MINISTRY OF FISHERIES

Te Tautiaki i nga tini a Tangaroa

ISSN 1175-1584

Management procedure evaluations for rock lobsters in CRA 3 (Gisborne)

Paul A. Breen
Susan W. Kim
Vivian Haist
Paul J. Starr

New Zealand Fisheries Assessment Report 2005/61
December 2005

Management procedure evaluations for rock lobsters in CRA 3 (Gisborne)

Paul A. Breen[1]
Susan W. Kim[1]
Vivian Haist[2]
Paul J. Starr[3]

[1]NIWA
Private Bag 14901
Wellington

[2]Haist Consultancy
1262 Marina Way
Nanoose Bay, B.C.
Canada V9P 9C1

[3]Starrfish
61A Rhine Street
Wellington

New Zealand Fisheries Assessment Report 2005/61
December 2005

Published by Ministry of Fisheries
Wellington
2005

ISSN 1175-1584

Citation:
Breen, P.A.; Kim, S.W.; Haist, V.; Starr, P.J. (2005).
Management procedure evaluations for rock lobsters in CRA 3 (Gisborne).
New Zealand Fisheries Assessment Report 2005/61. 71 p.

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

EXECUTIVE SUMMARY

Breen, P.A.; Kim, S.W.; Haist, V.; Starr, P.J. (2005).
Management procedure evaluations for rock lobsters in CRA 3 (Gisborne).

New Zealand Fisheries Assessment Report 2005/61. 71 p.

This report describes work done to fulfill Objective 6 of the MFish contract CRA2003-01, a three-year contract awarded to the New Zealand Rock Lobster Industry Council. The purpose of the work was to develop a management procedure to determine annual commercial catch limits in the depleted CRA 3 (Gisborne) rock lobster fishery.

Specifications for the management procedure were obtained from CRA 3 stakeholders at an industry meeting held in Gisborne in July 2004. These included a target CPUE of 0.75 kg/pot lift in the autumn-winter season, the goal of keeping biomass above *Bmin*, and an asymmetric latent year.

An operating model was constructed based on the most recent CRA 3 assessment, which had been done with a Bayesian length-based population model in 2004. Some changes were made to the model dynamics and structure: non-commercial catches were based on an exploitation rate in projections, the proportion of commercial catch taken in the AW season was assumed to be related to CPUE, recruitment was projected by sampling from a continuous distribution rather than be re-sampling recent recruitments, recruitment was serially autocorrelated, observation error was added to projected CPUE, unnecessary functions of the assessment model were stripped away and the model was modified to make projections one year at a time and use a harvest control rule to determine the next year's catch.

Initial exploration with constant catch and constant exploitation rate rules yielded some appreciation of the likely production characteristics of the stock: maximum mean yield under a constant exploitation rate strategy is about 250 t; attempting to obtain higher yields causes a significant percentage of years to have biomass less than *Bmin*. The mean catch associated with mean CPUE near 0.75 kg/potlift is about 200 t under a constant exploitation rate strategy.

Four different harvest rule families were developed and tested, all using observed CPUE in each year to determine what the catch limit should be in the next year. Each family has rule parameters that specify different members of the family. In all, we tested 215 rules. We defined a set of indicators, based on yield, safety, stability and performance with respect to the target, for comparing rules.

Each rule was tested by making a set of runs, with 1073 runs in the set, based on samples of the joint posterior distribution of parameters from the CRA 3 assessment. Rules were compared using "winnowing", which eliminated rules with patently sub-standard performance with respect to some indicators, "screening", in which the we compared the relative probability of delivering critical outcomes, and "choice frontiers", which can be used to find the "best" rules with respect to critical trade-offs.

Candidate rules were also evaluated for robustness by making additional sets of runs in robustness trials, with various changes to the operating model system.

The results comprise a set of candidate rules, and detailed data on their performance, that could be used as the basis of choice for the CRA 3 stakeholders.

Contents

1. INTRODUCTION

This work was conducted under Objective 6 of the MFish contract CRA2003-01, awarded to the New Zealand Rock Lobster Industry Council (NZ RLIC). The objective is *To evaluate new management procedures for rock lobster fisheries.* In discussion with the National Rock Lobster Management Group (NRLMG) it was agreed that this work should apply to CRA 3. It was also agreed that the work should use the 2004 assessment model for CRA 3 as the operating model. This decision caused changes to the reporting schedule for this Objective that were addressed by way of a contract variation.

1.1 CRA 3 fishery

The CRA 3 fishery extends from East Cape south to the Wairoa River. The 2004[1] total allowable catch (TAC) of 453 t, set in 1998, comprised 20 t allowances for amateur catch and customary harvest, 86 t for illegal removals and a total allowable commercial catch (TACC) of 327 t, distributed amongst 35 quota share owners. Effective 1 April 2005, the Minister of Fisheries reduced the TAC to 319 t by retaining the existing allowances for recreational and customary fishers, reducing the commercial TACC to 190 t and increasing the allowance for illegal catch to 89 t.

The recent history of the fishery (Figure 1) is more dramatic than in some other areas. The original TACC when the quota management system (QMS) was established in 1990 was 437 t, reducing to 327 t by 1992. In 1992 the fishery was in a seriously depleted state (Breen & Kendrick 1997), and a suite of management measures was imposed that included a TACC reduction to 164 t. The fishery showed strong recovery beginning in 1993, probably in response to the management measures combined with strong recruitment that was reflected in CPUE increases seen in CRA 1 and CRA 2 at the same time, and slightly later increases in CRA 4 and CRA 5. There was a strong shift from the spring-summer season, October through March (SS), to the autumn-winter season (AW) as CPUE increased. As a result of the strong recovery to levels higher than previously seen, the CRA 3 TACC was increased to 205 t in 1996 and back to 327 t in 1998.

Non-commercial catches are a major uncertainty in CRA 3. MFish have produced various estimates of illegal catches over time; the most recent is 89 t but the peak estimate was 250 t for 1992–93. MFish are unable to give any guidance on the nature of illegal catch, for instance whether it comprises mostly lobsters of all sizes and both sexes, or mostly scrubbed females, or mostly under-sized, etc. The assessment assumed that the illegal fishery operates on the full range of sizes and both sexes available to pots in both seasons, and assumed that the seasonal distribution of illegal catch follows the commercial catch. The assessment also assumed that a small proportion (5%) of the illegal catch is reported as legal catch, and adjusted the reported commercial catch downward by this amount.

---

[1] The fishing year for rock lobsters is from 1 April through 31 March of the following year. The convention used in this report is to name fishing years by the portion with 9 months, *viz.* the fishing year 2004–05 is referred to as "2004".

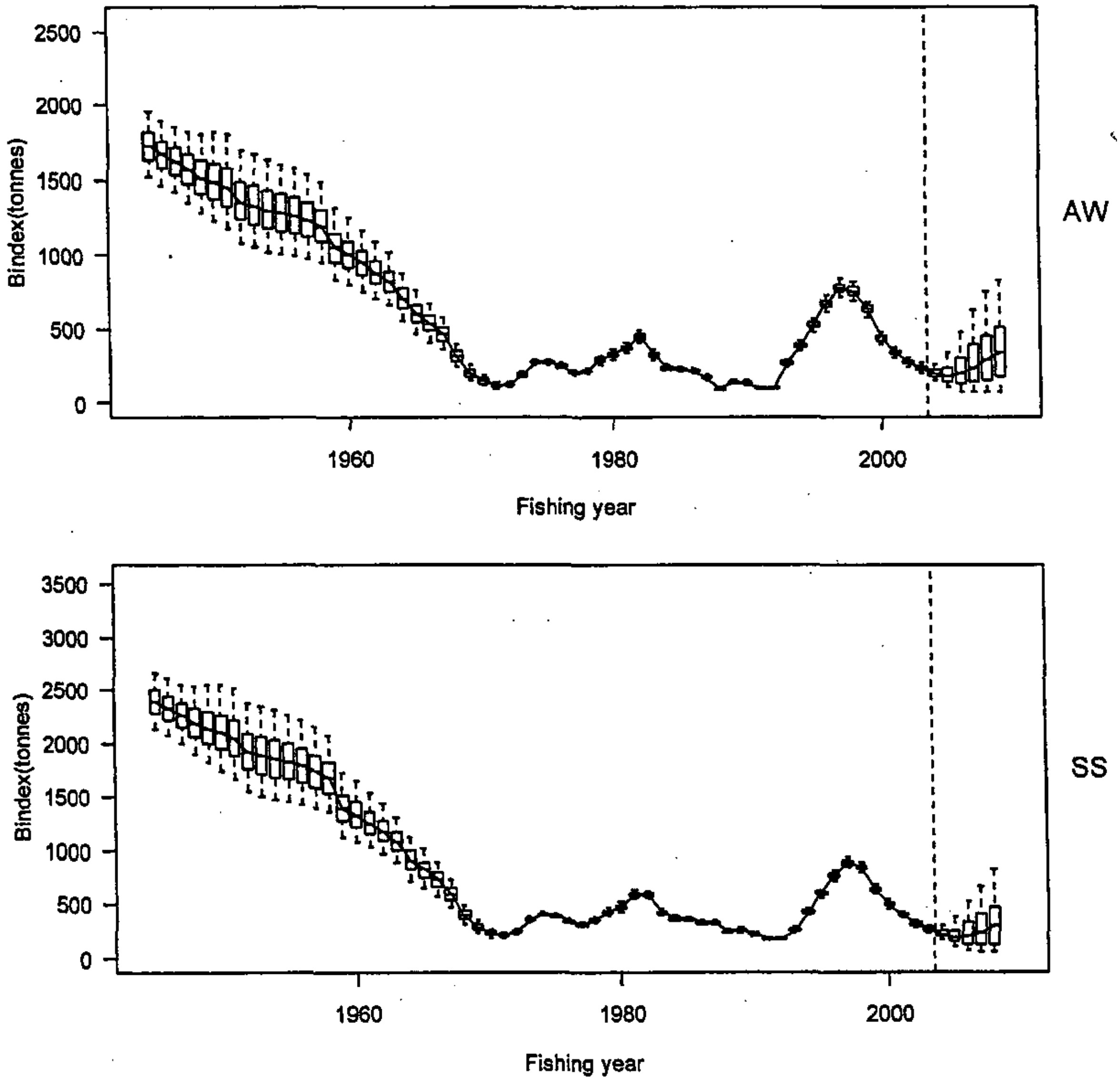

Figure 1: The posterior trajectory of mid-season biomass available to the fishery, by season (AW: autumn-winter; SS: spring-summer), from the CRA 3 base case Markov chain – Monte Carlo simulations. For each year the horizontal line represents the median, the box spans the 25th and 75th percentiles and the dashed whiskers span the 5th and 95th percentiles.

Recreational and traditional estimates are very uncertain. The Recreational Fishery Assessment Working Group (RFFAWG) rejected the 1994 and 1996 diary surveys of CRA 3 because of method problems. The RFFAWG recommended that the 2000 and the follow-up 2001 diary surveys be used with great caution because of uncertainty in the accuracy of diary catch estimates. The Rock Lobster Fishery Assessment Working Group (RLFAWG) rejected all three recreational diary surveys for the 2004 assessment and agreed to assume that recreational and customary catches had always been at their current allowances of 20 t; it noted that these assumptions are obviously highly uncertain.

CPUE peaked at very high levels in 1997 (AW) or 1998 (SS; this estimate is much less well determined because of small SS catches in the 1996–98 period) and has since declined, although not as far as the 1992 levels. Recent levels caused concern to MFish and to the CRA 3 stakeholders. In 2003 the CRA 3 Industry Association, after taking advice from fisheries research and management service providers, initiated an extensive consultation with industry members to determine a response to ensure a rebuild of stock abundance. With assistance and advice from FishServe, the NZ RLIC coordinated a quota shelving program limiting the commercial catch to 210 t in 2004.

The 2004 assessment used two reference levels: a target biomass that corresponds with 0.75 kg/potlift CPUE (standardised CPUE in the AW season), and a minimum biomass, *Bmin*, that corresponds with the lowest level of biomass in the stock's history. The assessment found that the current biomass (AW 2005) is likely to be higher than *Bmin*, but is less than the target reference point. Stakeholders agree that the fishery is depressed, especially compared with the high abundance of the late 1990s. CRA 3 stakeholders agreed to use a management procedure to determine commercial catch limits for each year, with the goal of rebuilding the fishery to the target level and maintaining it there.

1.2 Management procedures

The operational management procedure approach was developed in South Africa (Butterworth et al. 1997, Cochrane et al. 1998), was adopted by the International Whaling Commission (Kirkwood 1997), and is now reasonably widespread. Johnson & Butterworth (2005) described choosing management procedures to manage South African rock lobsters (*Jasus lalandii*). For some time, a management procedure (originally developed by Starr et al. 1997) has been used to control catch in the combined CRA 7 and CRA 8 fishery (NSS).

Management procedures pre-specify how management changes will be made based on some fishery data. A management procedure is "a fully specified feedback control system applied as part of a fishery management system" (McAllister et al. 1999) and specifies what data will be collected, how they will be collected and processed, what estimates will be made from the data, and how those estimates will determine harvest controls. Good reviews were provided by Butterworth & Punt (1999) and McAllister et al. (1999).

Designing good management procedures demands an understanding of what management is trying to do. Goals can usually be classified as belonging to one of these classes:

- yield: the average catch over time,
- abundance: stock size,
- safety: the risk that stock abundance will fall to low levels,
- stability: the amount by which catch limits change from year to year,
- diversity: the range of sizes of fish available,
- economics: involving both yields and costs and
- social goals, which are difficult to define.

A common perception is that fishers would like to maximise yields, and much of fisheries science was once grounded in this idea (e.g., maximum sustainable yield (MSY) was once a widespread quantitative goal). But, especially where fishers have property rights such as ITQs, fishers' goals are more complex.

The goals listed above cannot all be maximised. An obvious management trade-off, for instance, is between average catch and stability of catch. To maximise mean catch requires frequent catch changes; sustainably stable catches are lower catches than could otherwise be obtained. To maximise stability requires lower catch limits. Fishers who want stable catch limits cannot also have catches near their maximum average levels.

Another trade-off is between risk and catch: higher mean catch implies lower stock abundance and higher risk. Yet another is between mean catch and CPUE: higher mean catch implies lower CPUE; higher CPUE implies greater ease and lower costs of catching. One strategy may produce sustainable high catches from a low average abundance, another may result in lower average catches from higher average abundance (hence at lower cost). The choice between alternatives depends upon the relative importance of each goal.

The relative importance of competing objectives such as yield, safety, abundance and stability should be clearly defined, we believe, by stakeholders and certainly not by fisheries scientists. In a workshop held in New Zealand (Bentley et al. 2003a), lobster fishers placed importance on stability in catch quotas, maintenance of high CPUE (high abundance), and maintenance of a wide size range of lobsters so that fishers could respond to changes in market demand; i.e. abundance, stability and diversity. Fishers stated clearly that they were willing to trade some potential catch for stability and abundance goals.

A cynical view might be that it is easy for fishers to claim to value stability and abundance above high mean yield. Fishers in CRA 5 demonstrated their preference irrefutably in 2003 when they rejected an opportunity to pursue a TACC increase based on the favourable stock assessment of Kim et al. (2004).

It is essential that management procedures are extensively evaluated. Our approach to this objective follows work in designing and evaluating the 2002 NSS management procedure (Bentley et al. 2003b), evaluating procedures that might be useful for lobster fisheries (Breen et al. (2003) and procedures for managing sea lion bycatch (Breen et al. 2004, Breen & Kim 2005). This involves using a system of sub-models, including a population simulator or "operating model" (we base ours on the length-based stock assessment model), an observation model that simulates the population signal, and the harvest control rule model. Catch limits determined by the harvest control rule model are fed back into the population model in a feedback loop, to make a single run of 20–100 years. This whole process is repeated with simulated stochastic error, discussed further below, to make a large set of runs from which the distributions of a set of indicator values can be examined. In turn, that whole process can be repeated for different variants of a specific harvest control rule, for completely different rules, for variants of the population model, and for different simulated realities.

Management procedures were evaluated in previous work conducted under the MFish contract by NZ RLIC. The NSS decision rule was extensively re-evaluated and a new rule was adopted in 2002 (Bentley et al. 2003b). A variety of harvest control rules for use in management procedures were evaluated in 2002 (Breen et al. 2003). This study explored the literature and found a wide range of harvest controls that varied in how they used the index (usually CPUE). The major possibilities (rules may more than one of these) are:

- comparing CPUE with a target,
- using the CPUE gradient,
- estimating biomass from CPUE and
- estimating surplus production from two or more years.

Harvest control rules also vary in how they use the results of the initial calculation. Nearly all rules *modify (buffer) the response*, by one or more of

- combining calculations to make a hybrid rule,
- using a sensitivity parameter,
- combining the rule calculation with the previous catch limit,

- using thresholds to prevent small changes to catch limits,
- bounding the maximum changes for catch limits,
- bounding the minimum or maximum catch limits,
- *using alternative calculations based on the level of* CPUE,
- using moving averages to smooth out noise in CPUE and
- restricting changes to alternate years with "latent years".

In the Breen et al. (2003) study, many rules developed unstable oscillations; these results suggested that constant-exploitation rate strategies are, relative to others, well-behaved. They also showed (this was not original) that lags between observation and action, and latent years, degrade the performance of harvest control rules. When some catch components were not limited by the rule, such as illegal or recreational catches, catch tended to move from controlled to uncontrolled components, with poor results. This is a serious problem for some areas in New Zealand where non-commercial lobster fishing takes considerable catch but cannot be controlled through a management procedure. Customary fishing is legally immune from restriction; recreational fishing is limited by size and bag limits but effort is unrestricted; illegal catch is difficult to restrict.

The possibility of incorporating additional indices was explored in 2003 (Bentley et al. 2005). This work is incomplete, but it appears that pre-recruit indices might be useful for a management procedure when combined with a simple delay-difference approach. This work also suggested that length frequency data, CPUE trends and puerulus settlement indices are all less likely to be useful.

The NSS "decision rule" (Bentley et al. 2003b) is the only operational management procedure currently in use in New Zealand. This procedure has a CPUE target, which is an empirical target based on the fishery history, and a target trajectory from the initial position to the target. It uses CPUE as the indicator variable and calculates both a position reference (is current CPUE above or below the target?) and a gradient reference (how does the rate of CPUE change compare with the target trajectory's slope?). The rule then calculates a new TACC from these references and the rule parameters. The TACC is changed if the amount of change calculated under the rule exceeds a threshold value, and if the TACC had not been changed in the previous year's evaluation (a "latent year" component prevents change in two successive years).

This rule was chosen from a large family of alternatives (125 rules), each alternative specified by the set of rule parameters. Rules were run under different operating model assumptions that related to hypotheses about stock structure. For each rule for each operating model, thousands of stochastic runs were made from samples of the joint posterior distribution of parameters obtained by the stock assessment. The final rule was chosen by the NRLMG after considering the rules with the highest joint probability of delivering three critical outcomes: increasing CPUE, rebuilding the stock by the target time and having a low catch variation (Bentley et al. 2003b).

In this study we present a number of steps. We describe a basic operating model based on the most recent stock assessment. We choose a (very limited) number of harvest control rule families from the (very large) set of possibilities. We describe a set of performance indicators from which to judge rules. We describe some simple approaches to evaluating the large number of alternative candidate rules, and present results of such evaluation. Limited robustness testing is described.

1.3 CRA 3 specifications

The use of operational management procedures was discussed with CRA 3 stakeholders at a meeting in Gisborne on 20 May 2004. Another meeting was held in Gisborne on 29 July 2004 to obtain a mandate to proceed with management procedure development, and to obtain specifications for the particular needs of CRA 3 fishers.

At the first meeting we presented a study (Starr et al. unpublished) that explored the precision of standardised CPUE estimates that could be obtained from partial data available before the end of the season, estimates that could be used to determine a catch limit for the next year. This study addressed the problem of having a one-year lag between the data and the management decision. In the NSS management procedure, data from fishing year *y* are used in year *y*+1 to estimate standardised CPUE, and results are used to modify the commercial catch for year *y*+2. Such a lag can cause instabilities in the behaviour of harvest control rules, and almost always degrades rule performance. The study showed that good estimates could be obtained from the first six months of data from year *y*, using all the data from previous years, in time to have use in a management procedure in year *y* to modify the commercial catch for year *y*+1, thus eliminating one whole year of lag.

The second meeting discussed specific goals for the CRA 3 management procedure and approved the general goals of stability (fewer TACC changes than more), safety (staying above *Bmin*), and some level of abundance that reflects a good balance between costs and yield. Interestingly, nobody pursued the goal of maximising yield.

Specific items agreed were:

- the harvest control rule should be based on AW standardised CPUE;
- the target AW CPUE is 0.7 *to* 0.8 kg *per* pot-lift;
- the safety indicator should be based on *Bmin* from the assessment;
- if the rule mandates a decrease in commercial catch, that should happen without regard to a latent year and
- if the rule mandates an increase in commercial catch, that should be applied with a latent year.

The rule was first seen as a quota shelving rule, but the Minister's decision for 2005–06 suggests that any management procedure should be a TAC or TACC-adjusting rule. The asymmetric latent year allows consecutive catch limit decreases, but allows increases only when no change occurred in the previous year.

2. OPERATING MODEL

2.1 Overview

The operating model is a modification of the 2004 assessment model (Haist et al. 2005), which had been used to produce sets of projections with fixed projected catch values, based on the large set (7505) of samples from the joint posterior distribution. The assessment model had been modified slightly from the 2003 assessment model (Kim et al. 2004), most notably by addressing the existence of a marine reserve in CRA 3 since 1999. The large set of posterior samples was thinned uniformly for this study to a smaller set of 1073 samples.

In what follows, projected quantities are denoted by a head symbol, thus $\breve{N}^{g}_{t,l}$, to distinguish them from quantities estimated in or dependent on the assessment.

Assumptions of the assessment model were described by Kim et al. (2004). They are reasonably straightforward: for instance, growth for males occurs when they moult at the end of the AW and SS seasons, whereas female growth occurs at the end of SS only; mean increment-at-length is different for males and females and is a declining (not necessarily linear) function with normally distributed variation related to the expected increment. Natural mortality is constant across sizes and time. Handling mortality is 10% on lobsters returned to the sea by commercial fishers.

Additional assumptions were introduced for the 2004 assessment: it was assumed that the Te Tapuwae o Rongokako marine reserve has no stock-recruit effect and no yield-per-recruit effect, but that the removal of some part of the original stock (10% was assumed) had a negative effect on available stock size and subsequent recruitment.

For the operating model, several assumptions were varied, or new assumptions were made:

- non-commercial catches were assumed to be proportional to biomass in each season, through an exploitation rate calculated from catch (assumed non-commercial catches from the assessment, by season) and biomass (mean model biomass for 2001–03 by season);
- the proportion of commercial catch taken in the AW season was assumed to be related to CPUE;
- recruitment was simulated using recruitment deviations with the same mean and standard deviation as those seen in 1964–2000;
- recruitment deviations were serially autocorrelated and
- projected CPUE was assigned log-normally distributed observation error.

Changes made to the assessment model to produce the operating model involved

- coding the assumptions listed above,
- switching off or deleting large sections of the assessment model not needed by the operating model, such as predictions for comparison with data, likelihood calculations, normalised residuals, outputs for plotting, etc.,
- modification of the model to run projections one year at a time and
- incorporation of harvest control rules simulated at the end of each year's projection, based on recent CPUE.

Any aspect of the operating model dynamics not discussed below is described in the assessment model dynamics (Kim et al. 2004; Haist et al. 2005).

2.2 Non-commercial catches

The assessment team discussed various options for non-commercial catches: assuming constant values, assuming both initial values and trends in time, and assuming constant exploitation rates. We considered the last to be most realistic, and it was modelled based on the assessment model's estimated biomass and the estimates or assumptions of recent non-commercial catches.

The 2003 estimates of AW non-commercial catches are termed $C^{illegal}_{AW,2003}$, $C^{rec}_{AW,2003}$ and $C^{cust}_{AW,2003}$ for the illegal, recreational and customary catches respectively; similarly for the SS season. The assessment pre-processing work differentiates between reported and non-reported illegal catches, but the percentage of the former is small and the operating model dynamics ignores this category.

