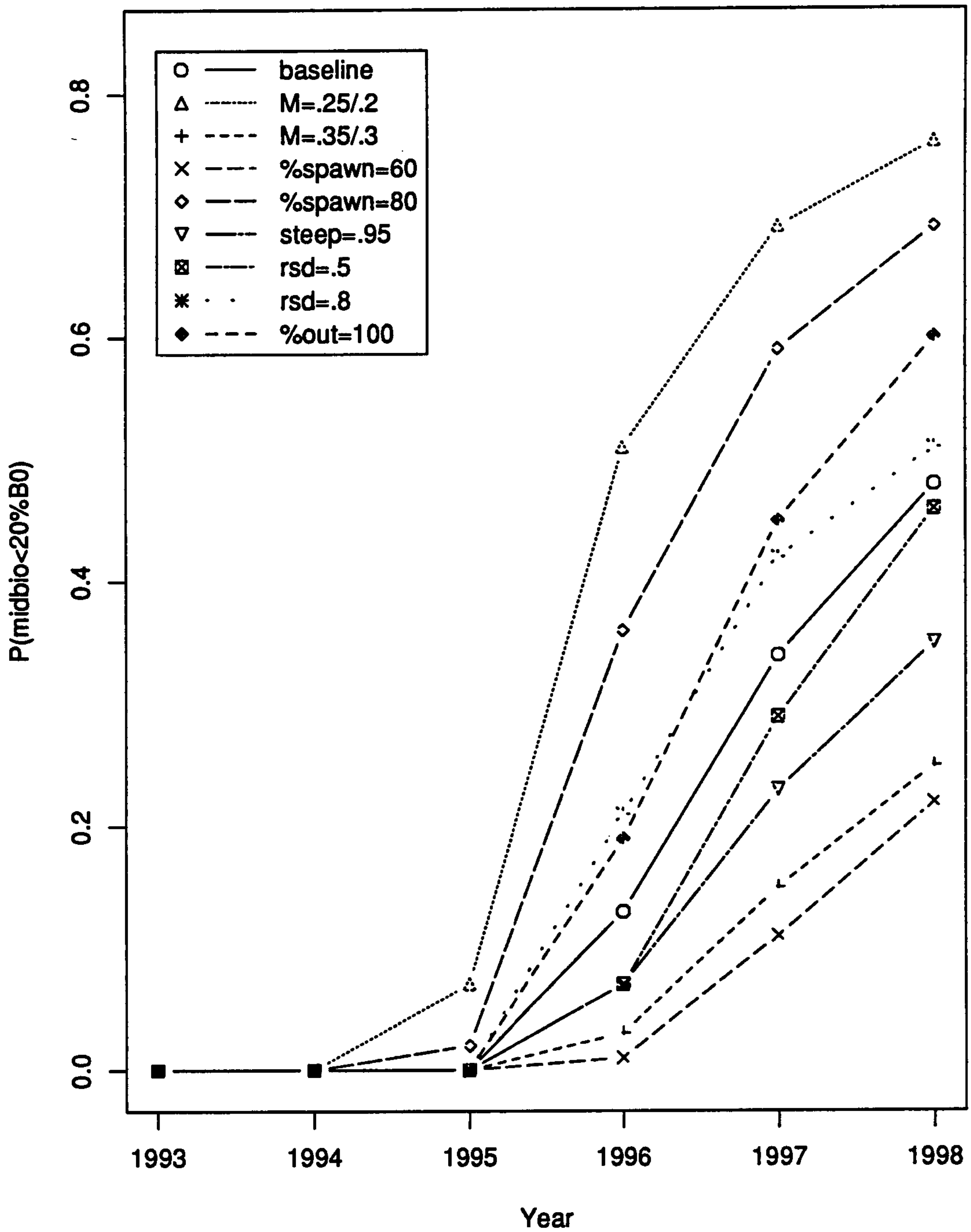

Figure 1. Maximum annual stock risk in 1993 to 1998 for each option modelled in the sensitivity analysis.

