Report on the review of the New Zealand hoki stock assessment model

Doug Butterworth, Richard Hillary and James Ianelli

4 March 2014

1 Introduction

The Ministry for Primary Industries convened an independent expert panel to review the stock assessment model for hoki during 17–21 February 2014. The Terms of Reference, agenda and a list of participants in the open sessions are provided in Appendices 1-3. Two days of formal presentations were followed by informal meetings during the following two days, culminating in a discussion with participants of a draft version of this report presented on the morning of the fifth day. The report was finalised by the Panel following that discussion.

The assessment model for the hoki resource is complex, with unique aspects related to the need for simultaneous estimation for West and East components and their different migration patterns. Hence time was limited to conduct informed diagnostic tests and evaluations during the review.

The Panel was impressed by the quality of the presentations provided and the work associated with the data collection and assessment analysis for the fishery. In broad terms, the Panel considers the results from the hoki assessment model to be satisfactory and robust in regard to resource status and trends. In summary in relation to the Terms of Reference, the Panel's findings are:

- *Inputs:* the extent to which the data, modelling assumptions, structure, priors, and penalties are appropriate (including assessment of both biological and fishery components)

  The Panel considered these to be generally appropriate, though the Recommendations below suggest some improvements.

- *Implementation*: whether the modelling has been implemented using best practice methods or, where other approaches or approximations have been used, the extent to which these could bias the results

  The Panel considered these to be generally appropriate, though the Recommendations below suggest some improvements.

- *Sensitivities*: the choice and execution of sensitivity runs and the extent to which these can be relied upon to describe the full uncertainty in the modelling results (including an assessment of the likely sensitivity of the results and conclusions to modelling choices not formally assessed through sensitivity trials)

  The Panel considered these to be immediately sufficient, but recommends consideration of a wider range of stock structure scenarios over the longer term.

- *Reliability*: noting that information to inform fisheries management is often uncertain, the extent to which the results can be relied upon as a basis for setting the catch limit for hoki

  The Panel considers that the results are reliable to provide a basis for setting catch limits for hoki, though issues associated with estimation of the most recent recruitments will need to be addressed (see Recommendations below).

- *Improvements to modelling*: any improvements that could (or must) be made to the modelling to increase the reliability of the results for future management decision making – including the potential use of alternative models or model structures

  See the Recommendations below.

- *Improvements to data and research*: other key areas of research or data collection that could decrease uncertainty or increase the utility of the modelling for future management decision making

  See the Recommendations below.

2 Recommendations

The following sections are structured as follows: section headings define general topic areas for the model, and the recommendations begin with a unique number (e.g. "**R-1**"), and conclude with an assigned priority (see table below). Italicised paragraphs provide further comments and rationale for each recommendation. The table below describes the rankings assigned to recommendations by the Panel.

Symbol	Description
U	Urgent - Complete in time for the 2014 assessment if possible
S-H	Short term - Complete in time for the 2015 assessment
S-L	Short term - Ideally complete for the 2015 assessment but lower priority
M-H	Medium term - Complete after the 2015 assessment
M-L	Medium term - Complete after the 2015 assessment but with lower priority

2.1 Data

2.1.1 CPUE

R-1. Continue to improve on developing CPUE indices for qualitative fishery assessment and potential inclusion (for some regions/fisheries) of future data of this nature within the assessment.
Priority: M-H

In their present form the standardised CPUE indices are inappropriate to use within the stock assessment—particularly for Cook Strait. However, the CPUE indices are used in a more exploratory data analysis framework when exploring issues around survey catchability and the general consistency in trends across the surveys as a whole.

CPUE series constructed from data collected in an appropriate manner in the future could be considered as abundance indices (interpretation problems almost certainly preclude the use of past data for this purpose). However, there are a number of issues that should be explored before such future indices might be adopted for use in the assessment.

Within the standardisations shown, the influence diagrams were very useful for exploring what covariates are driving the trends, but some of the choices made for the model structure are perhaps excluding some potentially important alternative structures. Foremost of these would be the effect of year-area and year-vessel interactions, as the former is almost always important in other standardisations (especially with migratory effects such as with hoki) and the role of the latter seemed quite clear on the influence diagrams, given the changes in fleet structure and vessel effect over time. While this does complicate extracting the annual effect in which one is really interested for abundance trends, it is still worth exploring further plausible - albeit perhaps more complicated – CPUE standardisation models.