The assessment model calculates the biomass of lobsters legally available to the fishery in period t as

Eq 1 $$B_t^{SL} = \sum_g \sum_s N_{s,t}^g W_s^g V_{s,k,z}^g L_{s,t}^g$$

where *g* indexes sex, *s* indexes size, *k* indexes season and the terms in the summation are numbers-, weight-, vulnerability- and legality-at-length respectively. This biomass is available to all catch sectors. The biomass vulnerable **only** to the illegal and customary fisheries, which do not respect the berried female and minimum legal size regulations, is

Eq 2 $$B_t^{NSL} = B_t^{total} - B_t^{SL} = \sum_g \sum_s N_{s,t}^g W_s^g V_{s,k,z}^g \left(1 - L_{s,t}^g\right)$$

These biomasses are calculated at the beginning of the season. The recreational and commercial catches are assumed to come only from the B_t^{SL}; the illegal and recreational catches are assumed to come from both the B_t^{SL} and B_t^{NSL} biomass components. Non-commercial exploitation rates, $U_k^{illegal}$ etc., for the AW and SS seasons are calculated from the 2003 catches and the average of the last three years' of model biomass:

Eq 3 $$U_{AW}^{illegal} = C_{AW,2003}^{illegal} \bigg/ \frac{1}{3} \sum_{y=2001}^{y=2003} \left(B_{AW,y}^{SL} + B_{AW,y}^{NSL}\right)$$

Eq 4 $$U_{AW}^{rec} = C_{AW,2003}^{rec} \bigg/ \frac{1}{3} \sum_{y=2001}^{y=2003} B_{AW,y}^{SL}$$

Eq 5 $$U_{AW}^{cust} = C_{AW,2003}^{cust} \bigg/ \frac{1}{3} \sum_{y=2001}^{y=2003} \left(B_{AW,y}^{SL} + B_{AW,y}^{NSL}\right)$$

and similarly for the SS season. Because these exploitation rates are calculated by season there is no need to divide the yearly non-commercial catch estimates among seasons; the procedure described locks the seasonal proportions into their 2001–2003 values. For the illegal catches, this is a change from the assessment model assumption that illegal catch follows the seasonal distribution of commercial catch.

2.3 Seasonal distribution of commercial catches

In projections made in the assessment, the projected commercial catch was constant and was assumed to be divided among the two seasons in the same proportions as the 2003 catch. The proportion of catch taken in AW was low before 1993 (20–30%, Figure 2). The fishery took a much higher proportion in the AW in 1993 as a result of the CRA 3 management package (Breen & Kendrick 1997) and this change persisted.

After 1993, CPUE, the tonnage taken in AW and the proportion of catch taken in AW all vary together (Figure 3). Both tonnage and the proportion taken in AW show a strong correlation with AW CPUE.

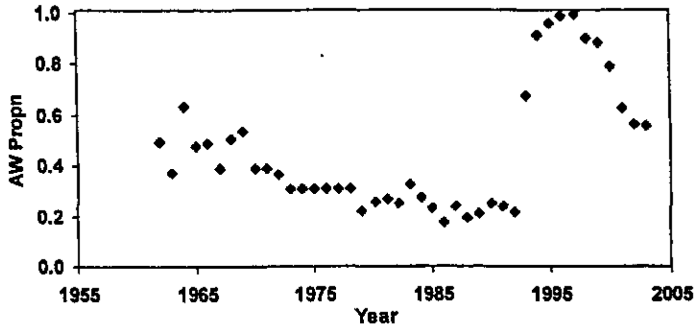

Figure 2: Proportion of the commercial catch taken in the AW season in CRA 3.

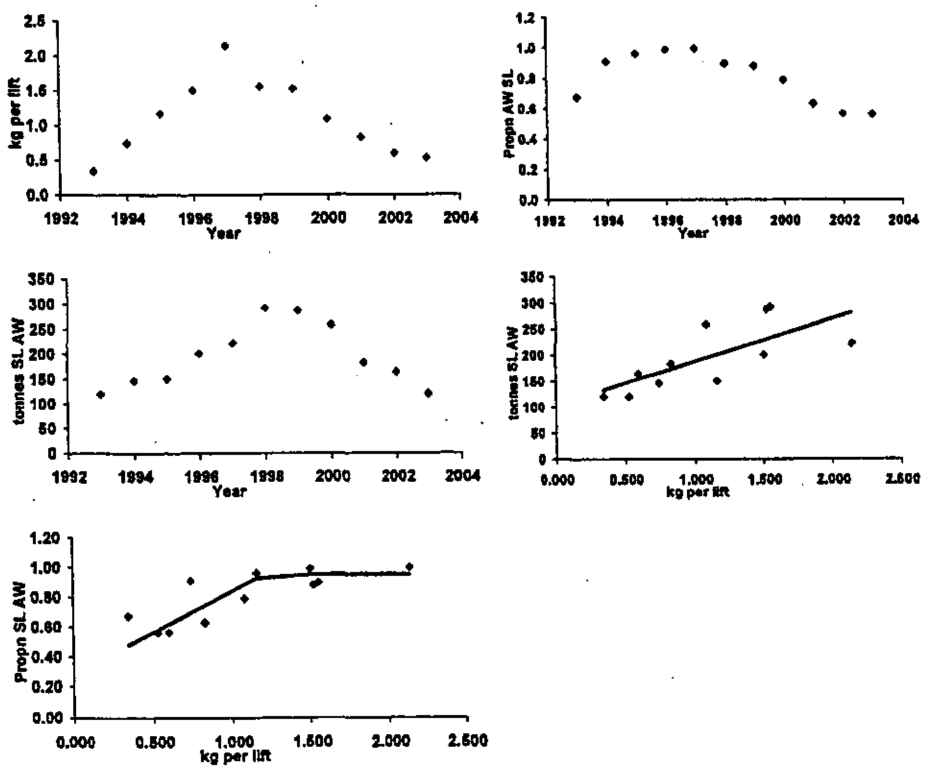

Figure 3: AW CPUE vs. year in CRA 3 after 1992 (top left), the proportion of commercial catch taken in AW (top right) and tonnage taken in AW (middle left); commercial tonnage taken in AW vs. AW CPUE (middle right), with a fitted relation (solid line); and proportion of the commercial catch taken in AW vs. AW CPUE (bottom left), with a fitted relation (solid line).

The lower panel of Figure 3 shows a simple model fitted to the data after 1993. It appears that the proportion taken in AW peaked at 95% at and above a CPUE of 1.25 kg per potlift. The relation up to that point is described by

Eq 6 $$P_{AW,y} = 0.288 + 0.548 I_{AW,y}$$

where $I_{AW,y}$ is CPUE. The operating model uses this sub-model to assign commercial catches to seasons, except that model CPUE is based on mid-season biomass, so an unavoidable one-year lag occurs:

Eq 7 $$\breve{P}_{AW,y} = 0.288 + 0.548\breve{I}_{AW,y-1}$$

2.4 CPUE observation error

Projected CPUE is used by the model's harvest control rules. Projected CPUE is based on AW mid-season biomass, or "index" biomass, using the estimated catchability coefficient, q, for each run. CPUE is observed with error, partly because CPUE is estimated from only part of the season's data, as would be the case in CRA 3 under the proposed management procedure. For use in the operating model's harvest control rules, to CPUE is added log-normally distributed error with mean zero and arbitrary standard deviation:

Eq 8 $$\breve{I}^{obs}_{AW,t} = 0.7578q\breve{B}^{index}_{t}\exp\left(\psi_t\sigma^{Iproj} - 0.5\left(\sigma^{Iproj}\right)^2\right)$$

where q is the catchability coefficient, $\psi_t = N(0,1)$ and σ^{Iproj} is specified. The constant 0.7578 converts CPUE from the assessment's relative index units (having a geometric mean of 1) to absolute kg/pot-lift; it is the geometric mean of CPUE over the period used in standardisation.

2.5 Recruitment

Recruitment is projected using recruitment deviations, in log space, that have the same properties as those seen in the assessment. For each of the thinned set of samples from the joint posterior (1073 samples), we calculated the mean, standard deviation and serial autocorrelation of the estimated ε_y parameters for 1964 through 2000. Years before 1964 have poorly estimated ε_y because the assessment model uses little data from before that year (Figure 4), and the last ε_y estimated was for 2000 because later estimates are unaffected by the data. The means of the mean, standard deviation and serial autocorrelation of ε_y from the assessment's posterior distributions were 0.1309, 0.6300 and 0.2734 respectively. Their posterior distributions are shown in Figure 5.

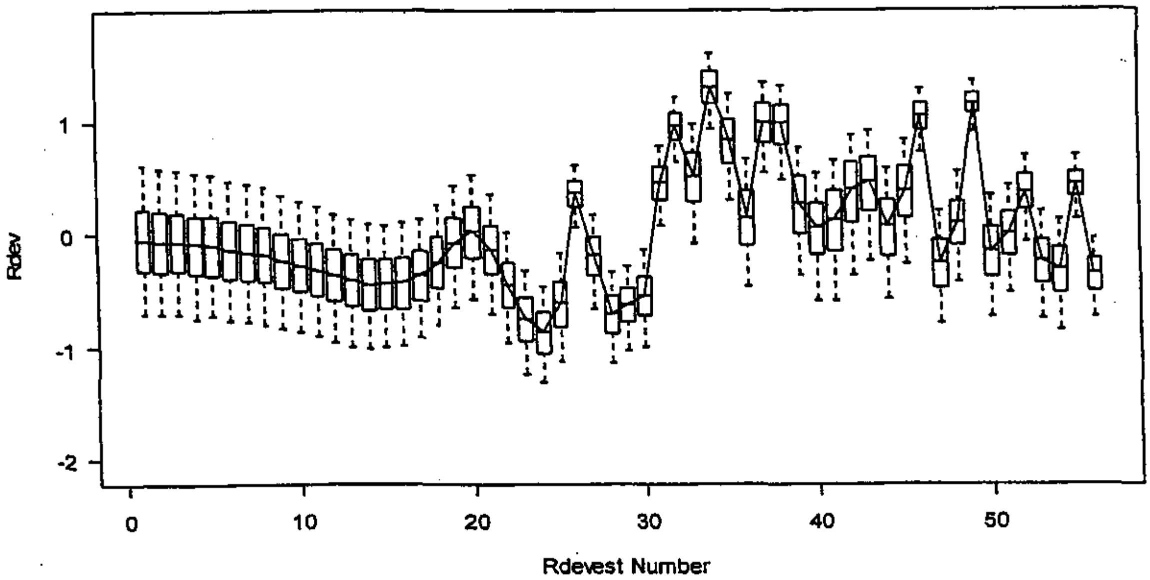

Figure 4: The posterior trajectory of recruitment deviation parameters ε_y from the CRA 3 base case Markov chain – Monte Carlo simulations. For each parameter the horizontal line represents the median, the box spans the 25th and 75th percentiles and the dashed whiskers span the 5th and 95th percentiles.

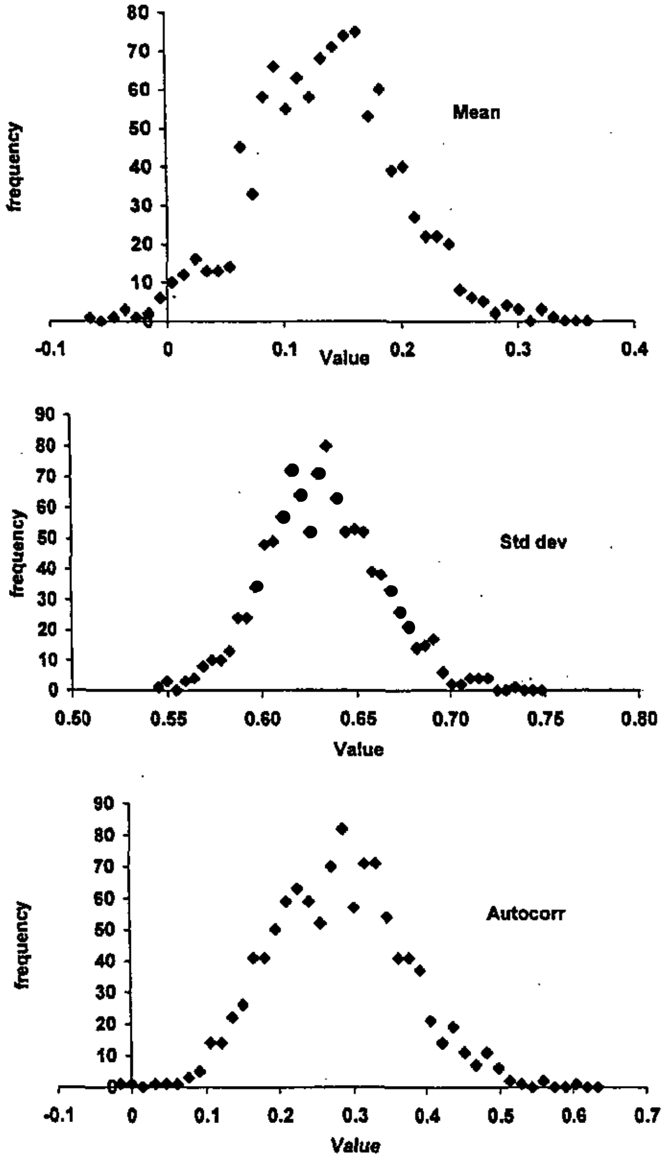

Figure 5: Posterior distributions of the mean (top), standard deviation and serial autocorrelation in ε_y for years 1964 through 2000 from the CRA 3 assessment base case.

For each sample, the operating model calculates the mean ($\bar{\varepsilon}$) and standard deviation (σ^{ε}) of the ε_y for 1964–2000, then calculates recruitment deviations ($\breve{R}dev$) in log space:

$$\text{Eq 9} \qquad \breve{R}dev_y = \rho \breve{R}dev_{y-1} + \sqrt{1-\rho^2}\left(\bar{\varepsilon} + \psi_y \sigma^{\varepsilon}\right) \text{ for } y \geq 2001$$

where $\psi_y \cong N(0,1)$ and ρ determines the amount of autocorrelation. Recruitment (it is the same in both seasons and for both sexes) is given by

Eq 10 $$\breve{R}^{g}_{AW} = \breve{R}^{g}_{SS} C^{reserve} R0 \exp\left(\breve{R}dev_y - 0.5\left(\sigma^{\bar{\varepsilon}}\right)^2 - 0.5\left(\sigma^{\varepsilon}\right)^2\right)$$

where $C^{reserve}$ is a constant that reduces annual recruitment by 10% after 1999 to simulate the effects of the Te Tapuwae o Rongokako marine reserve.

Mean recruitment from the model for 1964–2000 can be compared with the mean projected recruitment from the operating model in a 20-year run (Table 1). Projected recruited is about 7% less than the 1964–2000 mean recruitment before the effect of the marine reserve is applied: this is because the assessment assumed that $\sigma^{\varepsilon} = 0.4$ whereas the values used in projections, based on the assessment results, have a mean of 0.63, causing a lower mean recruitment (Eq 10).

Table 1: The mean, standard deviation and serial autocorrelation of recruitment multipliers from 1964-2000 and 2001-2005.

	1964-2000	2001-25
Mean	1.284*R0*	1.200*R0*
Standard deviation	0.839	0.845
Autocorrelation	0.176	0.175

2.6 Minimum legal size

In the assessment projections, minimum legal size (MLS) was taken to be the MLS regime in 2003 (Table 2). For the projections, MLS was specified by sex and season, and the specification remained constant for the whole projection. Projections described here all used the values shown in Table 2: a change in MLS for the AW season would require the CPUE target to be re-visited.

2.7 Assessment – projection transition

The assessment proper simulated the fishery through the 2003 season, ending 31 March 2004. For 2005, a TACC change was made, and the catches can be assumed to be the TACC. Thus the first true projection year in which catch could be determined by a harvest control rule, or in which MLS could be varied, is 2006.

The operating model is run from 2001 forward through 2004 and 2005. The 2001–03 "projections" are made because recruitment deviations cannot be reliably estimated for these years – no data give any information about them. As in the assessment, the projections generate stochastic recruitment for these years.

Assumptions for 2000-2005 are shown in Table 2. The commercial catch for 2004 was based on preliminary catch figures available in April 2005; the 2005 catch was assumed to be the TACC with the same seasonal split as the mean of 2001–2003; the 2006 catch was obtained with harvest control rules.

Table 2: Assumptions made in the projections from 2000 through the first true projection year, 2006. HCR: harvest control rule.

Year	Season	Commercial catch	Non-commercial catch	MLS (M,F)	Recruitment deviations
2000	AW	Observed	Assumed	52/60	Estimated
	SS	Observed	Assumed	54/60	Estimated
2001	AW	Observed	Assumed	52/60	Simulated
	SS	Observed	Assumed	54/60	Simulated
2002	AW	Observed	Assumed	52/60	Simulated
	SS	Observed	Assumed	54/60	Simulated
2003	AW	Observed	Assumed	52/60	Simulated
	SS	Observed	Assumed	54/60	Simulated
2004	AW	87 t	Calculated	52/60	Simulated
	SS	73 t	Calculated	54/60	Simulated
2005	AW	106 t	Calculated	52/60	Simulated
	SS	84 t	Calculated	54/60	Simulated
2006	AW	From HCR	Calculated	Specified	Simulated
	SS	From HCR	Calculated	Specified	Simulated

2.8 Performance indicators

Performance indicators permit of comparisons among harvest control rules. We chose a suite of indicators to reflect yield, abundance, safety, and stability (the classic management goals) and performance of the rules against the target CPUE. For all indicators, one value is derived for each of the 1073 runs in a set. Indicators from a set are summarised by the median of the marginal posterior distribution, in most cases, with exceptions noted below.

The yield indicators are:

- MeanCatch: the mean of annual projected commercial catch during each run,
- MinCatch: the minimum of annual projected commercial catch during each run and
- MeanNCCatch: the mean of annual projected non-commercial catch during each run.

The main safety indicator is

- *%<Bmin*: the percentage of years in which biomass fell below its 1992 level during each run (summarised for a set of runs by the **mean**).

Abundance indicators are

- MeanBiomass: the mean of $\breve{B}_y^{index}$ during each run,
- MinBiomass: the minimum of $\breve{B}_y^{index}$ during each run and
- BiomassRange: the difference between the minimum and maximum of $\breve{B}_y^{index}$ during each run.

The stability indicators are:

- CatchStdev: the standard deviation of commercial catch during each run and
- Nchanges: the number of TACC changes during each run.

The performance indicators that capture how well the procedure delivered the target are:

- % nearTarget: the percentage of years in which $\breve{I}_y^{rule}$ was within 10% of the target CPUE during the run (summarised for a set of runs by the **mean**),
- MeanCPUE: calculated from Eq 8 during the run and
- RebuildYr; where the "year of rebuilding" is the first year when $\breve{I}_{AW,y}$ meets or exceeds the target (summarised variously).

3. EXPLORATIONS WITH CONSTANT CATCH AND CONSTANT CATCH RATE

3.1 Specifications

Production characteristics of the operating model population, which ideally should be similar to the actual characteristics of the CRA 3 population, were examined by using simple constant-catch and constant-exploitation rate harvest control rules. These are not candidates for management procedures; they are used simply to explore the operating model productivity. The constant-catch rule could, in theory, be used as a real-life harvest control rule, although it wouldn't be a very good one, but a constant exploitation rate could not be used: vulnerable biomass is unknown because the scalar between biomass and CPUE, q, is unknown and CPUE is observed with error.

Sets of long (100-year) runs, through 2104 were made with different values of annual specified commercial catches (SCC) at intervals from zero to 900 t annually. A set of runs was 1073, comprising one run from each of the thinned set of base case joint posterior samples from the assessment.

In the second approach, the commercial catch target was calculated as a fixed percentage of the preceding year's mid-period AW biomass:

Eq 11 $\breve{C}_y^{target} = \breve{U}^{target} \breve{B}_{t-1}^{index}$

The specified target catch from this rule is based on AW mid-season biomass in the previous year. Mid-season biomass is less than pre-season biomass, recruitment occurs between the AW and SS seasons, and females are available in the SS but not the AW season; for all these reasons the target exploitation rate can reasonably exceed 1.

In these runs, the other projection variables were as shown in Table 3. The high MLS for females in AW reflects a prohibition on taking females in June, July and August, part of the CRA 3 management package introduced in 1993.

Table 3: Values for projection variables used for the exploratory runs of the operating model.

Variable	Value
Projections to	2104
MLS male AW	52
MLS male SS	54
MLS female AW	100
MLS female SS	60
ρ	0.10
σ^{ε}	0.00
σ^{Iproj}	0.00

3.2 Results from constant catch

MeanCatch increases linearly for specified commercial catches (SCC) less than about 210 t (Figure 6, Table 4). After this, MeanCatch continues to increase but at a decreasing rate and the results suggest an asymptote near 300 t.

The percentage of years where biomass falls below *Bmin* , %<*Bmin*, has median zero until SCC of 200 t, and rises rapidly towards an asymptote of 34%. MeanBiomass initially declines linearly as SCC increases. It is 800 t with an SCC of zero. At SCC of 200 t the rate of decline begins to decrease, and this indicator approaches an asymptote of about 118 t.

The non-commercial catch pattern is the inverse of the commercial catch pattern: at SCC of zero, MeanNCCatch is greatest at 228 t; this declines linearly as SCC increases to about 250 t, then declines more slowly towards an asymptote near 75 t.

MeanCPUE is near the target when SCC is near 200 t and actual mean catch is also near 200 t. The performance of this rule, if it were used as a management strategy, peaks at SCC of 150–210 t, when CPUE is close to the target 15% of the time. At lower specified catches, MeanCPUE tends to be greater than the target, and vice-versa. The median of RebuildYr increases exponentially with increasing SCC.

This set of runs shows first that there is no obvious "MSY". The relation between MeanCatch and SCC is not dome-shaped. This suggests that, at the current MLS and with the parameters estimated by the assessment model, fishery productivity is highly dependent on recruitment rather than on growth after recruitment. The "critical length" at which a size cohort would reach its maximum biomass may be less than the size at recruitment. For this fishery, a strict definition of *Bmsy* that tries to maximise yield may not be practical or desirable.

Second, these runs suggest that the average population productivity available to commercial catch, under the assumed dynamics of the operating model, may be 250 to 300 t, but with strongly diminishing returns (low biomass, CPUE less than the target, biomass falling below *Bmin*) when such high catches are sought.

Table 4: Indicators from exploratory runs of the operating model in which a constant catch was specified (SCC) for the entire 100-year run. The table shows median values from the posterior distribution of each indicator except for the percentages for <*Bmin* and CPUE near target, which are means. Median rebuild year of 3000 means that more than the runs did not rebuild by the end of the 100-year run.