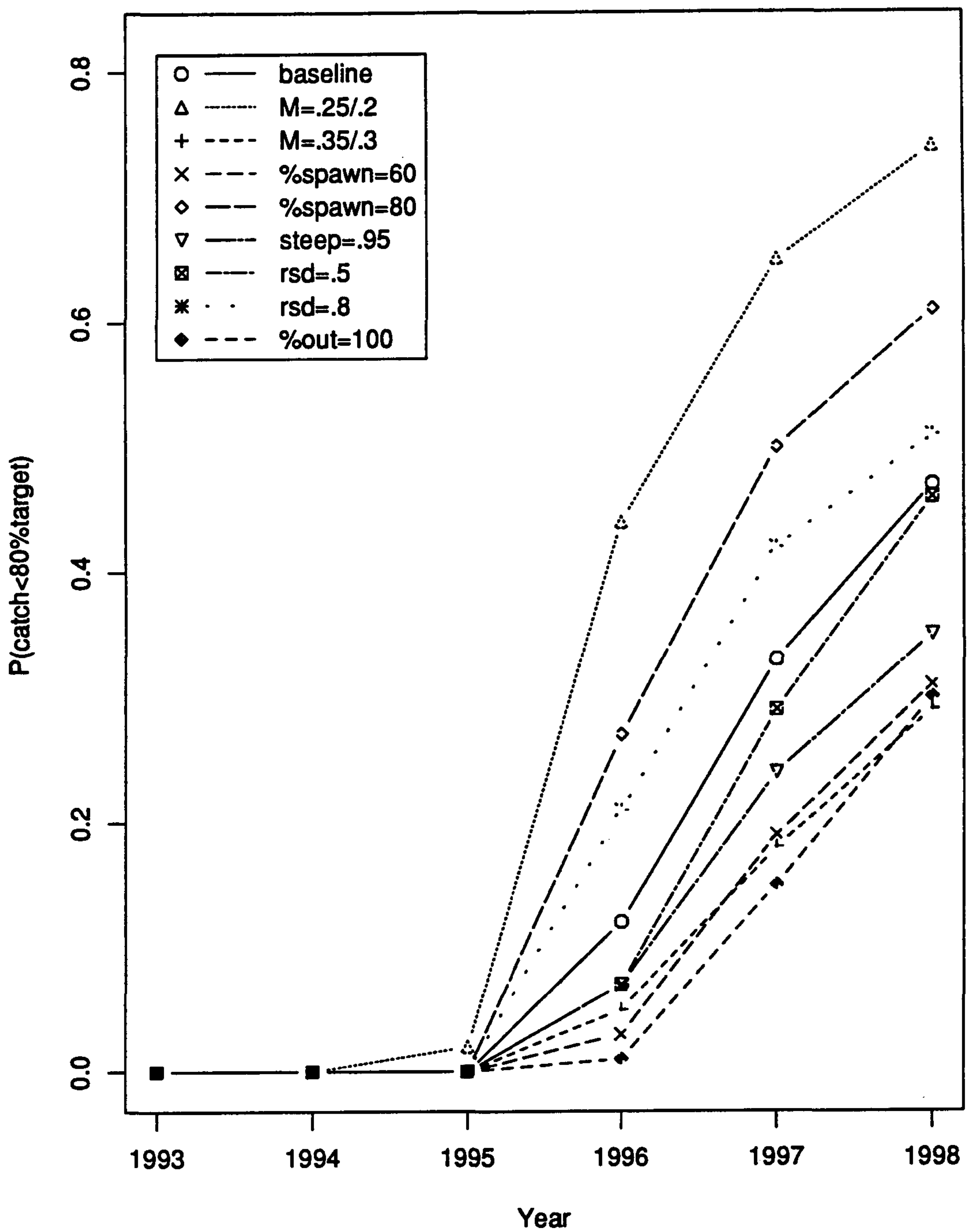

Figure 2. Maximum annual fishery risk in 1993 to 1998 for each option modelled in the sensitivity analysis.