Another issue to address is the use of effort covariates within the standardisation model itself. While catch-per-tow was typically the observation being predicted by the standardisation model, variables such as duration were included within the set of covariates used in the linear predictor. There is no problem with using duration as a predictor, rather than a simple offset, but when generating CPUE based solely around catch-per-tow, the index is going to be missing additional effort information that would likely be important in interpreting CPUE as some kind of measure of fish density over time.

The most basic production function where catch is the product of effort, biomass and catchability is implicitly more complicated with the models being explored for hoki CPUE. For a simple linear effect for, say, tow duration, one is already introducing a power relationship into the implicit production function; furthermore, with polynomial smoothers this gets even more complicated. Future analyses should focus on accounting for this more complicated use of effort in the standardisation, in terms of ensuring that all of the effort variables are used to generate CPUE as an index of relative abundance, not just the catch-per-tow. In terms of data collection protocols future improvements might include additional information such as a proper record of the time a tow is targeting fish (rather than keeping them cold and wet), the problem of 15 t escapement panels on smaller boats (and apparent changes in practice over time), and the general focus on fishing densities most appropriate for factory considerations.

2.1.2 Surveys

R-2. Evidence of changes in catchability for the Sub-Antarctic summer survey was provided and implemented within the model. While allowing for changes in catchability is generally problematic, the Panel considered that the evidence provided to support these changes was sufficient in this case to justify this implementation, and recommended that the two options for catchability changes should form part of the base case set, while the scenario of no such changes should be retained as a sensitivity test only. **Priority: U**

Surveys play a key role in the assessment and are one of the main contributors to the model result of near-term increases in abundance.

The relative importance of the different surveys to the assessment appears to be in rough order:

1) *SA summer bottom trawl*
2) *Chatham Rise bottom trawl*
3) *WCSI Acoustic*
4) *Cook Strait Acoustic*
5) *SA autumn bottom trawl (only 3 years)*

The Chatham Rise bottom trawl surveys appear to be important as an index of incoming year classes and the SA data play an important role as an index for the staging (maturation) of fish heading towards the WCSI spawning grounds.

The SA summer time series had a period of apparently low catchability/availability from 2004-2007 and evidence of a return to normal conditions is now apparent. The anomalous pattern of availability/catchability was supported by other data (e.g. similar patterns for other species). A key reason for conducting fishery independent surveys is that these are designed to be comparable over time, and therefore should be treated as such in assessments. However, occasionally circumstances can arise which are sufficiently compelling to justify departure from this practice. In this instance the Panel considered that the burden of demonstrating this had been met, so that allowance should be made for a catchability change to reduce bias in results even though this would be at the expense of some increase in variance because of the need to estimate a further parameter reflecting the extent of the change in question. Efforts to identify the reasons for the change should continue; for example, the use of environmental factors that may contribute to changes in availability/catchability. Attention should also be given to exploring variants of these base case scenarios that reflect catchability changes in other ways that are compatible with the data.

R-3. The survey catchability in the SA region should be considered the same between summer and autumn periods. Alternatively, a multiplier for the autumn survey with an upper bound of 1.0 on the prior to reflect fish having moved out of the survey area could be used. **Priority: S-L**

Since the same gear was used and availability should be moderated by movement patterns, assuming that the catchability was the same seems appropriate. However, the time steps are too coarse to allow movement to adjust the "availability" side of catchability. Therefore using a catchability multiplier to account for the difference between seasons with an appropriate upper bound for the associated prior might prove useful.

R-4. Consider revisiting the use of length frequency and age data collected from the acoustic survey trawls instead of the fishery. These types of data are used in two ways: to convert model numbers to predicted biomass estimates, and also as additional age composition data. **Priority: M-L**

Acoustic surveys use length frequency data from the fishery, in contrast to the more common practice of using samples collected by the survey trawls themselves.

Other data aspects are discussed within the context of model parameter estimation in the following sections.