SCC	Mean catch	Min catch	Mean nonC	%<*Bmin*	Mean biomass	Min biomass	Biomass range	Catch Stdev	%near Target	Mean CPUE	Rebuild year
0	0	0	228	0.0	793	220	1065	0	0.0	1.806	2007
50	50	50	205	0.0	679	219	949	0	1.0	1.549	2007
100	100	100	181	0.0	565	210	841	0	7.1	1.290	2007
150	150	150	157	0.0	452	166	761	0	15.2	1.031	2008
175	175	175	144	0.0	396	126	732	0	16.2	0.904	2009
190	190	190	137	0.0	363	99	717	0	16.2	0.827	2009
200	200	177	132	2.0	341	84	701	3	15.2	0.778	2010
210	209	171	127	3.0	320	76	684	6	15.2	0.728	2011
225	222	165	120	6.1	289	66	655	11	13.1	0.658	2012
250	240	160	108	11.1	243	59	595	21	10.1	0.556	2014
275	254	157	99	16.2	207	58	532	32	7.1	0.472	2018
300	264	156	93	22.2	180	58	474	42	5.1	0.410	2023
325	271	156	88	27.3	161	57	428	50	4.0	0.366	2030
350	276	157	84	30.3	148	57	388	56	3.0	0.336	2043
400	282	157	80	33.3	132	57	318	66	1.0	0.301	2068
500	288	158	78	33.3	121	57	240	75	0.0	0.277	3000
600	291	159	77	34.3	119	57	214	78	0.0	0.271	3000
700	293	160	77	34.3	118	57	204	80	0.0	0.270	3000
800	295	160	77	34.3	118	57	203	81	0.0	0.269	3000
900	296	160	76	34.3	118	57	203	81	0.0	0.269	3000

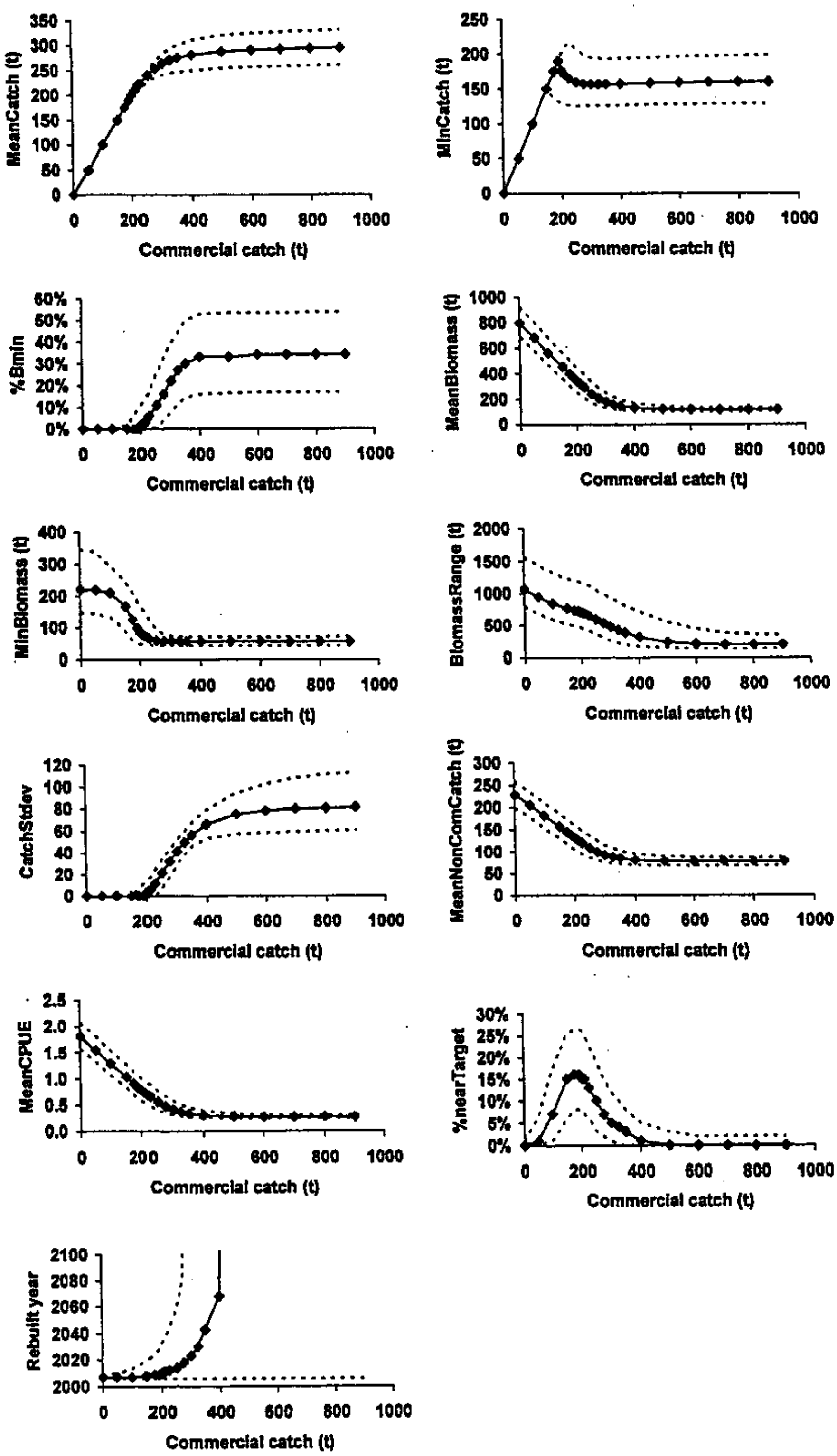

Figure 6: Summaries of posteriors of indicators from exploratory runs of the operating model in which a constant catch was specified (SCC, shown as "Commercial catch" in the x-axis labels) for the entire 100-year run. The figures show the medians of posteriors (central line with diamonds) and the 5th and 95th quantiles. The last figure shows the median of RebuildYr.

3.3 Results from constant exploitation rate

With constant exploitation rates (Figure 7, Table 5), MeanCatch increases at a decreasing rate towards an asymptote of 275 t as U^{target} increases. The %<*Bmin* is zero until specified exploitation rate reaches 1.2, then increases steeply. MeanNCCatch behaves much as it did under the constant catch rule: it is greatest at U^{target}=0, and decreases towards an asymptote near 75 t as U^{target} increases. MeanCPUE is near the target when U^{target} is near 0.6, associated with a mean catch of 200 t. The rule shows peak performance when U^{target} is 0.5 to 0.7, with 25% of years near the target CPUE. As with constant catch, these results also show that there is no simple MSY.

This rule shows some better performance aspects over the constant-catch rule: it delivers substantially higher %nearTarget (25% vs 15%); when fishing levels are higher than optimum in terms of target CPUE, it has far lower %<*Bmin* (the constant catch rule has an asymptote near 34%; this rule was 12% at the maximum value tested).

Table 5: Indicators from exploratory runs of the operating model in which a constant exploitation rate was specified for the entire 100-year run. The table shows median values from the posterior distribution of each indicator except for the percentages for <*Bmin* and CPUE near target, which are means. Median rebuild year of 3000 means that more than the runs did not rebuild by the end of the 100-year run.

U^{target}	Mean catch	Min catch	Mean nonC	%*Bmin*	Mean biomass	Min biomass	Biomass range	Catch stdev	%near Target	Mean CPUE	Rebuild year
0.1	64	23	198	0.0	646	220	880	17	1.0	1.473	2007
0.2	108	45	177	0.0	545	218	750	29	5.1	1.241	2007
0.3	141	65	161	0.0	471	212	653	39	14.1	1.071	2007
0.4	165	80	149	0.0	415	200	582	47	22.2	0.945	2007
0.5	185	91	139	0.0	372	183	531	54	26.3	0.846	2007
0.6	201	99	131	0.0	337	166	489	60	26.3	0.766	2008
0.7	215	104	125	0.0	308	150	457	66	24.2	0.701	2009
0.8	226	107	119	0.0	284	135	434	71	20.2	0.645	2010
0.9	236	110	114	0.0	263	123	414	76	16.2	0.598	2012
1.1	251	111	105	0.0	228	101	382	85	10.1	0.520	2016
1.2	257	110	101	1.0	214	92	367	88	8.1	0.488	2019
1.3	262	111	98	3.0	201	85	355	91	6.1	0.459	2022
1.4	266	112	95	4.0	190	80	342	94	5.1	0.434	2026
1.5	269	117	93	6.1	181	78	329	96	4.0	0.411	2031
1.6	271	121	90	7.1	172	76	315	96	3.0	0.392	2039
1.7	273	126	88	9.1	165	74	303	96	3.0	0.375	2047
1.8	274	131	87	10.1	159	73	292	95	2.0	0.361	2057
1.9	276	135	85	12.1	153	71	283	93	2.0	0.348	2071

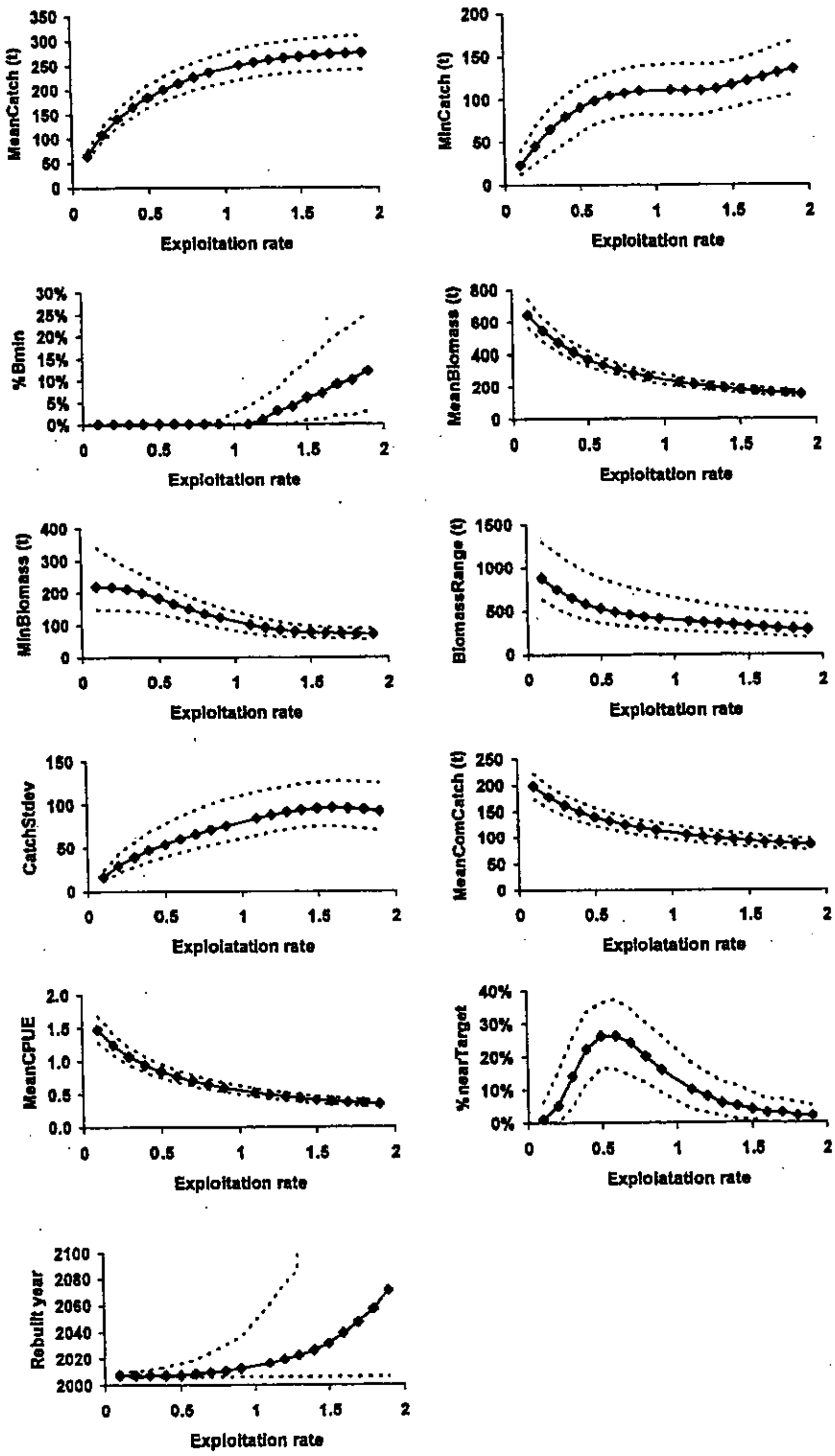

Figure 7: Summaries of posteriors of indicators from exploratory runs of the operating model in which a constant exploitation rate was specified for the entire 100-year run. The figures show the medians of posteriors (central line with diamonds) and the 5th and 95th quantiles.

3.4 Effect of non-commercial exploitation

Posteriors of non-commercial exploitation rates are shown in Figure 8. The largest component is the illegal catch, with the median of $U_{AW}^{illegal}$ near 7%. $U_{SS}^{illegal}$ and U_{SS}^{rec} are similar in scale, but the illegal catch is taken from both the SL and NSL biomass components.

Although these rates seem small, they are significant to the operating model production results. With no commercial catch, biomass increases to reach equilibrium with these rates, resulting in average annual non-commercial catches near 230 t. However, productivity is not simply the sum of the commercial and non-commercial catches, because these come from different segments of the population.

To explore the effect of non-commercial catches on commercial catch productivity, a special set of exploratory runs was made by turning off the non-commercial catch dynamics in the operating model. The rule used was the constant specified commercial catch (SCC) rule, so the results (Figure 9, Table 6) should be compared with Figure 6 and Table 4.

Without non-commercial catch, the asymptotic MeanCatch is 340 t, vs 300 t with non-commercial catch. However, this comparison is somewhat misleading because non-commercial catch is least when SCC is highest.

MeanBiomass is 1500 t when SCC is zero: this is nearly twice the value seen with the non-commercial catches assumed by the operating model. With very high SCC, %<*Bmin* reaches 23%, considerably less than the 34% seen in the base case explorations. MeanCPUE is near the target value at much higher SCC (325 t), with MeanCatch averaging 311 t compared with 200 t in the base case. The last comparison shows the true effect of the operating model's non-commercial catch on the potential commercial catch: without the non-commercial catch the commercial fishery would take about 100 t more while maintaining the average target CPUE.

In this trial, %nearTarget was low (11%) compared with the base case (16%). This is probably because of the higher variability induced when the entire catch is governed by the SCC: in the base case a substantial part of the catch is governed by an exploitation rate, and this constant-rate dynamics has a stabilising effect.

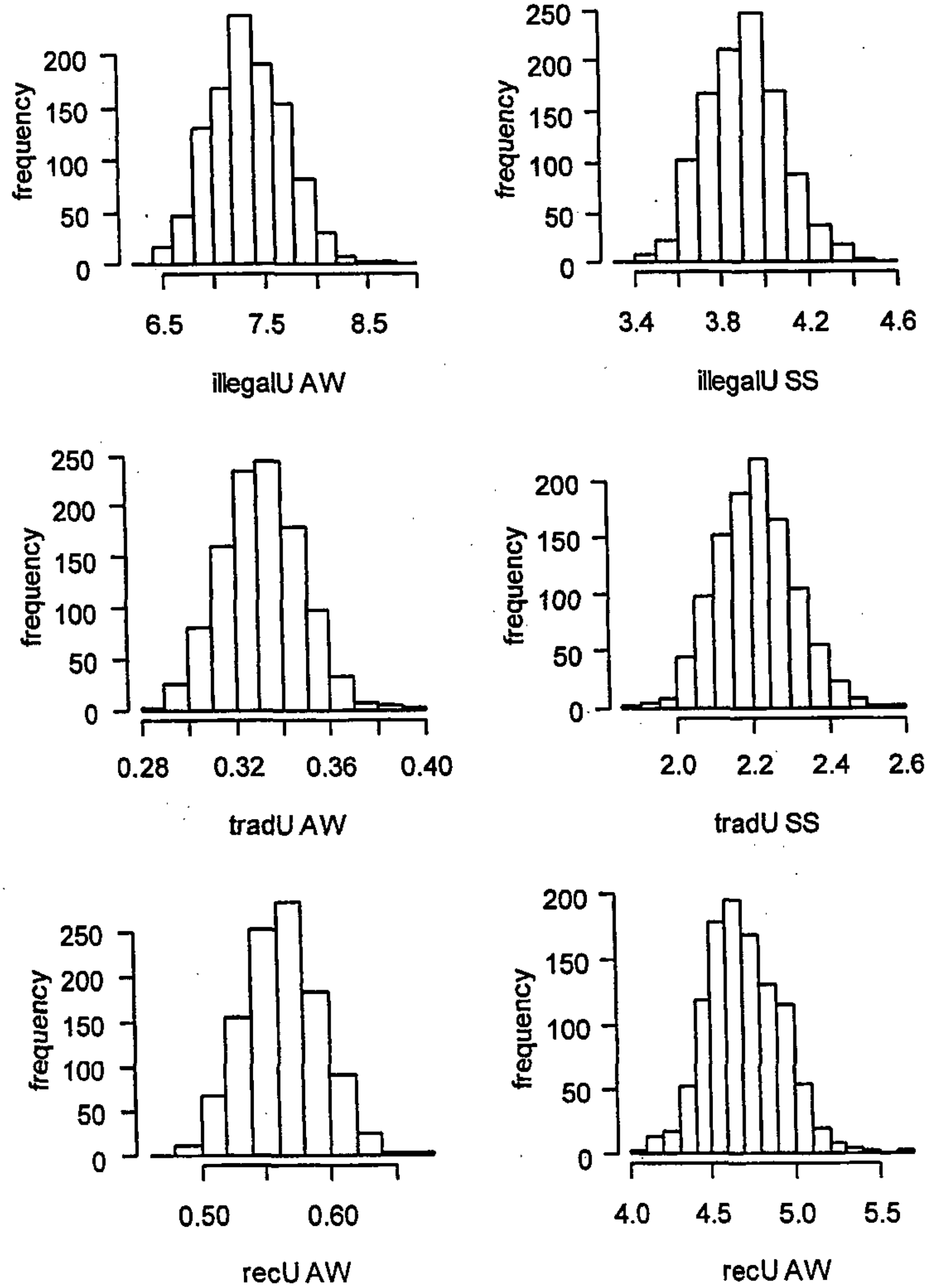

Figure 8: Posterior distributions of the illegal (top), customary (middle) and recreational exploitation rates (in percentage) for the AW (left) and SS seasons.

Table 6: Indicators from exploratory runs of the operating model in which non-commercial catch was zero and a constant catch was specified (SCC) for the entire 100-year run. The table shows median values from the posterior distribution of each indicator except for the percentages for <*Bmin* and CPUE near target, which are means.

SCC	Mean catch	Min catch	Mean nonC	%*Bmin*	Mean biomass	Min biomass	Biomass range	Catch stdev	%near Target	Mean CPUE	Rebuild year
0	0	0	0	0.0%	1512	231	2 021	0	0.0%	3.440	2006
50	50	50	0	0.0%	1344	231	1 841	0	0.0%	3.053	2006
100	100	100	0	0.0%	1168	231	1 654	0	0.0%	2.652	2006
150	150	150	0	0.0	979	230	1 455	0	0.0	2.231	2006
175	175	175	0	0.0	883	228	1 352	0	1.0	2.009	2006
190	190	190	0	0.0	824	223	1 292	0	2.0	1.873	2006
200	200	200	0	0.0	782	219	1 257	0	2.0	1.783	2006
210	210	210	0	0.0	743	212	1 224	0	3.0	1.691	2006
225	225	216	0	0.0	682	198	1 180	0	5.1	1.552	2006
250	250	216	0	0.0	583	156	1 110	0	8.1	1.330	2006
275	274	216	0	1.0	488	94	1 039	6	10.1	1.112	2006
300	295	203	0	4.0	402	72	954	19	11.1	0.910	2006
325	311	192	0	8.1	326	67	855	32	10.1	0.739	2007
350	322	186	0	12.1	267	65	750	45	9.1	0.605	2008
400	334	183	0	19.2	195	64	577	65	5.1	0.445	2012
500	339	182	0	22.2	145	63	370	87	2.0	0.331	2036
600	340	181	0	23.2	135	63	275	94	1.0	0.307	2095
700	340	181	0	23.2	132	63	243	97	0.0	0.301	3000
800	340	181	0	23.2	132	63	232	97	0.0	0.299	3000
900	340	181	0	23.2	131	63	229	98	0.0	0.299	3000

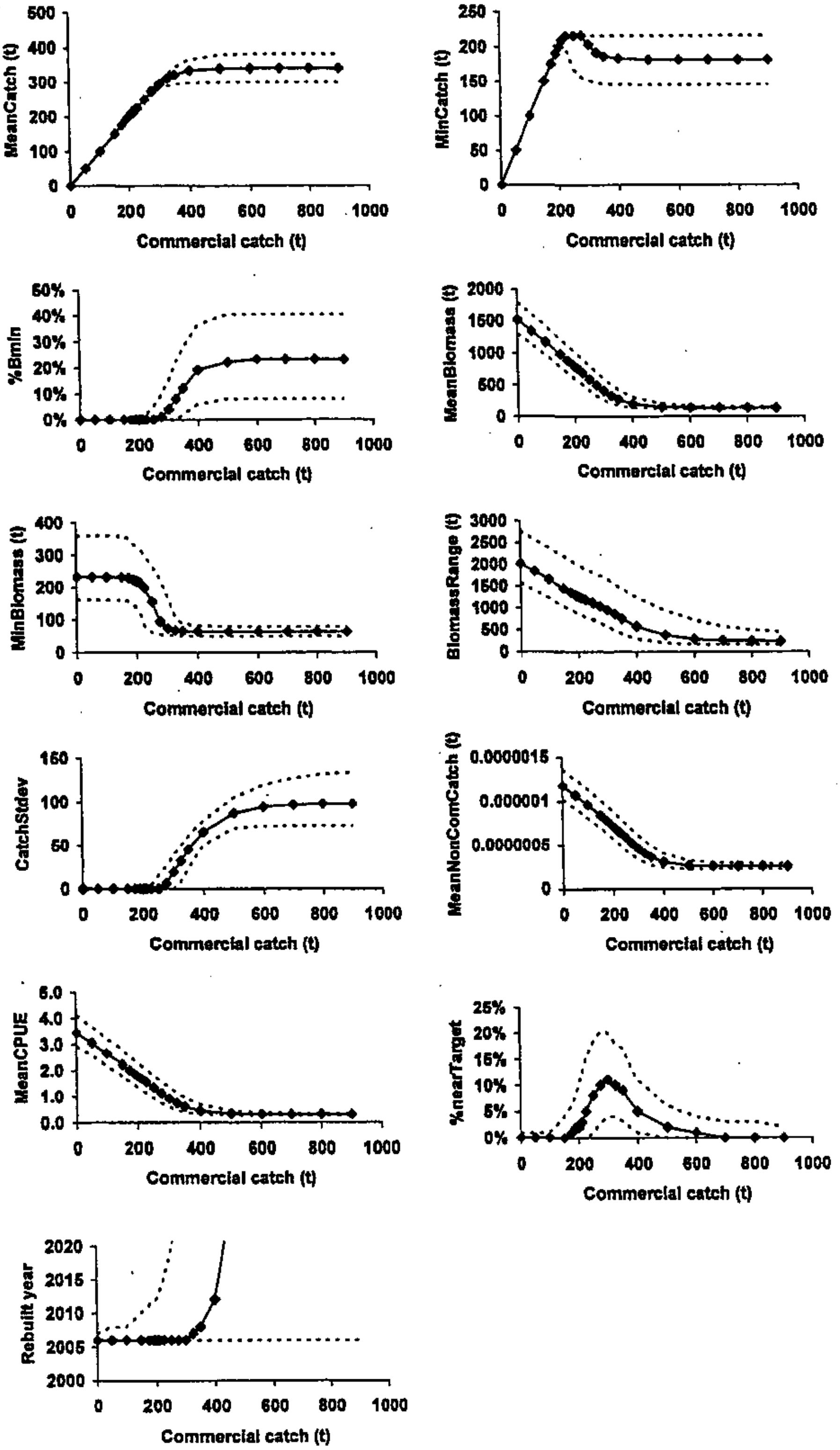

Figure 9: Summaries of posteriors of indicators from exploratory runs of the operating model with different values for specified annual commercial catch (SCC) with zero non–commercial catch. The figures show the medians of posteriors (central line with diamonds) and the 5th and 95th quantiles.

4. HARVEST CONTROL RULES

We explored four rule families. Preliminary evaluations used families that we call the Bentley, Breen and Kim families. Final evaluations discarded the Kim family and included the "modified BK" family. In this section we describe the rules and show some examples of their performance.

4.1 Bentley rule

What we call the Bentley rule is the basis of the current CRA 7 and CRA 8 (NSS) management procedure (Bentley et al. 2003b). It has a CPUE target and a "target trajectory" of CPUE, which is a straight line from the CPUE observed in 1997 to the target (Figure 10).

The rule compares the current CPUE with the target CPUE and compares the slope of the CPUE trajectory with the slope of the target trajectory, then averages and combines the comparisons to obtain a multiplier that determines the new specified commercial catch from the existing specified catch:

Eq 12 $SCC_{y+1} = Z_y SCC_y$

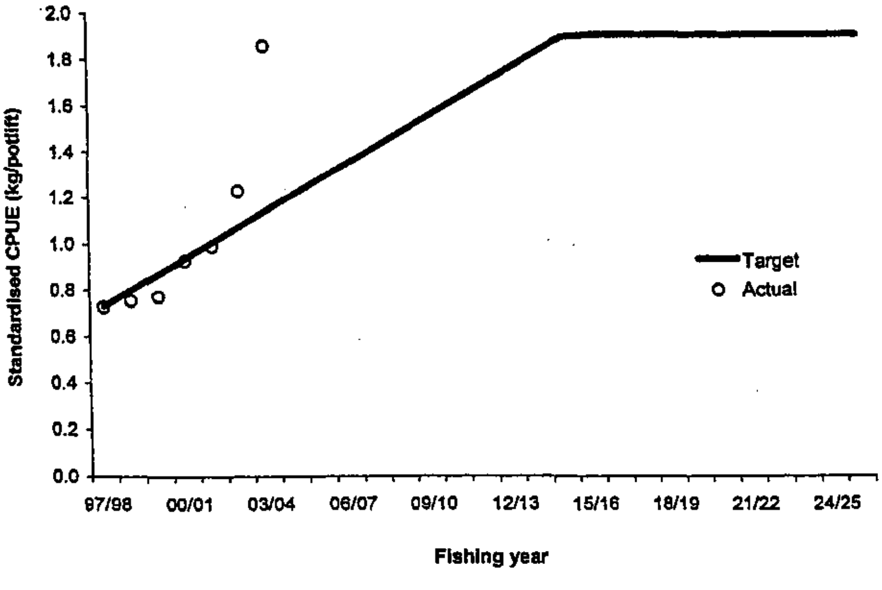

Figure 10: Showing the NSS management procedure's target trajectory for CPUE (bold line). Open circles show annual CPUE estimates.

In what follows, some parameters are denoted by the same letter as quantities described for the assessment model. To prevent confusion, harvest rule quantities with the same symbol as a model variable are denoted with a prime.

The Bentley rule has seven parameters:

- the target CPUE (treated as a constant 0.75),
- the target rebuilding year,

- N', the number of years used for averaging observed CPUE,
- W', the weight given to the distance between observed and target CPUE, relative to the difference between observed and target gradients,
- S', a sensitivity parameter used to determine the rule's response,
- *Min*, the minimum multiplier Z_y allowed, and
- *Max*, the maximum multiplier Z_y allowed.

These parameters define a family of candidate harvest control rules. For each year the target CPUE, I_y^{target} is calculated from the straight line defined by the final target, the target rebuild year, the initial CPUE (in AW 2004) and the initial year (2004), as was done for CRA 8 (Figure 10). For this study was assumed that the target rebuild year is 2012. In the control rule, the difference between the target and the observed CPUE is calculated for a "status indicator":

Eq 13 $A_y^s = \breve{I}_y^{obs} / I_y^{target} - 1$

Similarly, the difference between the target and observed gradient is calculated for a "gradient indicator:

Eq 14 $A_y^g = \left[\left(\breve{I}_y^{obs} - \breve{I}_{y-1}^{obs}\right) / \breve{I}_{y-1}^{obs}\right] - \left[\left(I_y^{target} - I_{y-1}^{target}\right) / I_{y-1}^{target}\right]$

The status indicator is averaged for N' years:

Eq 15 $\bar{A}_y^s = \frac{1}{N'} \sum_{d=y-N'+1}^{d=y} A_d^s$

and similarly for A_t^g to obtain $\bar{A}_t^g$. Then the mean gradient and status indicators are combined, using the relative weight W':

Eq 16 $A_y^* = W'\bar{A}_y^s + (1-W')\bar{A}_y^g$

and the combined mean indicator is used with the sensitivity parameter S' to determine a response:

Eq 17 $R_y' = S'A_y^*$

Then this response is used to determine the multiplier Z_t, taking into account the sign of R_y' and using the *Min* and *Max* threshold parameters.

$$\begin{array}{ll}
Z_y = 1 & \text{for } -Min \le R_y' \le Min \\
Z_y = 1 + R_y' & \text{for } -Max \le R_y' \le Min \\
 & \text{and for } Min < R_y' \le Max \\
Z_y = 1 - Max & \text{for } R_y' < (-Max) \\
Z_y = 1 + Max & \text{for } R_y' > Max
\end{array}$$

A "latent year" is operative, prohibiting changes to the specified catch in two consecutive years.

In the results shown in this section, runs were made from the thinned base case Markov chain - Monte Carlo (McMC) results from the CRA 3 assessment using the quantities shown in Table 7. Four parameters of the rule were varied through two to five values each (values were chosen after some very simple explorations not shown), and we ran every combination in 90 sets of runs.

The indicators (Figure 11) show small to moderate contrast among the rules, except for Nchanges, CatchStdDev and MinCatch, all showing substantial variation among rules. Correlations among the 90 average indicators are very high (Table 8), suggesting that selection of a rule could be based on a smaller set of indicators. For instance, mean non-commercial catch is almost perfectly negatively correlated with mean non-commercial catch and mean CPUE, and strongly correlated with minimum catch.

The median of RebuildYr was 2008 or 2009. MeanCatch ranged from 148 to 200 t. The %<*Bmin* was under 4% for all rules. MeanCPUE ranged from 0.83 to 1.02; %nearTarget ranged from 13 to 17% and Nchanges ranged from 5 to 12.

Table 7: Values for projection variables used for the exploratory runs of the operating model for the Bentley rule.

Variable	Value
Projections to	2025
MLS male AW	52
MLS male SS	54
MLS female AW	100
MLS female SS	60
ρ	0.30
$\sigma^{\bar{\varepsilon}}$	0.00
σ^{Iproj}	0.10
Final target	0.75
Target year	2012
N'	1, 2 and 3
W'	0.2, 0.4 and 0.6
S'	0.50, 1, 1.5, 1.75, and 2
Min	0.05 and 0.10

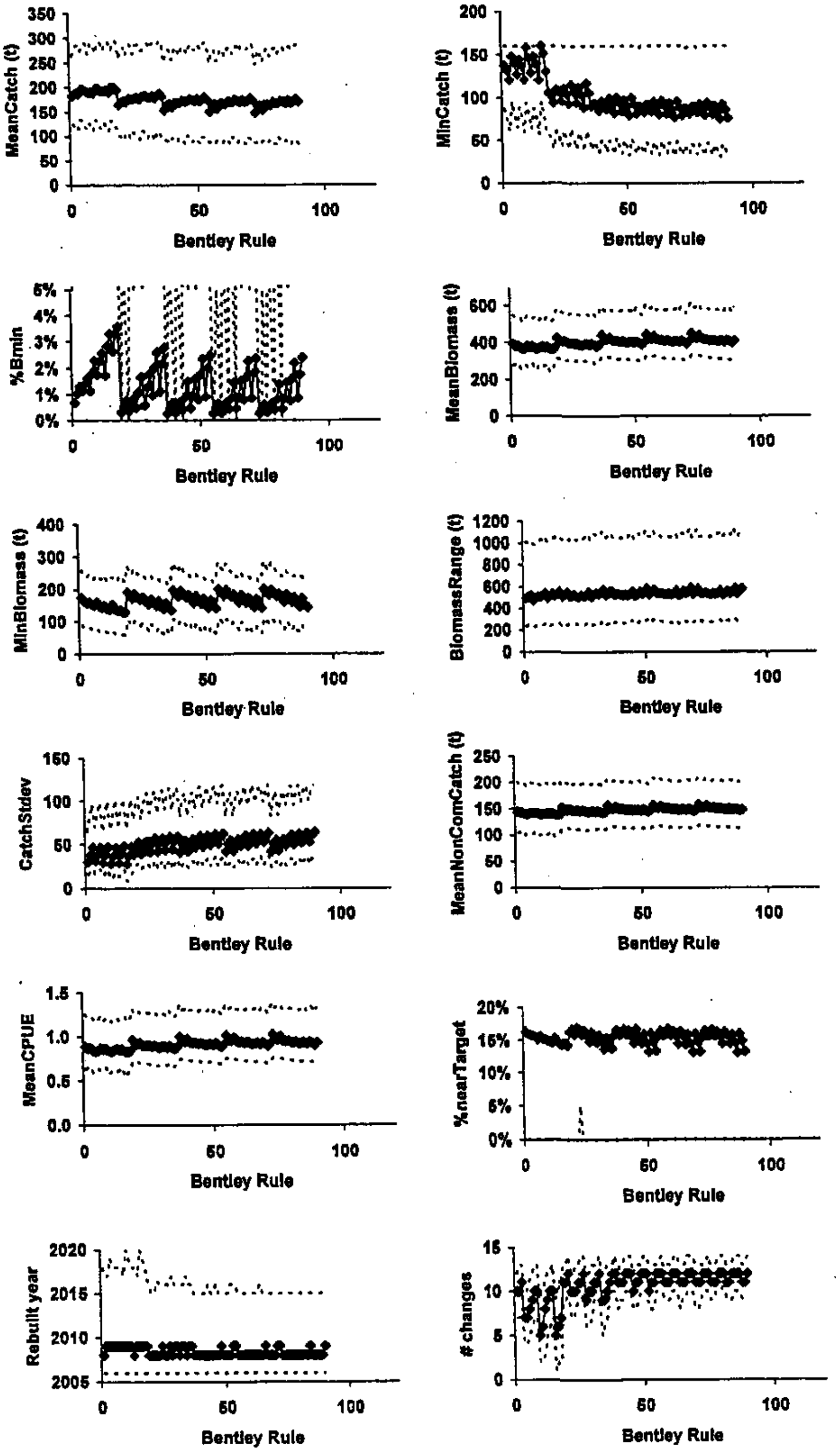

Figure 11: Indicator values from each of the 90 versions of the Bentley rule tested with the base case operating model. Diamonds show the median of the indicator (except for % < *Bmin* and %nearTarget, which are the means) and the dashed lines show the 5th and 95th percentiles.

Table 8: Correlations among the medians (for % < *Bmin* and %nearTarget, the mean) indicators from exploratory runs of Bentley rule versions with the base case operating model. Shading indicates an absolute value greater than 0.5 and boxes have been used for negative correlations less than -0.50.

	Mean catch	Min catch	Non-comm catch	% < *Bmin*	Mean biomass	Min biomass	Biomass range	Catch std dev	% near target	Mean CPUE	N changes
MeanCatch	1.00										
MinCatch	-0.81	1.00									
MeanNCCatch	0.99	-0.76	1.00								
%<*Bmin*	0.67	0.30	0.69	1.00							
MeanBiomass	-0.99	0.75	-1.00	-0.67	1.00						
Min biomass	-0.78	-0.34	-0.80	-0.96	0.80	1.00					
BiomassRange	-0.30	-0.68	0.26	0.45	0.27	-0.35	1.00				
CatchStdev	-0.38	-0.85	0.32	0.15	0.30	-0.16	0.87	1.00			
%nearTarget	-0.29	0.11	0.31	-0.81	0.29	0.70	-0.77	-0.44	1.00		
MeanCPUE	-0.99	0.74	-1.00	-0.66	1.00	0.77	0.28	0.30	0.28	1.00	
Nchanges	-0.78	-0.86	0.75	-0.41	-0.70	0.44	0.52	0.64	0.07	-0.74	1.00

A crude examination of average indicators calculated across the various rule parameter values (Table 9) showed a relative insensitivity to N', W' and *Min*, and high sensitivity to S'. Higher S' implied lower and more variable catches, more catch changes, higher CPUE and better rebuild performance. Higher values of *Min* implied slightly higher catches.

Table 9: Mean values, averaged across the rule parameter values indicated, of the medians (for % < *Bmin* and % near target, the mean) indicators from runs of the Bentley rule versions with the base case operating model. In this table, "%rebuilt" is the percentage of runs in which biomass is rebuilt in 2011.

	Mean Catch	Min Catch	Mean NCCatch	%<*Bmin*	Catch Stdev	Nchanges	Mean CPUE	%near Target	% Rebuilt
N'									
1	169.1	100.2	149.4	0.6	46.8	10.77	0.934	16.0	76.5
2	176.6	102.1	146.2	1.2	50.5	10.60	0.903	15.4	75.4
3	179.0	100.8	145.1	1.9	52.8	10.53	0.894	14.6	74.8
W'									
0.2	172.0	108.4	148.2	0.7	42.3	10.33	0.923	15.9	76.4
0.4	177.5	102.8	146.0	1.3	51.1	10.60	0.900	15.5	75.3
0.6	175.1	92.0	146.6	1.7	56.6	10.97	0.908	14.6	74.9
S'									
0.50	192.5	138.2	140.4	2.0	38.0	8.22	0.849	15.1	68.2
1.00	177.8	103.0	145.8	1.2	50.0	10.50	0.897	15.5	74.2
1.50	170.3	91.0	148.6	1.1	53.4	11.33	0.926	15.4	77.5
1.75	167.9	87.6	149.5	1.0	54.1	11.50	0.936	15.3	78.5
2.00	165.8	85.4	150.3	1.0	54.7	11.61	0.944	15.3	79.4
Min									
0.05	172.7	99.2	147.8	1.2	49.7	11.42	0.919	15.4	76.3
0.10	177.0	102.9	146.0	1.3	50.3	9.84	0.901	15.3	74.8

Four typical TACC and CPUE trajectories from the Bentley rule, chosen randomly from the first 20 runs of the base case operating model, are shown in Figure 12. The first set shows some lag in responding to a sharp drop in CPUE. The second shows a much better response to CPUE and a good performance; the fourth also shows a very good performance.

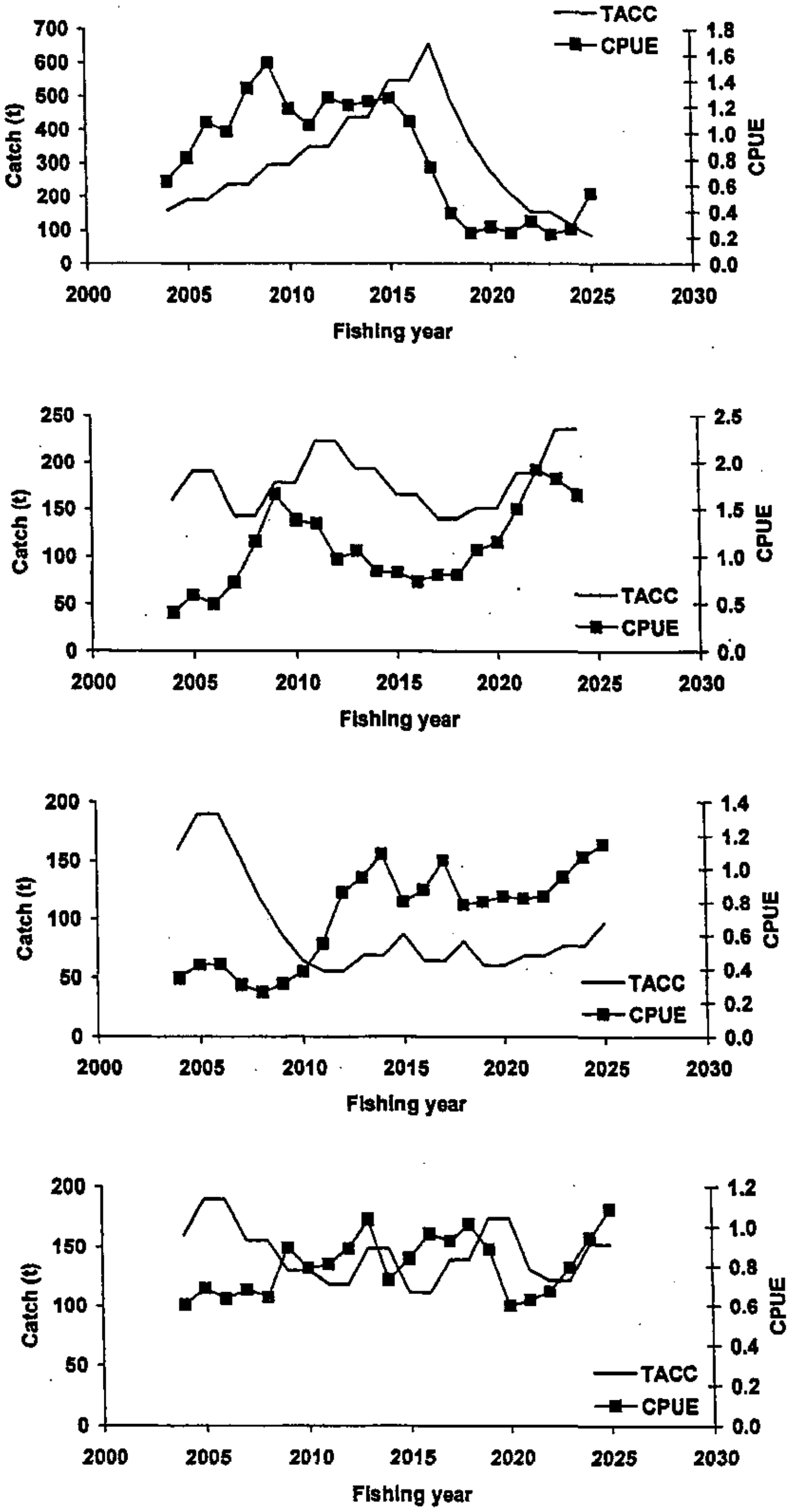

Figure 12: Four typical catch and CPUE trajectories from the Bentley rule under the base case operating model. The four panels show the 5th, 9th, 11th and 13th runs respectively (these were chosen randomly from the first 20 runs). In this version of the rule, later called P1271, $N'=1$, $W'=0.4$, $S'=2.0$ and $Min=0.05$.

4.2 Breen rule

The Breen rule is a simple constant-rate rule. The concept is that some information is available from the stock assessment about the relation between mean CPUE and mean catch (discussed above in Section 3.3). The estimate of mean catch associated with the target CPUE defines a proportional relation that can be used in a harvest rule.

The Breen rule has four parameters:

- target CPUE, I^{target} (treated as a constant 0.75),
- a level of catch, T', that is assumed to be associated with the target CPUE,
- a sensitivity parameter, S' and
- a power term, p.

These parameters define a family of candidate harvest control rules. For each year, the rule's suggested catch, SCC', is calculated from T' and the ratio between observed and target CPUE in the previous year:

Eq 18 $$SCC'_{y+1} = T'\left(\frac{\tilde{I}_y^{obs}}{I^{target}}\right)^p$$

Essentially, a target exploitation rate is defined by T' and I^{target}. The difference between the rule output, SCC'_{y+1}, and the current catch limit, SCC_y, is compared with the sensitivity parameter S' to decide if the catch limit should be changed:

$$SCC_{y+1} = SCC_y \qquad \text{for } \left|SCC'_{y+1} - SCC_y\right| < S'$$
$$SCC_{y+1} = SCC'_{y+1} \qquad \text{for } \left|SCC'_{y+1} - SCC_y\right| \geq S'$$

Thus, if S' were 20 t, no change less than 20 t would be made.

Unlike more complex rules, this simple rule can be shown graphically with SCC' as a function of CPUE (Figure 13 and Figure 14).

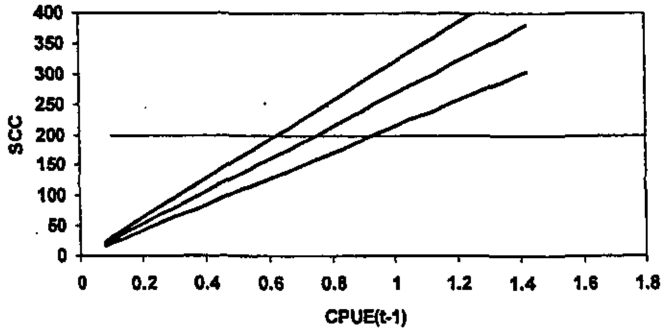

Figure 13: Three members of the Breen rule family, all with $p = 1$. From top to bottom, T' = 240, 200 and 160 t.

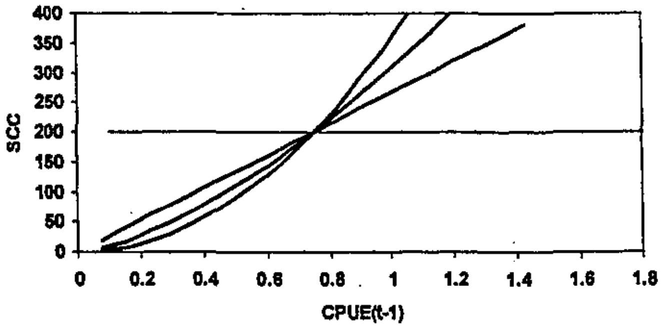

Figure 14: Three members of the Breen rule family, all with T' = 200 and with p = 2.0, 1.5 and 1.0.

Exploratory runs were made from the thinned base case McMC results from the CRA 3 assessment using the quantities shown in Table 10. In the results shown in this section four levels of T' parameter, five of S' and three of p were used, for 60 different runs.

The indicators (Figure 15) show limited contrast among the rules except for Nchanges, ranging from 7 to 17, and the variability of catch indicators: . Most of the rules resulted in median RebuildYr 2007 to 2009. MeanCatch ranged from 195 to 224 t. The %<*Bmin* was under 1% for all rules. MeanCPUE ranged from 0.71 to 0.83. The %nearTarget ranged from 20 to 25%.

Table 10: Values for projection variables used for the exploratory runs of the operating model for the Breen rule.

Variable	Value
Projections to	2025
MLS male AW	52
MLS male SS	54
MLS female AW	100
MLS female SS	60
ρ	0.30
$\sigma^{\bar{\varepsilon}}$	0.00
σ^{Iproj}	0.10
I^{target}	0.75
T'	180, 200, 220, and 240
S'	15, 25, 35, 45 and 55 t
þ	1.0, 1.2 and 1.4

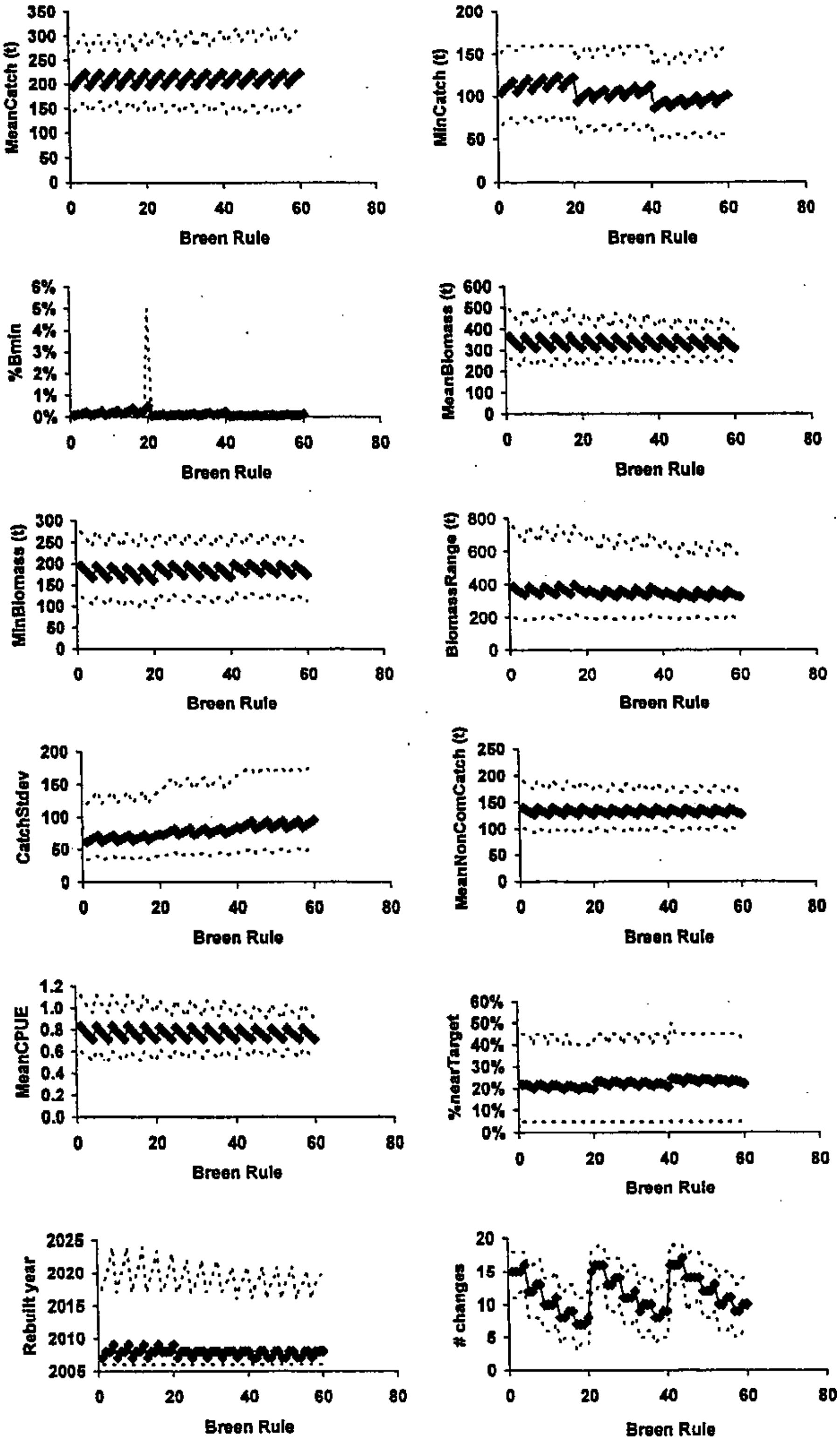

Figure 15: Indicator values from each of the 60 versions of the Breen rule tested with the base case operating model. Diamonds show the median of the indicator (except for % < *Bmin* and % near target, which are the means) and the dashed lines show the 5th and 95th percentiles.

A crude examination of average indicators versus the rule parameter values (Table 11) showed that MeanCatch was related to T'; MeanCPUE and the percentage of runs rebuilt by 2011 were inversely related to T'. Other results were not sensitive to this rule parameter. Most results except Nchanges were insensitive to S'; Nchanges decreased as this parameter increased. Higher values of the power parameter p gave better performance in nearly all indicators except the MinCatch and CatchStdev.

Table 11: Mean values, averaged across the rule parameter values indicated, of the medians (for % < *Bmin* and % near target, the mean) indicators from runs of the Breen rule versions base case operating model.

T'	Mean catch	Min Catch	Mean NCCatch	%<*Bmin*	Catch Stdev	Nchanges	Mean CPUE	%near Target	% Rebuil
180	197.8	97.7	138.2	0.1	72.8	11.27	0.817	22.8	81.
200	207.1	102.4	134.1	0.1	76.3	11.40	0.776	22.8	77.
220	215.6	105.9	130.5	0.1	79.9	11.80	0.740	22.1	72.
240	223.1	109.2	127.1	0.2	83.2	12.20	0.707	21.2	68.
S'									
15	210.5	100.4	132.5	0.1	77.1	15.75	0.760	22.5	74.
25	210.6	102.3	132.4	0.1	77.4	13.33	0.760	22.5	74.
35	210.9	103.8	132.5	0.1	77.8	11.25	0.760	22.3	75.
45	211.1	105.8	132.6	0.1	78.4	9.58	0.760	22.1	75.
55	211.3	106.9	132.5	0.2	79.4	8.42	0.759	21.7	75.
p									
1.0	209.9	114.5	133.0	0.2	67.7	10.75	0.766	20.9	73.
1.2	211.0	103.4	132.4	0.1	78.1	11.75	0.760	22.3	75.
1.4	211.8	93.6	132.0	0.1	88.4	12.50	0.754	23.6	76.

Typical TACC and CPUE trajectories from the Breen rule under the base case operating model are shown in Figure 16. The rule is closely responsive to CPUE, although some unavoidable lag is evident.

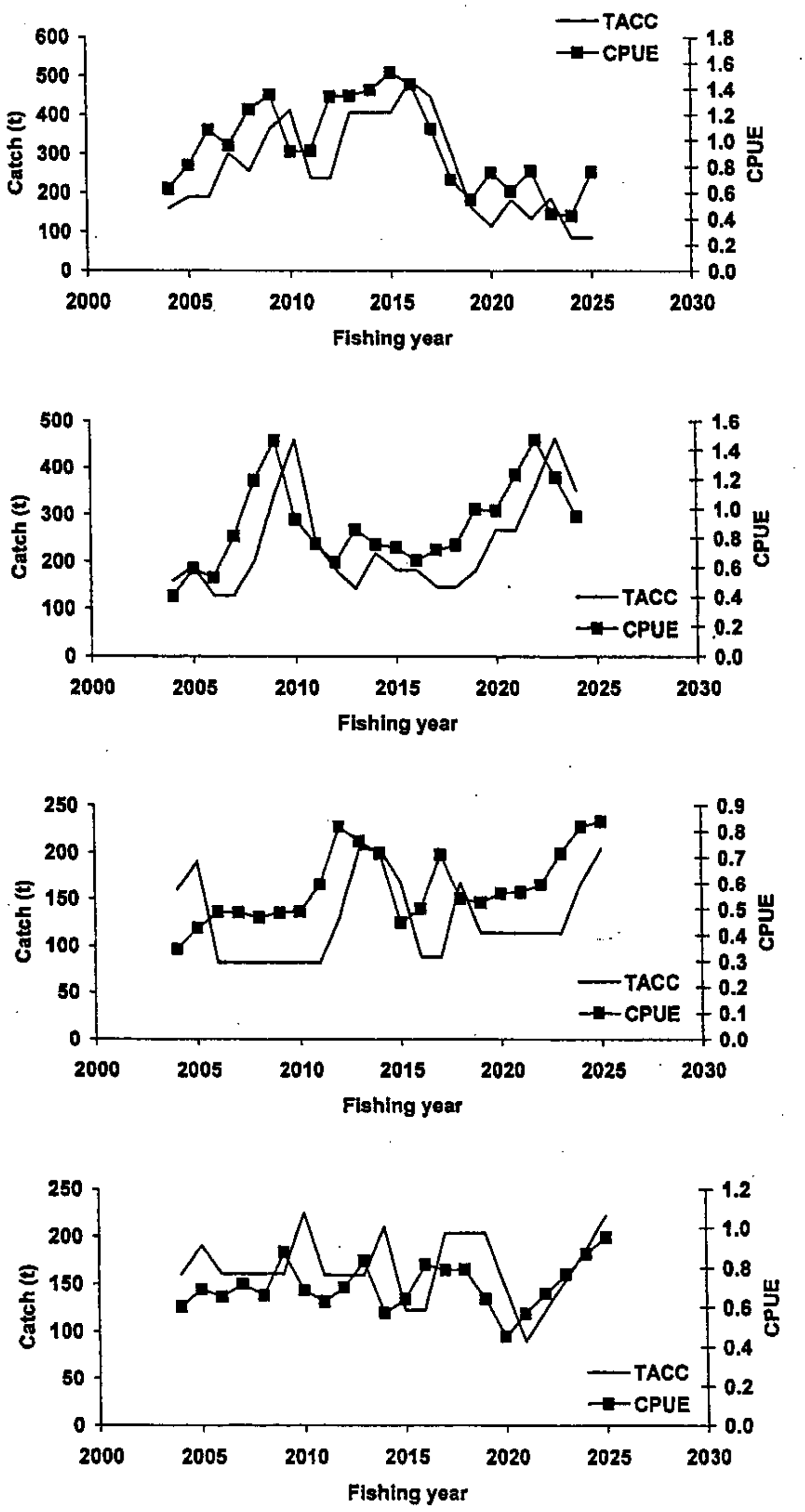

Figure 16: Typical catch and CPUE trajectory from the Breen rule under the base case. The panels show the 5th, 9th, 11th and 13th runs, as for the Bentley rule (Figure 12). In this rule, later called P123, T'=180, S'=15 and p=1.4.

4.3 Kim rule

The Kim rule has a target CPUE and a maximum SCC. The rule compares two consecutive CPUE observations and produces the SCC from their ratio and the previous SCC. The new SCC is then capped by the maximum SCC.

In what follows, some parameters are denoted by the same letter as quantities described for the assessment model and, to prevent confusion, harvest rule quantities with the same symbol as a model variable are denoted with a prime.

The Kim rule has two parameters:

- a target CPUE, I'^{target}, considered a constant 0.75 in this study and
- maximum allowable catch, $T^{\max}$

These parameters define a family of candidate harvest control rules. For each year, the new catch limit suggested by the rule, SCC'_{y+1}, is calculated from the ratio between the two most recent years of CPUE and the current catch limit:

Eq 19 $$SCC'_{y+1} = \min\left(\frac{\breve{I}_y^{obs}}{\breve{I}_{y-1}^{obs}} SCC_y, T^{\max}\right)$$

The specified catch, SCC, is capped at $T^{\max}$:

$SCC_{y+1} = SCC'_{y+1}$	if $\breve{I}_y^{obs} \geq I'^{target}$ and either there was increase in CPUE or there were two consecutive decreases in CPUE, or if $\breve{I}_y^{obs} < I'^{target}$ and either there was decrease in CPUE or there were two consecutive increase in CPUE
$SCC_{y+1} = SCC_y$	if conditions above are not satisfied.

Exploratory runs were made from the thinned base case McMC from the CRA 3 assessment using the quantities shown in Table 12. The parameter $T^{\max}$ was varied through 12 levels.

Table 12: Values for projection variables used for the Kim rule with base case operating model.

Variable	Value
Projections to	2025
MLS male AW	52
MLS male SS	54
MLS female AW	100
MLS female SS	60
ρ	0.30
$\sigma^{\bar{\varepsilon}}$	0.00
σ^{Iproj}	0.10
I^{target}	0.75
T^{max}	160 to 270 t in 12 steps

The indicators (Figure 17) show relatively small contrast among the rules, except for the CatchStdDev. The median of RebuildYr is 2008 or 2009. Mean catch ranged from 147 to 201 t. The %<*Bmin* was under 2% for all rules. Mean CPUE ranged from 0.82 to 1.04. The %nearTarget ranged from 13 to 15%. Nchanges ranged from 7 to 8.

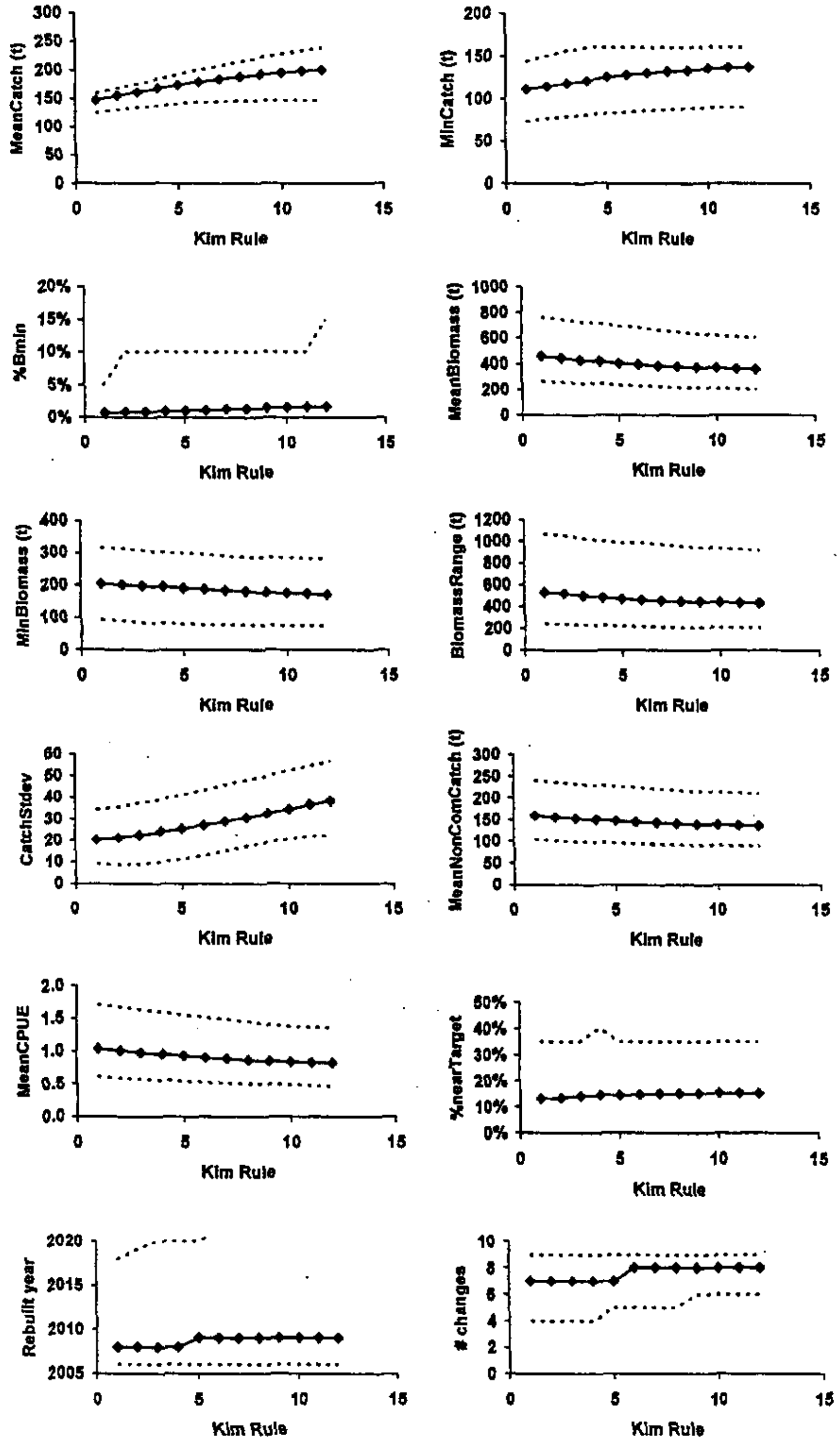

Figure 17: Indicator values from each of the 12 versions of the Kim rule tested with the base case operating model. Diamonds show the median of the indicator (except for % < *Bmin* and %nearTarget, which are the means) and the dashed lines show the 5th and 95th percentiles.

An examination of indicators against the rule parameter values (Table 13) showed that higher T^{max} implied higher and more variable catches, more catch changes but lower CPUE.

Table 13: Mean values, averaged across the rule parameter values indicated, of the medians (for % < *Bmin* and % near target, the mean) indicators from exploratory runs of the Kim rule versions with the high catch in 2004 and 2005 operating model.

T^{max}	Mean Catch	Min Catch	Mean NCCatch	%<*Bmin*	Catch Stdev	Nchanges	Mean CPUE	%near Target	% Rebuilt
160	147	111	159	0.7	20	7	1.04	13.2	72.4
170	155	114	156	0.8	21	7	1.01	13.4	70.5
180	161	118	153	0.9	22	7	0.97	14.1	68.8
190	167	121	150	0.9	24	7	0.95	14.6	67.6
200	174	125	148	1.0	25	7	0.93	14.6	66.3
210	179	128	145	1.1	27	8	0.90	14.8	65.3
220	185	130	143	1.3	29	8	0.88	15.0	65.3
230	189	132	141	1.3	31	8	0.87	15.1	64.7
240	192	133	140	1.4	33	8	0.85	15.2	64.4
250	195	135	138	1.5	34	8	0.84	15.3	64.1
260	198	136	137	1.6	37	8	0.83	15.3	63.9
270	201	137	136	1.7	38	8	0.82	15.3	63.6

Typical TACC and CPUE trajectories from the Kim rule under the base case operating model are shown in Figure 18.

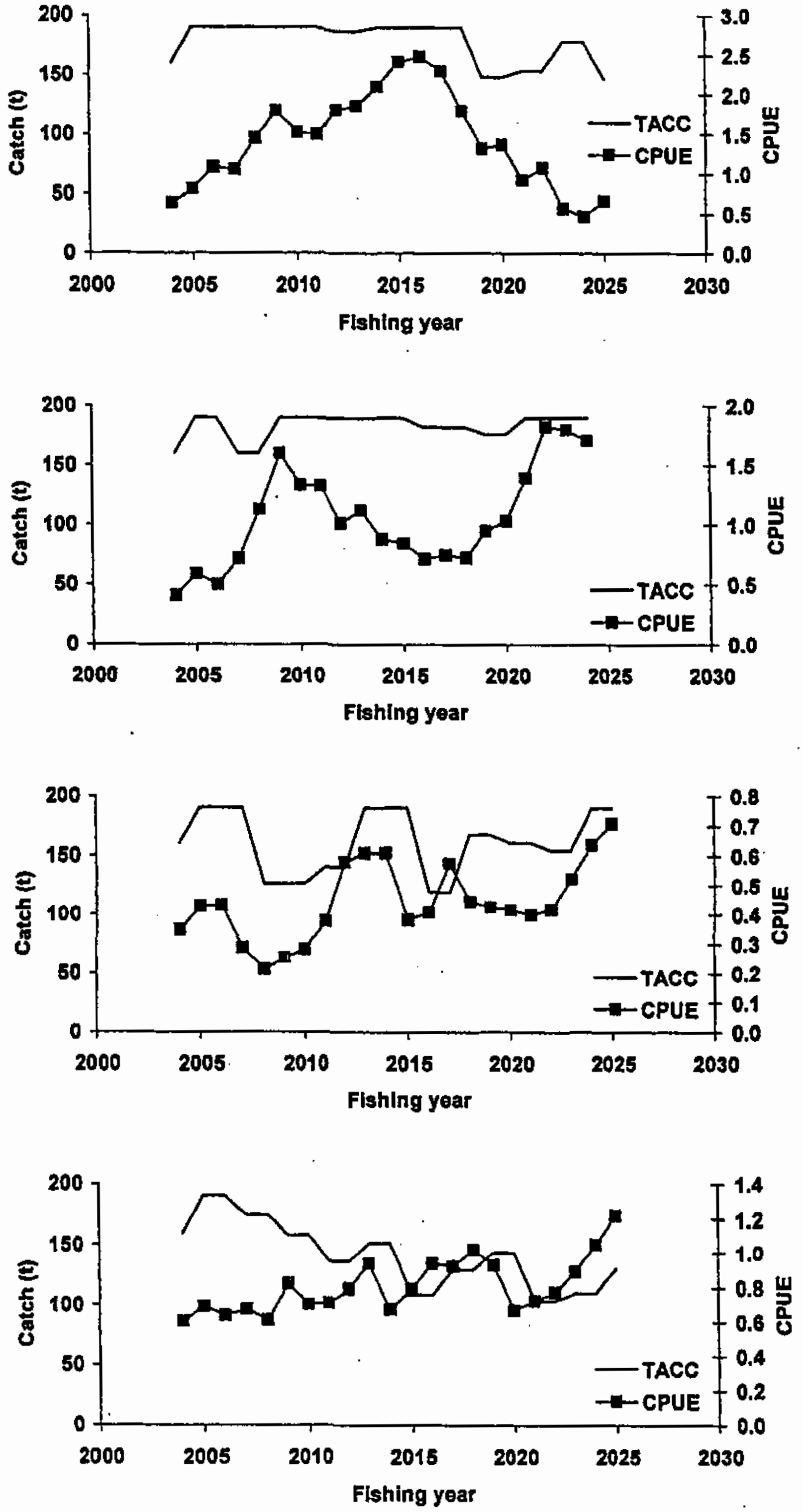

Figure 18: Typical catch and CPUE trajectory from the Kim rule under the base case. The panels show the same runs as for the Bentley and Breen rules (Figure 12 and Figure 16).

4.4 Modified BK rule

This rule is modified from the Breen rule. The absolute sensitivity parameter S' in the Breen rule is replaced by a sensitivity parameter that limits the relative change from the current SCC. As for the Breen rule, the modified BK rule has parameters I^{target}, T', a sensitivity parameter, S'' and a power term p. These parameters define a family of candidate harvest control rules.

For each year, the rule's suggested catch, SCC', is calculated from the target catch and the ratio between observed CPUE in the previous year:

$$\text{Eq 20} \quad SCC'_{y+1} = T'\left(\frac{\breve{I}_y^{obs}}{I^{target}}\right)^p$$

The relative difference between the rule output, SCC'_{y+1}, and the current catch limit, SCC_y, is compared with the sensitivity parameter S'', and a minimum change level of 0.05, to determine the new SCC:

$$SCC_{y+1} = (1+S'')SCC_y \quad \text{for } 1-SCC'_{y+1}/SCC_y > S''$$

$$SCC_{y+1} = SCC'_{y+1} \quad \text{for } 0.5 < 1-SCC'_{y+1}/SCC_y < S''$$

$$SCC_{y+1} = SCC_y \quad \text{for } -0.5 < 1-SCC_{y+1}/SCC_y < 0.5$$

$$SCC_{y+1} = SCC'_{y+1} \quad \text{for } -S'' < 1-SCC'_{y+1}/SCC_y < -0.5$$

$$SCC_{y+1} = (1-S'')SCC_y \quad \text{for } 1-SCC'_{y+1}/SCC_y < -S''$$

Exploratory runs were made from the thinned base case McMC results from the CRA 3 assessment using the quantities shown in Table 14. Four levels of T' parameter, five of S' and three of p were used, for 48 different runs. The indicators (Figure 19) show no contrast among the rules in the median Nchanges, which were all 12. Most of the rules resulted in median RebuildYr 2007 or 2008. MeanCatch ranged from 160 to 194 t. The %<*Bmin* was under 1% for all rules. MeanCPUE ranged from 0.84 to 0.98, well above the target, and %nearTarget ranged from 16 to 19%.

Table 14: Values for projection variables used for the exploratory runs of the operating model for the Breen rule.

Variable	Value
Projections to	2025
MLS male AW	52
MLS male SS	54
MLS female AW	100
MLS female SS	60
ρ	0.30
$\sigma^{\bar{\varepsilon}}$	0.00
σ^{Iproj}	0.10
I^{target}	0.75
T'	160, 180, 200, and 220 t
S'	15, 20, 25, and 30 t
p	1.0, 1.2 and 1.4

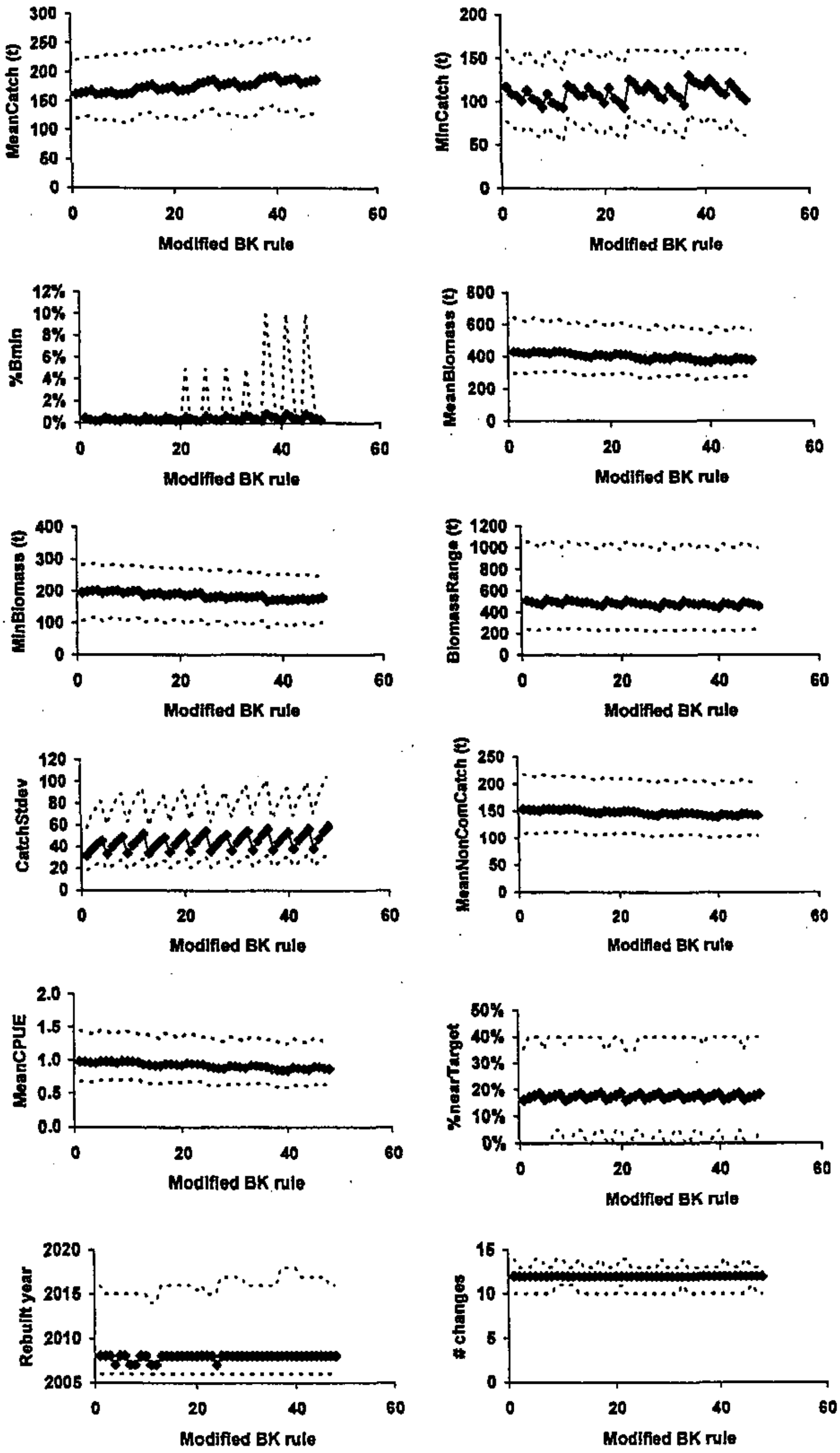

Figure 19: Indicator values from each of the 48 versions of the modified BK rule tested with the base case operating model. Diamonds show the median of the indicator (except for % < *Bmin* and % near target, which are the means) and the dashed lines show the 5th and 95th percentiles.

A crude examination of average indicators versus the various rule parameter values (Table 15) showed that MeanCatch was related to T' while MeanCPUE and the percentage of runs rebuilt by 2011 were inversely related to T'. Other results were not sensitive to this rule parameter. MeanCatch, %nearTarget and percentage of runs rebuilt by 2011 increased with S'', although not strongly, and MinCatch decreased. Higher values of the power parameter p gave lower MeanCatch and MinCatch and higher CPUE and catch variability.

Table 15: Mean values, averaged across the rule parameter values indicated, of the medians (for %<*Bmin* and % near target, the mean) indicators from runs of the Breen rule versions base case operating model.

T'	Mean catch	Min Catch	Mean NCCatch	%<*Bmin*	Catch Stdev	Nchanges	Mean CPUE	%near Target	% Rebuilt
160	163.7	103.2	152.7	0.3	41.5	12	0.968	17.3	85.2
180	172.5	107.7	148.6	0.3	44.0	12	0.928	17.6	83.1
200	180.5	112.2	145.0	0.4	46.0	12	0.895	17.6	80.2
220	187.5	116.5	141.9	0.5	48.0	12	0.864	17.4	77.3
S''									
15	173.4	119.1	148.3	0.6	35.6	12	0.927	16.3	79.5
20	175.2	111.7	147.5	0.4	42.8	12	0.919	17.1	81.2
25	176.9	106.8	146.6	0.3	48.2	12	0.909	17.8	82.1
30	178.8	102.2	145.8	0.3	52.9	12	0.901	18.6	83.0
p									
1.0	178.8	115.3	145.9	0.4	42.8	12	0.902	17.6	79.8
1.2	175.8	109.5	147.1	0.4	45.1	12	0.914	17.5	81.5
1.4	173.6	105.0	148.1	0.4	46.8	12	0.925	17.4	83.0

Typical TACC and CPUE trajectories from the Breen rule under the base case operating model are shown in Figure 20. The rule is closely responsive to CPUE, although some unavoidable lag is evident. Compared with the Breen rule (see Figure 16), this rule responds more slowly to changing CPUE because of the limitation on the amount of change that is possible in any one year.

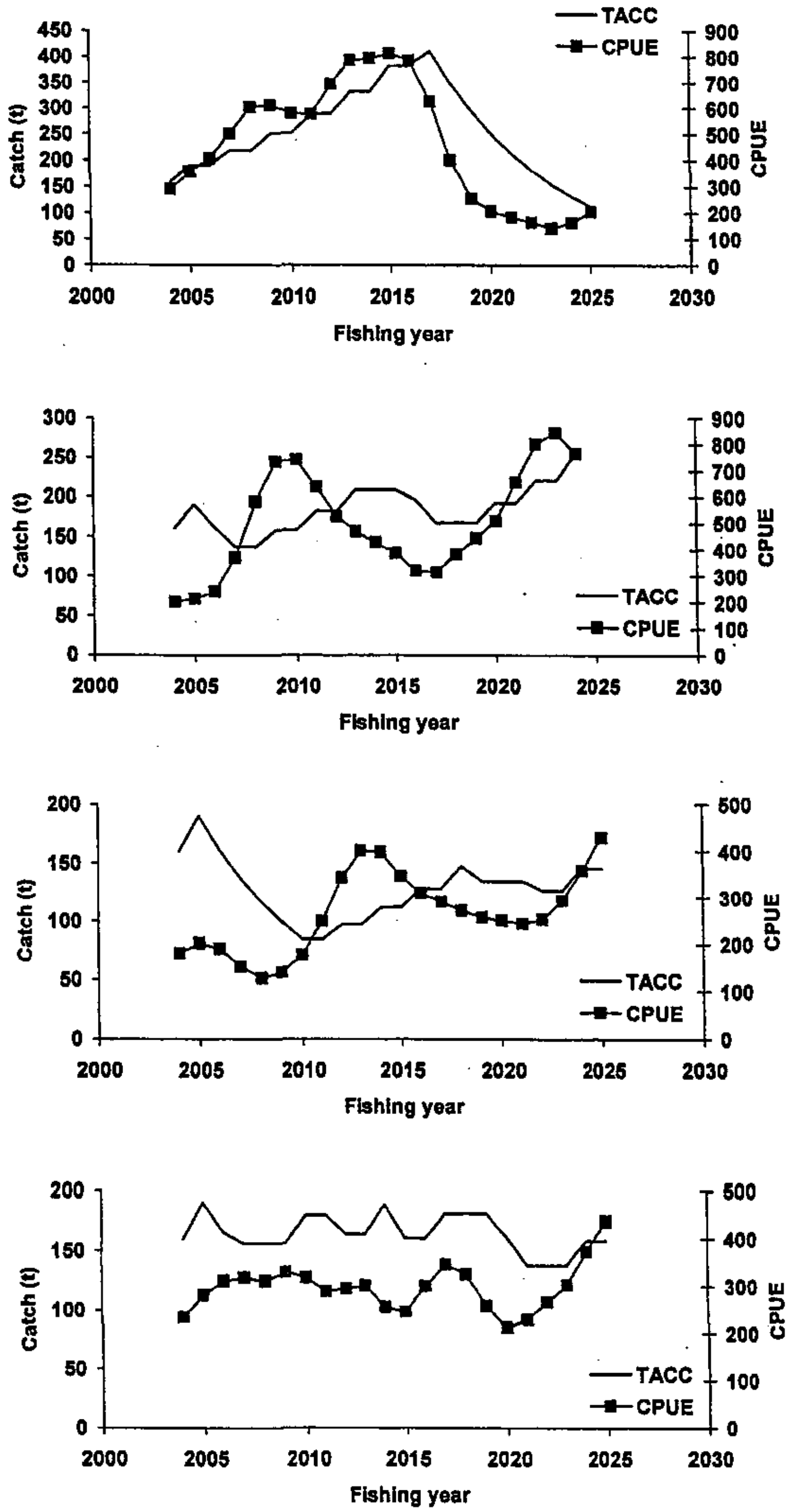

Figure 20: Typical catch and CPUE trajectory from the Modified BK rule under the base case. The panels show the 5th, 9th, 11th and 13th runs, as for the Bentley rule (Figure 12). In this rule, later called BK121, T'=180, S''=15 and p=1.

5. BASE CASE RESULTS

5.1 Results from 215 rules

We ran a base case set of runs for 215 different rules from the four families. These rules included the 90 Bentley rules, 12 Kim rules and 48 modified BK rules described above, and 65 Breen rules (a few additional rules were added to those described above to explore the edges of the results surfaces).

To facilitate comparisons of rules we developed a simple rule naming system, in which an initial letter designates the rule family (B, P, K and BK for Bentley, Breen, Kim and modified BK respectively) and digits give the level of each parameter. The values used to define parameter levels are shown for each rule in Table 16 in the order they are used in the name. Thus, for instance, Rule B2353 is the Bentley rule with level 2 of $N'=2$, level 3 of $W'=0.6$, level 5 of $S'=1.50$ and level 3 of $Min=0.15$.

Table 16: For each rule family, values of parameters and the level used in naming the rules. Not all possible combinations were used.

Levels	0	1	2	3	4	5	6	7	8	9	10	11	12
Bentley													
N'		1	2	3									
W'		0.2	0.4	0.6									
S'		0.50	0.75	1.00	1.25	1.50	1.75	2.00					
Min		0.05	0.10	0.15									
Breen													
T'	160	180	200	220	240	260							
S'		15	25	35	45	55							
p		1.0	1.2	1.4	1.6	2							
Kim													
T^{max}		160	170	180	190	200	210	220	230	240	250	260	270
modBK													
T'	160	180	200	220	240	260							
S''		0.15	0.20	0.25	0.30	0.35							
p		1.0	1.2	1.4	1.6	2							

The range of indicators seen in this large set of runs is shown in Table 17. Some showed high variability: MeanCatch varied from 143 to 223 (from very bad to very good if the sustainable yield is near 200 t under the CPUE target), MinCatch from 63 to 160 (63 is only about one-third of the current TACC). The %nearTarget varied through a factor of two, MeanCPUE from 0.7 to 1.0. The safety indicator %<*Bmin* was above 95% for all rules.

Table 17: Summary statistics for each indicator across the 215 runs described

	Index	Min	Mean	Max
MeanCatch	Median	147.4	185.9	223.8
MinCatch	Median	63.2	104.6	160.0
MeanNCCCatch	Median	126.9	142.6	158.9
MeanBiomass	Median	309.0	380.7	457.7
MinBiomass	Median	128.6	176.7	205.9
BiomassRange	Median	313.8	462.5	586.8
CatchStdDev	Median	20.3	56.6	113.2
%nearTarget	Mean	13.1	17.9	27.7
MeanCPUE	Median	0.705	0.867	1.037
Nchanges	Median	5	11.1	17
RebuildYr	Median	2007	2008.1	2009
%>*Bmin*	Mean	96.4	99.3	100.0
%Rebuilt by 2011	Median	63.6	76.5	89.3

Results from these runs are shown in Figure 21 through Figure 25 in whi
we considered to be of interest are plotted against one another. We discuss

- MinCatch vs MeanCatch shows a generally level relation, with so
have high minimum as well as high mean catches.
- MeanCPUE vs MeanCatch shows a strong negative relation that c
- catch relation. This relation suggests that MeanCPUE is redun
the base case.
- CatchStdDev vs MeanCatch demonstrates a weak trade-off (weak
in a central mass): low variability of catch is associated with lo
thus a trade-off between yield and stability.
- %>*Bmin* shows little contrast, with all rules greater than 95%, an
be discussed.
- %nearTarget is highest for rules with the highest catches.
- %rebuilt vs MeanCatch shows a trade-off between abundance
highest rebuilding success tend to have lower catches and vice-ver
- CatchStdDev vs MinCatch shows a relation: rules with high vari
minimum catches.
- %nearTarget vs MinCatch shows a trade-off: the rules with best s
target CPUE are also those with the lowest minimum catches; thi
stability and yield.
- %rebuilt vs MinCatch also shows a tradeoff between stability a
rules with low minimum catches tend to have better rebuilding su
- MeanCPUE vs MinCatch shows no relation.
- %nearTarget vs CatchStdDev shows a strong trade-off between
of the rule at keeping CPUE near its target. The highest suc
indicator, comes from rules with high variability.
- MeanCPUE vs CatchStdDev has a weak negative relation: rules
tend to have lower variability.
- %rebuilt shows a weak positive relation with catch variability.
- MeanCPUE vs %nearTarget shows a moderate negative relati
many rules are allowing CPUE to obtain levels higher than the ta

[redacted]

- %rebuilt shows a positive relation with %nearTarget, but little relation with MeanCPUE.
- %rebuild shows a positive relation with %>*Bmin*.

Thus, although safety is not a factor in these base case runs, there are some strong trade-offs among the rule results: rules with highest yield and success at achieving the target CPUE are those with highest variation in catches and lowest minimum catches. Success at rebuilding in a reasonable time shows weaker trade-offs with catch and stability.

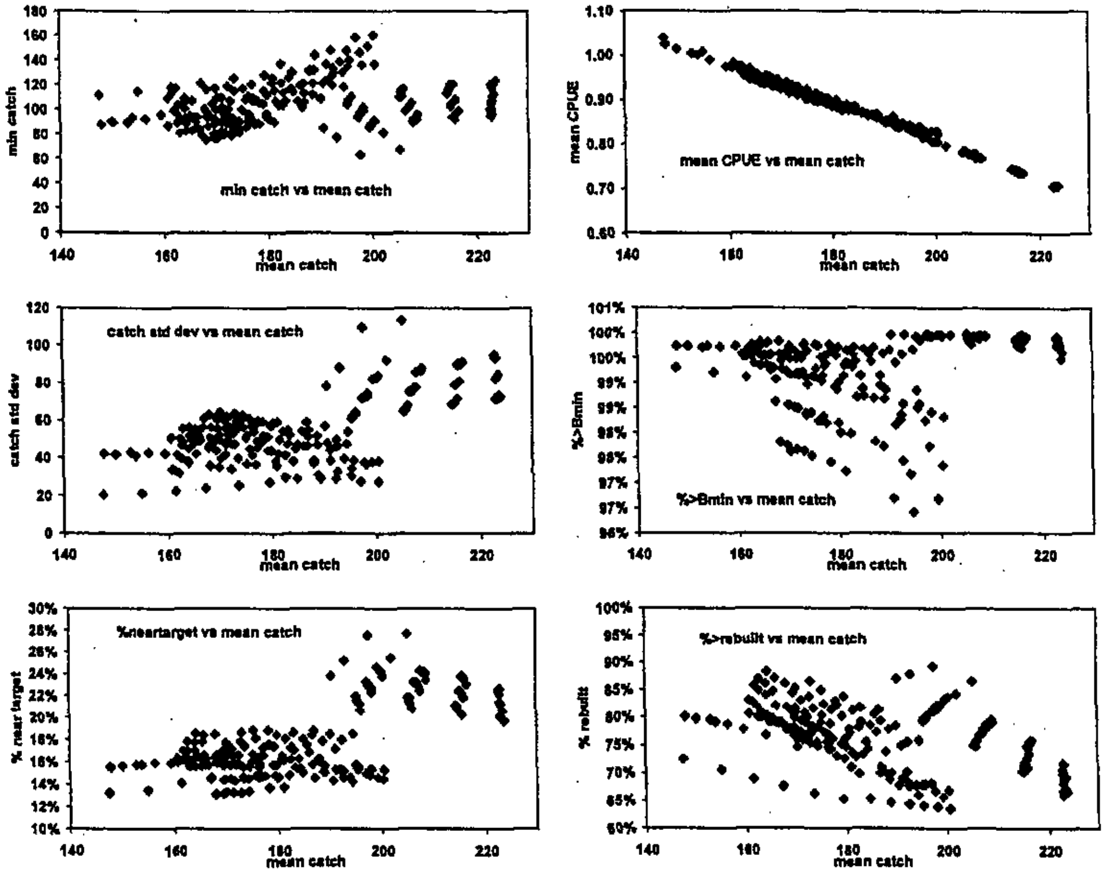

Figure 21: Major indicators from base case runs plotted against MeanCatch for the 215 rules described in the text.

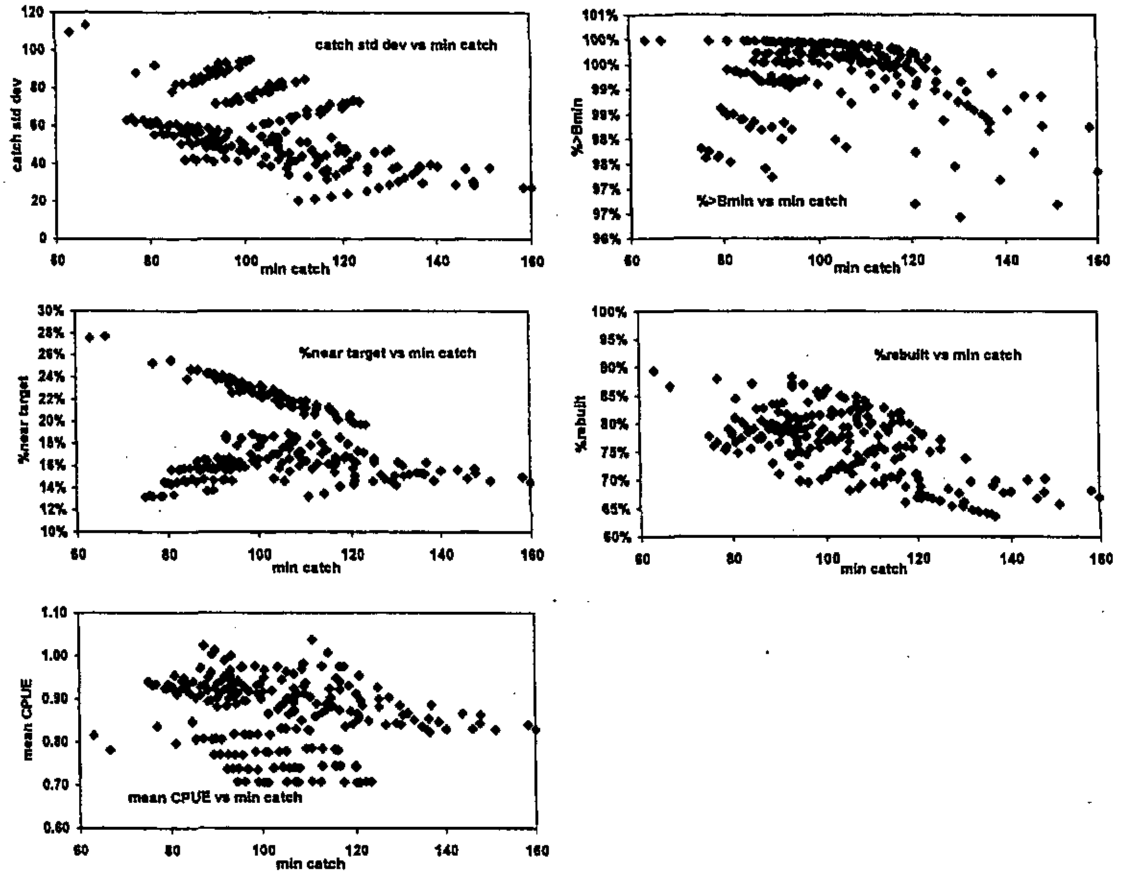

Figure 22: Major indicators from base case runs plotted against MinCatch for the 215 rules described in the text.

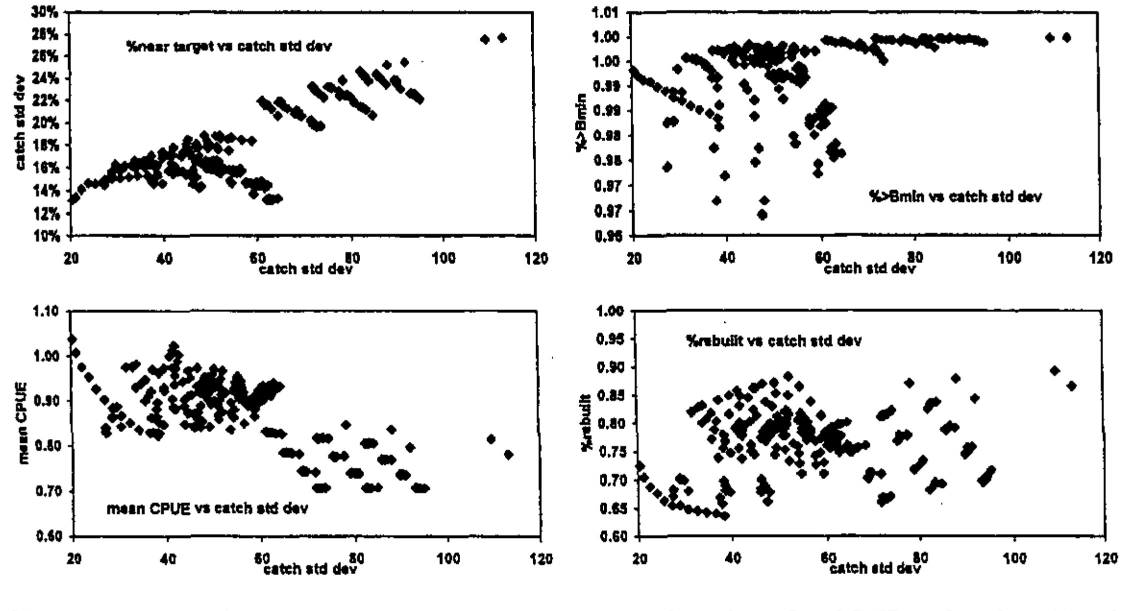

Figure 23: Major indicators from base case runs plotted against CatchStdDev for the 215 rules described in the text.

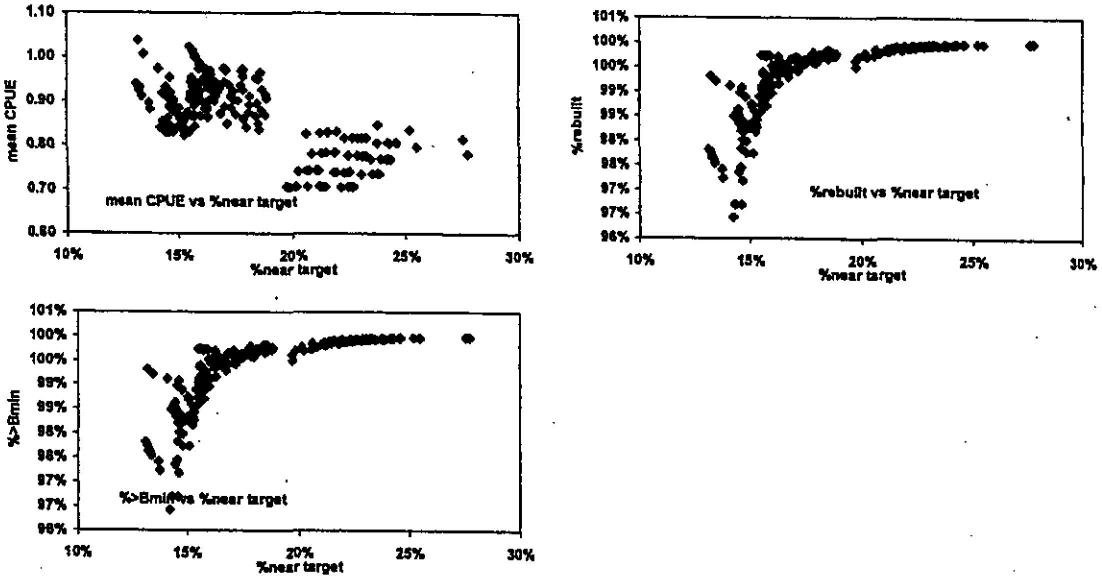

Figure 24: Major indicators from base case runs plotted against %nearTarget for the 215 rules described in the text.

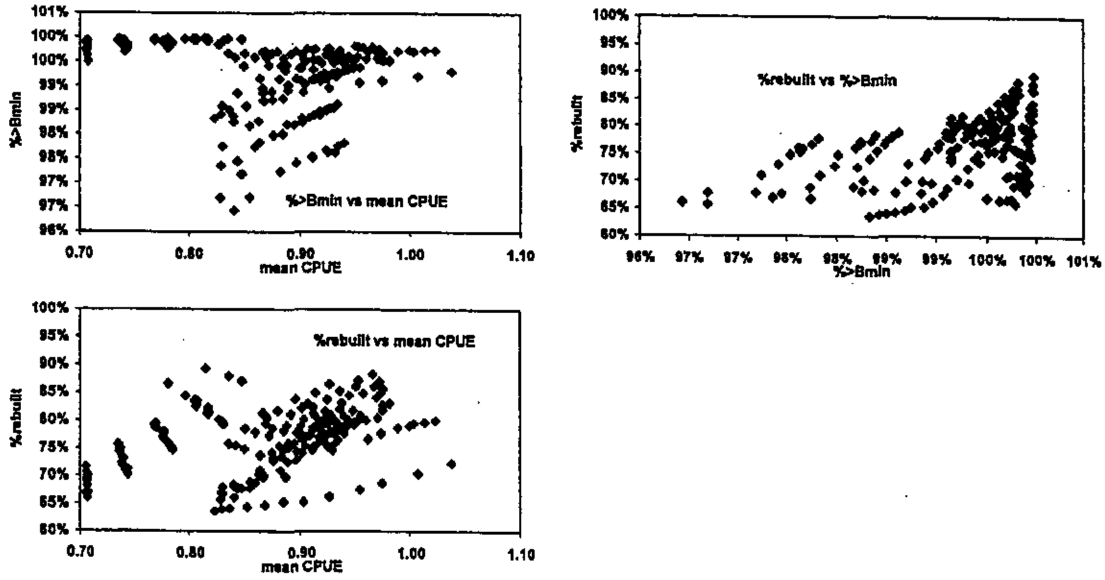

Figure 25: Major indicators from base case runs plotted against MeanCPUE (left) and %>*Bmin* for the 215 rules described in the text.

5.2 Towards choosing a rule

There are several possible approaches to using the large mass of results to find candidate harvest control rules that are better than the others. The data are massive: 215 rules and 13 indicators (although not all indicators are equally interesting).

The simplest approach would be to inspect a table with one row for each rule and one column for each indicator. With the large numbers of rules and indicators used in this study, this is not a viable approach.

Another approach might be to define a utility function for a set of major indicators. This approach was described and developed for some indicators by Bentley et al. (2003a). For instance, industry might have a target CPUE of 0.75, and they might think any CPUE below 0.45 is so undesirable that the utility is zero for rules producing MeanCPUE less than this; they might think utility is 1 for any rule delivering MeanCPUE higher than 0.90. One could assign a linear utility to values between 0.45 and 0.90 for each rule, zero to values less than 0.45 and 1 to values higher than 0.9. Taking this approach for other indicators, one could calculate an overall utility by multiplying the individual factors so that rules with zero utility for any one factor would have zero utility overall.

Problems with the utility approach lie in the complexity of developing the specific utility functions and in weighting the different indicators in a way that is compatible with the actual goals of the managers or stakeholders.

The approach taken for the CRA 7/8 decision rule (Bentley et al. 2003b) was to screen rules using the probabilities associated with a subset of the indicators. This approach essentially compares the relative probabilities of each rule delivering the desired outcome. For the CRA 7/8 decision rule, Bentley et al. (2003b) used the probabilities of rebuilding the stock by a specific date, obtaining a net increase at the end of the projections, and having a low inter-annual variability of catch.

The probabilities for each of these criteria were multiplied to obtain the joint probability; rules were ranked by the joint probability and the chosen rule was taken from among those near the top. The chosen rule in that situation was not the rule with the very highest joint probability, because the top few rules were very similar in their values: the screening procedure reduced a very large number of rules to a dozen or so final candidates, from which the final rule was chosen by looking at the performance of the top few candidates from the whole decision table.

We adopted this screening approach. First we imposed another first procedure on the rules that we called winnowing. Some rules perform badly in a key indicator (Table 17), and some performances seem sufficiently bad that the rule could be eliminated from further consideration no matter how well it performs on any other indicator. For instance, in the base case results, some rules produce a median of mean catch that is less than 150 t, 20–25% less than a good rule. It is unlikely that the industry would accept such rules.

Another example occurs when one considers what catches would be suggested by a rule if current CPUE were half the target level. The Bentley rule would produce an SCC of 143 t because of the maximum decrease parameter. The Breen rule family produces SCC as low as 40 t when T' is low and p is high. This is about a fifth of the current TACC and would essentially shut the fishery down.

We reduced the number of rules for further consideration by finding and flagging rules with low MeanCatch (less than 170 t), that would produce a low SCC (less than 120 t) next year if this year's CPUE were 0.38, and rules that had less than 70% chance of rebuild by 2011. The rules surviving this cut were 77. Of these, 43 were Bentley rules and 34 were modified BK rules. No Breen or Kim rules survived.

Choosing criteria for screening is somewhat arbitrary but can affect the relative ranking of results. Balancing this is the possibility of choosing a rule from among the top dozen or so (as determined by screening) by looking at criteria not used in screening. Our philosophy was:

- yield is already ensured by winnowing;
- safety is not an issue in the base case but could be an issue in the robustness trials;
- stability is a major issue identified by the industry;
- one aspect of performance was addressed by winnowing on %rebuild, but
- a good rule should keep CPUE near the target.

The criteria we chose for screening were:

- C_1: the probability that, for any year, biomass was greater than *Bmin.*
- C_2: the probability that, for any year, CPUE was within 10% of the target level and
- C_3: the stability of catch as measured by (1-CatchStdDev/max(CatchStdDev)).

The screening equation was

$$C^* = KC_1C_2C_3$$

where K is the flag (zero or one) resulting from winnowing. Of the three criteria, C_3 appears to the one having the highest effect on ranking (Figure 26).

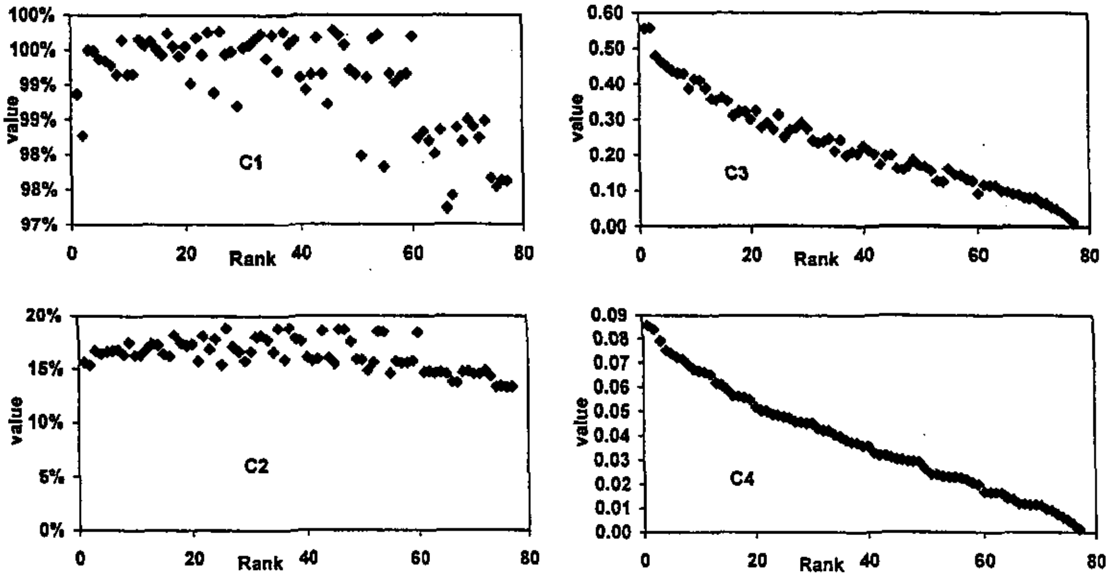

Figure 26: Values of the screening criteria C_1 through C_3, plotted against rank for the 77 rules, and the C* criterion used to rank rules.

Sensitivity to the screening criterion was explored with a different set of three. In this trial we used:

- C_1: the probability that biomass was greater than *Bmin*.
- C_2: the probability that biomass had rebuilt by 2011 and
- C_3: mean catch as measured by (1-MeanCatch/max(MeanCatch)).

Instead of valuing stability and CPUE near the target, this screening places value on yield and rebuilding. Results (Table 18) show quite different rankings, with only five rules appearing in the top 20 in both trials. Sensitivity to screening is a problem that can be resolved only by getting stakeholder agreement on what criteria to use for screening. For the remainder of the study we used the first screening procedure described above, for the reasons described above.

Table 18: Rankings from the screening criteria described above (base) and from the simple sensitivity trial using different criteria (trial). Shading shows rules that were in the top 20 in both exercises.

Rule	Base	Trial
B2111	1	72
B3111	2	75
BK121	3	12
BK122	4	5
BK221	5	53
BK222	6	36
BK223	7	27
BK321	8	70
BK131	9	18
BK322	10	61
BK323	11	49
BK132	12	11
BK231	13	56
BK133	14	4
B2131	15	37
B2132	16	54
BK141	17	28
BK232	18	45
BK331	19	74
BK233	20	33

The top 20 rules and their major indicators are shown in Table 19. Five, which included the top two, were Bentley rules and the rest modified BK rules. For all these top 20, the median rebuild year was 2008. In this instance the top two rules have mean catches near the maximum of 194 t among these 77 candidates (the 4th-ranked rule barely escaped the winnowing) and they also have the highest minimum catches. There is little contrast in the biomass indicators among these rules. The top rules have relatively low %nearTarget indicators, but the differences between these and other rules are relatively small. The top two rules would change the catch limit in roughly every other year. In this instance, either of the top two rules might be acceptable if stakeholders approved of the screening philosophy.

Table 19: The top 20 rules from the screening procedure, and major indicators.

Rule	Rank	Mean Catch	Min Catch	Mean Biomass	Min Biomass	Range Biomass	Catch StdDev	%near Target	Mean CPUE	N changes	%> *Bmin*	% Rebuilt
B2111	1	*189.0*	*143.9*	*379.1*	*160.4*	*501.5*	*28.86*	*15.6*	*0.866*	*9*	98.9	70.0
B3111	2	192.0	147.8	377.2	152.4	513.0	28.85	15.3	0.862	9	98.3	70.2
BK121	3	172.4	118.6	411.8	188.0	499.9	33.93	16.6	0.930	12	99.5	80.0
BK122	4	170.0	116.7	415.0	187.7	508.0	35.03	16.4	0.939	12	99.5	80.8
BK221	5	180.9	125.3	395.1	180.3	484.6	35.83	16.6	0.899	12	99.4	77.1
BK222	6	177.0	120.2	399.7	179.3	497.6	36.61	16.6	0.912	12	99.3	78.8
BK223	7	175.3	116.7	407.2	181.3	508.2	37.04	16.7	0.922	12	99.3	80.3
BK321	8	187.9	130.7	379.1	170.1	479.8	37.15	16.3	0.863	12	99.2	73.8
BK131	9	173.6	114.5	407.9	190.3	484.8	39.80	17.4	0.923	12	99.6	81.5
BK322	10	184.0	125.3	386.5	171.7	485.3	38.19	16.3	0.881	12	99.1	75.5
BK323	11	180.7	121.2	391.5	172.5	494.5	38.37	16.3	0.896	12	99.2	77.8
B1132	12	170.4	109.0	409.9	185.7	507.9	39.77	16.8	0.929	10	99.7	74.7
BK231	13	182.6	121.6	389.8	182.1	473.8	41.70	17.3	0.885	12	99.6	78.1
BK132	14	*171.2*	*109.4*	413.0	*190.6*	*491.4*	*42.00*	*17.2*	*0.937*	*12*	99.6	83.0
B2131	15	176.3	110.4	396.8	175.2	508.6	41.53	16.4	0.904	11	99.5	75.5
B2132	16	181.3	115.2	388.6	169.1	504.7	42.19	16.2	0.883	9	99.4	73.6
BK141	17	175.9	107.8	400.7	192.9	471.3	44.86	18.1	0.911	12	99.7	82.3
BK232	18	179.2	114.8	395.2	183.6	481.2	44.12	17.4	0.903	12	99.5	80.2
BK331	19	191.0	123.0	373.6	173.5	464.1	44.14	17.1	0.849	12	99.4	74.9
BK233	20	177.2	109.1	400.6	184.2	490.3	45.50	17.3	0.912	12	99.5	81.3

Several examples of the top-ranked rule are shown in Figure 28, and of one of the modBK rules, ranked eighth, in Figure 29. These figures show the runs based on the same series of stochastic effects, and so the effects of the two rules can be compared directly.

Bentley et al. (2003b) also described the "choice frontier" approach for choosing rules. Suppose, by way of example, that the major trade-off of interest is between stability and yield. Stakeholders might want the best combination of mean catch and minimum catch. The upper left diagram in Figure 27 plots minimum against mean catch for each rule. A line running along the upper surface of the plot would connect the highest minimum catch indicators; rules below this line produce a lower minimum catch than is possible for the same mean catch. Based on this criterion alone, one would choose a rule from the upper right-hand corner.

Under the screening procedure described here, the two highest-ranked rules were close to the optimum under the criterion just described (Figure 27) and for standard deviation of catch (where the optimum line would run along the bottom of the points); they are below the optimum for %nearTarget and %rebuilt. The third highest ranked rule is not generally near the optimum position under any of the criteria illustrated.

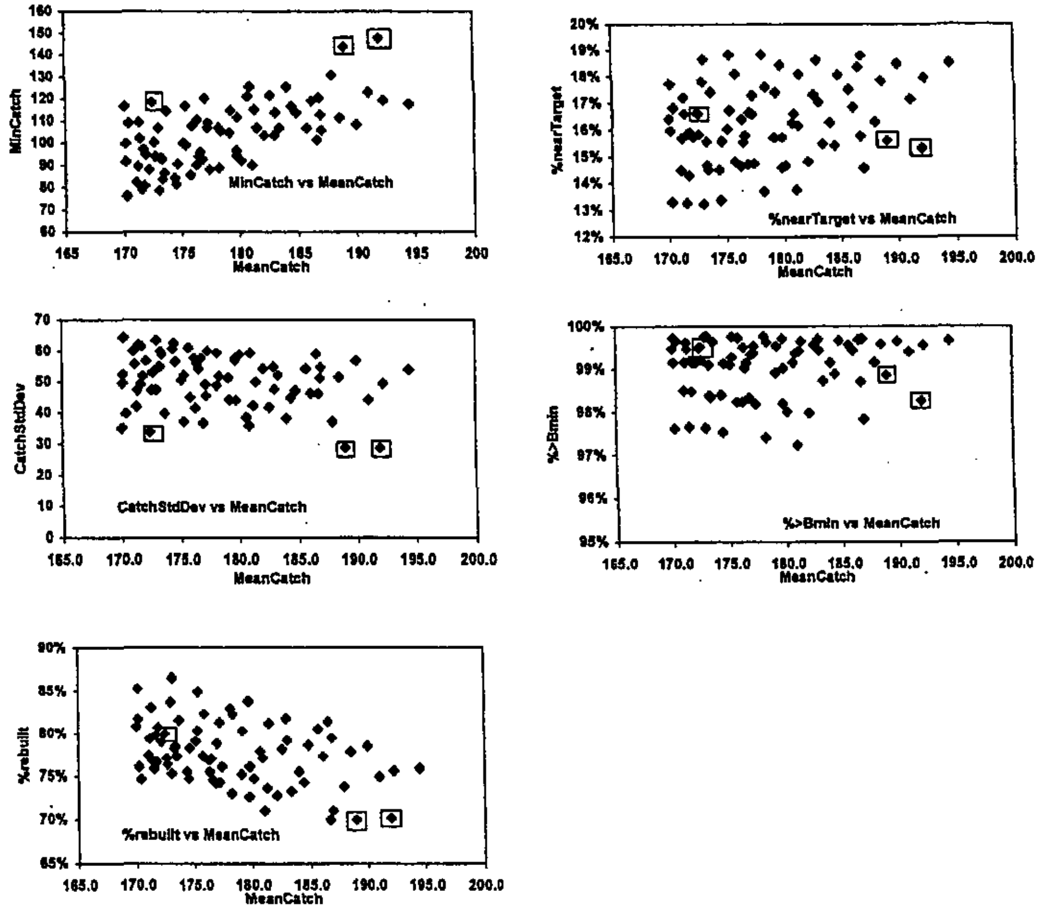

Figure 27: Indicators plotted against mean catch for the 77 rules that survived winnowing, to illustrate the "choice frontier" concept discussed in the text. Rectangles show the three highest-ranked rules under the screening procedure described in the text: they are, from left to right in all diagrams, numbers 3, 1 and 2.

The behaviour of the two rules in each specific run (in terms of the catch limits they set) is similar. Both rules are limited in the amount of increase or decrease in catch they allow from year to year (Rule B2111 is limited to 25% change and BK321 to 20% change), and this leads to a lag: decreases in TACC lag behind long-scale decreases or increases in CPUE, as the rule makes a series of smaller decreases instead of one large adjustment.

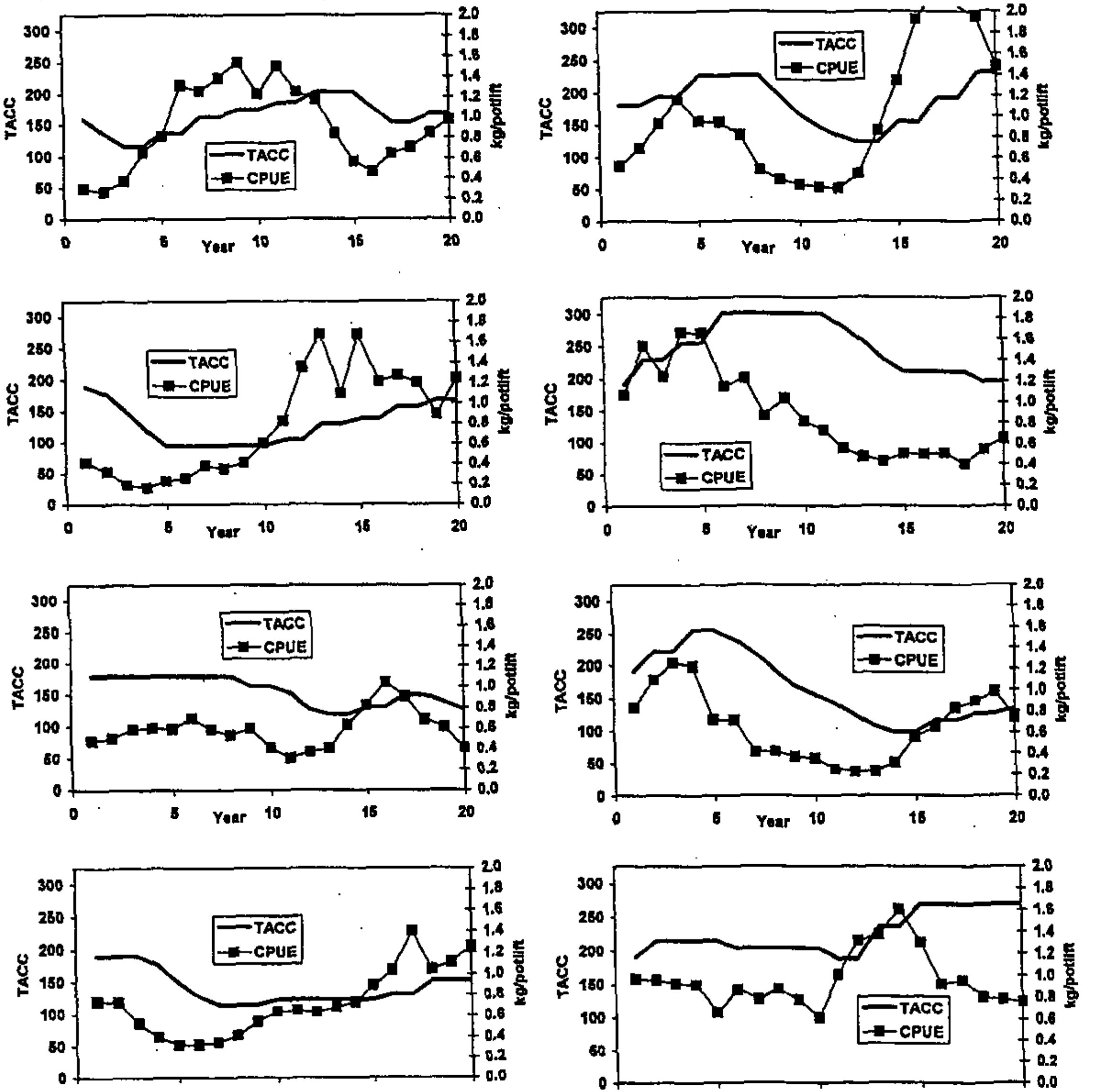

Figure 28: Examples of rule B2111, each showing TACC (t) and CPUE (kg/potlift) from 8 of the first 20 runs (numbers 2 through 16).

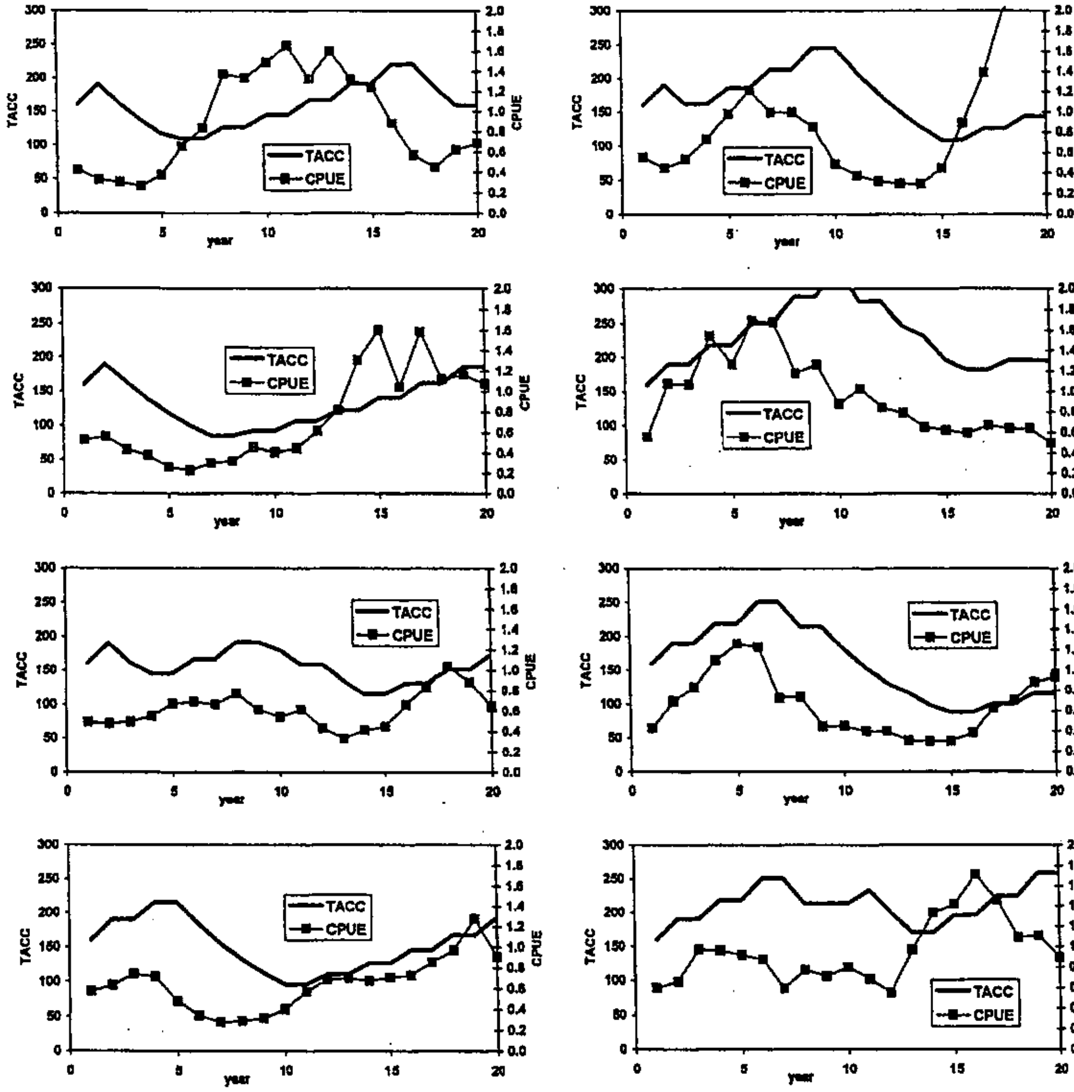

Figure 29: Examples of rule BK321, each showing TACC (t) and CPUE (kg/potlift) from 8 of the first 20 runs (numbers 2 through 16).

6. ROBUSTNESS TRIALS

Our base case operating model used the McMC results from the CRA 3 stock assessment (Haist et al. 2005) with the values shown in Table 2. Management procedures should be tested with a variety of operating models that represent a range of alternative population dynamics and productivity (McAllister et al. 1999).

A concern might be that the population could become much more depressed than the assessment predicted for the start of 2006. We simulated this situation crudely by assigning arbitrarily large catches (double the base case values, (Table 20) to the 2004 and 2005 fishing years to cause the model to reach its maximum permitted exploitation rate in these years, resulting in a low biomass at the start of 2006. This trial was called R1.

Table 20: Values (t) for projected commercial catches in 2004 and 2005 in the R1 robustness trial.

	Base case	Trial R1
AW 2004	87	174
SS 2004	73	146
AW 2005	106	212
SS 2005	83	168

In the course of this work we also used an alternative, less optimistic than the base case, result from the 2004 assessment for CRA 3 (Haist et al. 2005). Those authors made an McMC sensitivity trial with fixed growth, called "fixed growth A3", and we used the 2000 samples of the joint posterior distribution. This trial was called R2. Because these results were similar to the base case results we do not report them here.

The base case operating model uses a specific level of recruitment autocorrelation. We turned this serial autocorrelation of by making $\rho = 0.001$ and called this trial R3. These results were also similar to the base case results and we do not report them here.

Finally, we arbitrarily reduced the operating model's recruitment by multiplying all projected recruitments by 0.70 and called this trial R4.

Each of the 215 candidate harvest control rules was run for robustness trials R1 and R4. Not all were run for R2 and R3, so we do not report the results: these trials produced results in the partial set of runs that were similar to the base case results.

The R1 and R4 trials both resulted in poorer performance(Table 21): lower biomass and CPUE, higher catch variability, longer rebuild times and fewer years near to the target. Trial R4 was the more severe of the two.

Table 21: Comparative summaries of indicators across all rules from the base case and two robustness trials.

Trial		Mean Catch	Min Catch	Min Bio	Catch Stdev	%near Target	Mean CPUE	N changes	Rebuild Year	%> *Bmin*	% Rebuilt
Base	Min	147.4	63.2	128.6	20.3	13.1	0.705	5	2007	96.4	63.6
	Mean	185.9	104.6	176.7	56.6	17.9	0.867	11.1	2008.1	99.3	76.5
	Max	223.8	160.0	205.9	113.2	27.7	1.037	17	2009	100.0	89.3
R1	Min	155.8	24.2	80.4	35.8	10.4	0.606	6	2008	88.2	24.1
	Mean	189.9	93.8	116.8	64.2	16.0	0.784	12.0	2010.8	95.7	54.5
	Max	227.2	166.5	162.1	120.7	26.7	0.933	17	2016	99.7	89.8
R4	Min	87.5	27.1	92.4	24.0	9.5	0.482	5	2008	86.9	27.6
	Mean	123.6	67.8	131.2	41.6	15.4	0.648	11.0	2013.4	97.0	40.4
	Max	196.7	105.7	167.1	120.7	26.7	0.782	15	2019	99.7	86.4

Output was collated, keeping results from each robustness trial separate, and the screening procedure described in Section 5 was conducted for each trial. Winnowing that had been conducted on the base case results was not repeated: the flags were retained and used as for the base case. For all rules, we determined the ranking of the rule for each robustness trial by sorting the rules from 1 (the rule with the highest joint probability) to 77 (the rule with the lowest), as had been done for the base case.

The ranks had a similar pattern among the three trials (Figure 30). Low-ranking rules in one trial had a strong tendency to be low-ranking in the other trials. Many of the highest-ranked rules from the base case were in the top 20 rules in the other trials, with some exceptions (Table 22).

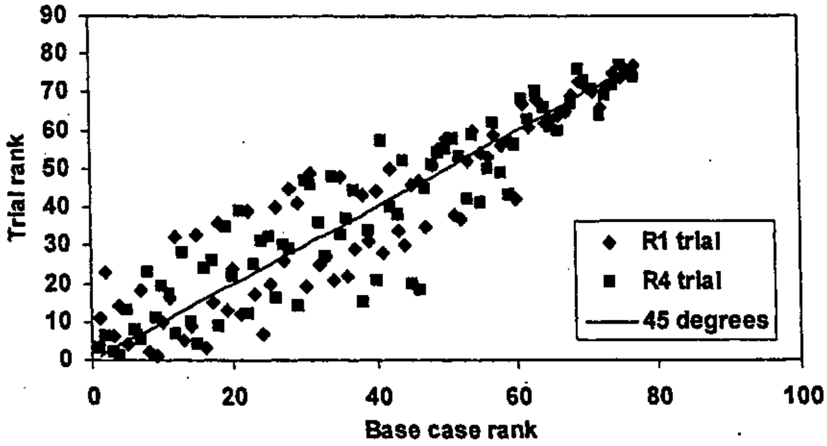

Figure 30: Ranks for each of the 77 rules remaining after winnowing, in the two robustness trials plotted against the base case rank.

Table 22: Comparison of ranks among the base case and two robustness trials for the top 20 rules from the base case.

Base case rank	R1 rank	R4 rank
1	11	3
2	23	6
3	6	2
4	14	1
5	4	13
6	8	8
7	18	5
8	2	23
9	1	11
10	10	19
11	16	17
12	32	7
13	5	28
14	9	10
15	33	4
16	3	24
17	15	26
18	36	9
19	13	35
20	24	22

Although the ranges of indicators varied in these trials, the rules performed similarly in the trials (Figure 31). For instance, rules with high minimum catch in the base case tended to have high minimum catch in the robustness trials; similarly for catch variability. A notable exception was for safety, %>*Bmin*, where the highest-ranked rules from the base case had the lowest values in the R4 trial. For the percentage rebuilt, intermediately ranked rules did better than the high-ranked rules from the base case.

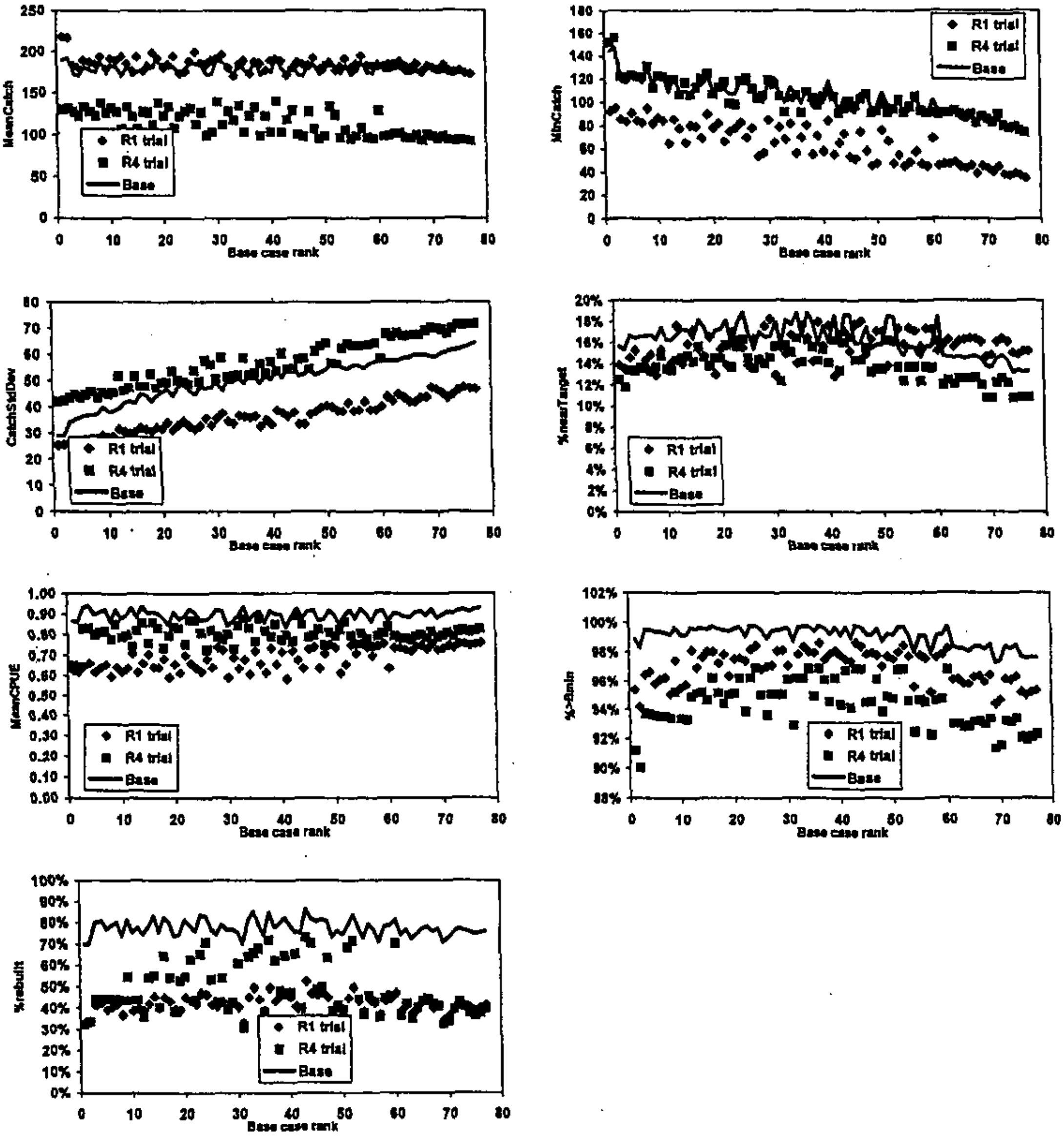

Figure 31: Indicators from each of the three trials, plotted against the base case rank.

7. DISCUSSION

Management procedures are a relatively new tool for fisheries management and have clearly demonstrated benefits. Johnson & Butterworth (2005) described the reduction of time spent discussing catch limit proposals each year for South African rock lobsters. In New Zealand's quota management system, catch changes are relatively rare but are extremely time-consuming for all parties when assessment results cause them. The NSS management procedure produced two decreases and one increase in TACC since 1997, accompanied by very little of the debate and controversy that is usual when such changes are made *ad hoc*.

Potentially the most important benefit lies in forcing stakeholders and managers to confront management objectives. To adopt a management procedure requires clear definition of objectives (Robb & Peterman 1998). There must be a shift from "tactical" thinking ("what should the TACC be?) to "strategic" thinking ("what should the harvest control rule be?") (Butterworth & Punt 1999). These authors also suggested that management procedures tend to shift the attention of interested parties away from catch levels onto the data that are used by the management procedure, such as CPUE, a process already underway in New Zealand.

The results we describe underscore the need for stakeholders to agree on their goals. After winnowing we have a set of 77 harvest rule candidates, of which the "best" rule can be defined only in terms of a set of competing indicators. We demonstrate (see Table 18) that altering the set of screening indicators alters the ranking of rules. We show (Figure 21 through Figure 25) that quite different "choice frontiers" (Bentley et al. 2003b) could be used for rule selection, and that rules that are well placed on one choice frontier are badly placed on others.

Complexity of management procedures and complexity of evaluation is a problem. In the New Zealand system, the drive for management procedures must come from stakeholders, who therefore must understand and accept them. Technical complexity in evaluation is beyond most people who are not au fait with current assessment technology. Some rules require understanding a set of equations, and even purely arithmetic equations put off many (not only stakeholders). In a system where a single management goal has been the main focus of legislation for a decade, many are confused to be confronted with a choice of alternative management goals. A focus of continuing work should be to develop communication techniques so that stakeholders can become comfortable with these issues.

The results of simple explorations with our operating model suggest that *Bmsy* is not a simple concept for CRA 3. The highest mean catch was always obtained with the highest specified catch limit or exploitation rate. This result implies that, under the current MLS regime, a cohort of lobsters is shrinking in weight by the time it recruits to the fishery: mortality exceeds growth. What is estimated as "mortality" by the assessment model may include other processes – emigration, decreased vulnerability, or some other mechanism – but the effect is the same. The strict "maximum sustainable yield" is obtained by fishing very hard; the biomass associated with that is very low, replenished regularly by recruitment. Stakeholders would not be happy with strict *Bmsy* management because of the low catch rates associated with it. Strictly speaking, the history-based target identified by CRA 3 stakeholders is associated with a biomass **above** *Bmsy*. A different target is essential, and that chosen by stakeholders appears highly workable.

Simulations conducted during this study comprised more than a million model projection runs, including some exploratory work not reported. The rules tested showed a wide range of performance: many seemed acceptable and many showed poor performance. The study explored only three rule families (if constant catch is eliminated), which is small given the variety of possible families and the variety of ways to buffer rule behavior (Breen et al. 2003). However, expansion of rule families causes enormous magnification of the number of runs required. It was encouraging that the "best" rules identified in base case evaluations were a mixture of rules from two families: this suggests that choice of rule family may not be critical if a wide enough range of members is considered. Some authors (e.g. Polacheck et al. 1999) consider that constant exploitation rate rules should perform best. "Best" cannot be objectively defined, but in this study the rule family closest to a constant-rate rule produced the best mean catches and CPUE with respect to the target.

The rules we describe and evaluate here all compare CPUE with a target. The Bentley rule also uses rate of change of CPUE. The Breen family of rules (also the basis for the modified BK family) does not estimate biomass, but does attempt to maintain a constant catch rate strategy (or, with $p > 1$, an adaptive rate strategy in which fishing rate increases with biomass). It might be possible to include the pre-recruit index derived from catch sampling as an index, as suggested by Bentley et al. (2005), but the most recent stock assessment suggested that the pre-recruit index in CRA 3 may not contain much information; in any case evaluating such a rule was beyond the resources of this project.

In this study we modelled non-commercial catches, we believe, in the most realistic way. It is very likely that these fisheries are adaptive: catch increases as biomass increases. Recreational fishers are more likely to target lobsters when they know they have a reasonable chance of catching some, and they are more likely to take their bag limit when lobsters are more abundant. As abundance increases, so will numbers of fishers, fishing days and lobsters per trip; similarly for illegal fishers.

The effect of non-commercial catches on commercial catches is difficult to estimate: customary and illegal fishers are not limited by the size limits or prohibitions on berried females or winter females. When we explored this effect in a set of special runs, where we turned off non-commercial fishing in the operating model, non-commercial fisheries appeared to be the equivalent of just over 100 t of commercial catch from a stock near the target CPUE level.

If non-commercial fisheries do operate adaptively, in the way we modelled them, then commercial stakeholders must consider their CPUE target carefully. A lower target effectively "allocates" less catch to the non-commercial sector, and vice-versa. There would be scope for a bio-economic examination of the balance between yield and costs under various management procedures.

The concepts of screening and choice frontiers, which we used to compare rules, come from Bentley et al. (2003b). Winnowing, not used for rock lobster work before, is closely related to the concept of minimum acceptable standards for management procedures. Which of the three approaches is most useful depends on the reaction of stakeholders.

A lesson learned in this study relates to choice of indicators. We chose a set based on the classic concerns of yield, safety and stability. We did not initially consider what the initial TACC set by a rule might look like: the current CPUE in CRA 3 is, very roughly, half the target level. Depending on its parameters, most Breen rules would set a TACC of less than 100 t for the first year of operation. This would be unacceptable to stakeholders and would make the management procedure unacceptable. Johnson & Butterworth (2005) discussed the importance of the first year's catch limit in gaining acceptance for management procedures – they use the term "bribery"! When we did consider this indicator in winnowing, it wiped out nearly the entire Breen rule family.

Winnowing underscores the trade-off between stability and yield. The Breen rules had quite high mean catches (well above those of other rules) and performed well in other respects. The relatively lower performance of rules that remained after winnowing, especially in terms of yield, is a price paid for stability. Stability is produced by buffering rule behaviour, introducing lags and slowing rule behaviour. When CPUE is above the target, the rule cannot quickly increase catch, so lobsters die or are caught by non-commercial fishers instead of being caught by the commercial fishery. When CPUE is below the target, catches are not decreased quickly and biomass falls to lower levels than under a faster rule. Against this, of course, faster-acting rules such as the Breen rule family would create economic hardship for marginal fishers.

The work described here assumed that the MLS regime would remain the same. There has been much discussion and debate recently about the winter MLS for males, 52 mm tail width in June through September. If this changed, the target CPUE based on the AW fishery would require re-consideration. When we changed the male AW MLS in an experimental set of runs (not reported) without changing the target CPUE, rule performance was substantially degraded and 30 to 40 t of catch was transferred to the non-commercial fishery.

Cooke (1999) suggested that management procedures must be tested against a wide range of mis-specifications of, or uncertainties about, the underlying reality. Specifically, he suggested testing with a range of productivities, different starting conditions, misreported catches, regime shifts, incorrect stock structure, trends in bias of the abundance indices, alternative stock-recruitment hypotheses, linear or cyclic trends in productivity, and episodic events.

The robustness testing reported here is not fully representative of that range because of time constraints. Some trials in exploratory work used alternative productivity models by using the results of McMC trials from the assessment and used alternative recruitment models (no autocorrelation). These trials did not differ substantially from the base case.

In the trials we report, and with the screening criteria we chose, rules performed generally similarly: a "good" rule in the base case tended to be "good" in the robustness trials (Table 22). The effect on individual indicators was also similar, an important exception being safety (Figure 31), where the "best" rules had the worst safety performance in the R4 trial, although safety levels were still high. In general, the effects of robustness trials on rule ranking were smaller than the effects of changing the screening criteria.

This study explored only the medium-term performance of harvest control rule candidates. Management procedures are unlikely to remain in place for longer than about five years without a review, because in five years the operating model used to evaluate rules will be obsolete and performance should be re-evaluated. Such a review was written into the 2002 NSS management procedure (Bentley et al. 2003b). It can be argued, therefore, that only the short-term behaviour of a rule is important.

Against this view, Breen et al. (2003) explored harvest control rules with a simple model over 100 years, and showed that some families demonstrate very poor long-term stability. The worst rules get out of phase with biomass and create oscillations with increasing amplitude, often crashing the fishery. Many rules, although not as pathological, create rather than damp oscillations, or are slow to react to long-term change in biomass. In the short term, such rules may behave apparently acceptably, but such rules can produce a balance between biomass and catch that is not optimal.

This study has identified a range of candidate harvest control rules that could be used in a CRA 3 management procedure. If the MLS regime is not changed, these should enable a choice to be made by the CRA 3 stakeholders.

8. ACKNOWLEDGEMENTS

This work was conducted under Objective 6 of MFish contract CRA2003/01, awarded to the New Zealand Rock Lobster Industry Council. We are indebted to Nokome Bentley for the screening and choice frontier concepts. Thanks to Daryl Sykes for his support.

9. REFERENCES

Bentley, N.; Breen, P.A.; Kim S.W.; Starr, P.J. (2005). Can additional abundance indices improve harvest control rules for New Zealand rock lobster fisheries (*Jasus edwardsii*)? *New Zealand Journal of Marine and Freshwater Research 39(3):* 629–644.

Bentley, N.; Breen, P.A.; Starr, P.J.; Sykes, D.R. (2003a). Development and evaluation of decision rules for management of New Zealand rock lobster stocks. *New Zealand Fisheries Assessment Report 2003/29.* 14 p.

Bentley, N.; Breen, P.A.; Starr, P.J. (2003b). Design and evaluation of a revised management decision rule for red rock lobster fisheries (*Jasus edwardsii*) in CRA 7 and CRA 8. *New Zealand Fisheries Assessment Report 2003/30.* 44 p.

Breen, P.A.; Hilborn, R.; Maunder, M.N.; Kim, S.W. (2003). Effects of alternative control rules on the conflict between a fishery and a threatened sea lion (*Phocarctos hookeri*). *Canadian Journal of Fisheries and Aquatic Sciences 60:* 527–541.

Breen, P.A.; Kendrick, T.H. (1997). A fisheries management success story: The Gisborne, New Zealand, rock lobster fishery. *Marine & Freshwater Research 48(8):* 1103–1110.

Breen, P.A.; Kim S.W. (2005). Exploring alternative management procedures for controlling bycatch of Hooker's sea lions in the SQU 6T squid fishery. Final Research Report For Ministry of Fisheries Project M0F2002/03L, Objective 3. Revision 4 May 2005. 88 p. (Unpublished report held by Ministry of Fisheries, Wellington, New Zealand.)

Breen, P.A.; Kim, S.W.; Bentley, N.; Starr, P.J. (2003). Preliminary evaluation of maintenance management procedures for New Zealand rock lobster (*Jasus edwardsii*) fisheries. *New Zealand Fisheries Assessment Report 2003/20.* 65 p.

Butterworth, D.S.; Cochrane, K.L.; De Oliveira, J.A.A. (1997). Management procedures: A better way to manage fisheries? The South African experience. pp. 83 – 90 *In* Pikitch, E.K.; Huppert, D.D. ; Sissenwine, M.P. (eds.) Global trends: Fisheries management. *American Fisheries Society Symposium 20.*

Butterworth, D.S.; Punt, A.E. (1999). Experiences in the evaluation and implementation of management procedures. *ICES Journal of Marine Science 56:* 985–998.

Cochrane, K.L.; Butterworth, D.S.; De Oliveira, J.A.A.; Roel, B.A. (1998). Management procedures in a fishery based on highly variable stocks and with conflicting objectives: Experience in the South African pelagic fishery. *Reviews in Fish Biology and Fisheries 8:* 177–214.

Cooke, J.G. (1999). Improvement of fishery-management advice through simulation testing of harvest algorithms. ICES Journal of Marine Science 56: 797–810.

Haist, V.; Breen, P.A.; Kim, S.W.; Starr, P.J. (2005). Stock assessment of red rock lobsters (*Jasus edwardsii*) in CRA 3 in 2004. *New Zealand Fisheries Assessment Report 2005/38.* 126 pp.

Johnston, S.J. ; Butterworth, D.S. (2005). Evolution of operational management procedures for the South African West Coast rock lobster (*Jasus lalandii*) fishery. *New Zealand Journal of Marine and Freshwater Research 39*: 687–702.

Kim, S.W.; Bentley, N.; Starr, P.J.; Breen, P.A. (2004). Assessment of red rock lobsters (*Jasus edwardsii*) in CRA 4 and CRA 5 in 2003. *New Zealand Fisheries Assessment Report 2004/8.* 165 p.

Kirkwood, G.P. (1997). The revised management procedure of the International Whaling Commission. pp. 91–99 *In* Pikitch, E.K.; Huppert, D.D. ; Sissenwine, M.P. (eds.) Global trends: Fisheries management. *American Fisheries Society Symposium 20.*

McAllister, M.K.; Starr, P.J.; Restrepo, V.R.; Kirkwood, G.P. (1999). Formulating quantitative methods to evaluate fishery-management systems: What fishery processes should be modelled and what trade-offs should be made? *ICES Journal of Marine Science 56:* 900–916.

Polacheck, T.; Klaer, N.L.; Millar, C.; Preece, A.L. (1999). An initial evaluation of management strategies for the southern bluefin tuna fishery. *ICES Journal of Marine Science 56:* 811–826.

Robb, C.A.; Peterman, R.M. (1998). Application of Bayesian decision analysis to management of a sockeye salmon (*Oncorhynchus nerka*) fishery. *Canadian Journal of Fisheries and Aquatic Sciences 55:* 86–98.

Starr, P.J.; Breen, P.A.; Hilborn, R.; Kendrick, T.H. (1997). Evaluation of a management decision rule for a New Zealand rock lobster substock. *Marine and Freshwater Research 48(8):* 1093–1101.