2.2 Model structure

R-5. Develop a simple aggregated form of the Statistical Catch-at-Age method which, for example, may use selectivity at age for different series as a surrogate for age-dependent distribution patterns. Such an approach would probably first be applied to the no natal fidelity scenario treated as a single stock. **Priority: S-L**

The purpose here, in addition to providing readier insights into model behaviour, is first to check that a different (and simpler) model structure does not lead to results that differ considerably from those for the existing assessment model. It might prove helpful to extend this approach (typically allowing for time-varying effects and given an appropriate split of the catches), to consider two (West and East) stocks. The YCS correlation between stocks suggests single rather than separate stock-to-recruitment processes.

R-6. As a candidate for consideration, develop a model that treats the 1-year-old numbers for the Chatham Rise survey as a separate index from that provided by the older fish from that survey. **Priority: S-H**

The 1-yr-olds from the Chatham Rise survey are important as they provide the only information on the most recent year-class to join the fishery. The estimated strength of this year-class can have an appreciable impact on projections. Since different mechanisms likely control the availability of these younger fish (which have not fully settled to the bottom) to the survey, they would probably best be considered separately in the assessment to allow for a different variance-covariance structure to the older fish. Two-year old hoki may also require consideration for similar separate treatment, since their presence appears to be poorly correlated with that of 3-year olds the following year—this option should also be investigated.

2.3 Assessment specifications

2.3.1 Diagnostics

R-7. Develop diagnostics for evaluating how different data components affect trends. **Priority: U**

As an example, by specifying a pseudo index with a flat trend (and low CV), the assessment can be configured for both East and West populations to see for which data components the likelihood changes most. This can provide a means to communicate which information has the most influence on any trend pattern. For one of the MPD reference cases, a trial run with a flat index indicated that fits to multiple datasets—both composition and index information—contribute to the observed increase (based on how different likelihood components degraded when the pseudo-index was included). These were the SA summer bottom trawl survey index, and the composition data from both the CR and SA summer bottom trawl surveys.

R-8. Explore alternative MCMC algorithms for non-converged chains, or other diagnostics, to improve convergence. **Priority: S-L**

The presentations provided evidence that some MCMC chains may have failed to converge (separate chains sometimes tended to diverge rather than show similar results). This is due to either the type of MCMC sampling routine or the parameterisation of the model.

R-9. Retrospective analyses should become a routine part of assessment presentations. **Priority: S-L**

A standard diagnostic for model self-consistency tests is to conduct retrospective analyses which involve peeling off the most recent year's data and re-running the model for a given configuration. This can indicate if there is a tendency for the model to have periods of over- and under-estimation of stock trends. If a poor pattern is found, then further evaluations of alternative model specifications may be required.

R-10. Evaluate the model estimated population-level effective proportion mature at age for each stock as a self-consistency check. **Priority: S-L**

Few fish stock assessment models compute spawning biomass based on estimates of the proportion of fish (at age) that are moving from feeding grounds to spawning grounds. As an accounting cross check, computing the total number spawning at age over the total population (for each stock) over time may be useful. In particular, comparing these values between the two stocks would be interesting (particularly if the values diverged appreciably or were correlated). Furthermore, these estimates would be interesting to compare with estimates (or assumptions) common in other related stocks.

R-11. Examine the impact of selectivity (and other, e.g. B_{mean}) parameter specifications between MPD and MCMC runs. **Priority: S-H**

The Panel was concerned about the MPDs being near the edge (for critical parameters) of the credible intervals from the posterior distributions, and hence recommends that the reason for this be investigated and presented. For example, the two posterior samples provided indicated that a number of selectivity parameters were held fixed in the sampling of the posterior distribution, but were apparently estimated for the MPD runs. Documenting these characteristics of the assessment could be improved, and in instances when the MPD is near the credible limits some presentation on how this has occurred would be useful. This is considered important because MPD results are often presented and used for data weighting, and for the model selection process leading to the final MCMC base case set that is used to provide advice.

2.3.2 Data weighting

R-12. Consider alternative means for re-weighting data components and whether pooling process errors within observation errors (e.g. multinomial sample sizes for well sampled fisheries) is appropriate. **Priority: M-H**

Using the implied (by input sample size) distribution of mean age as a way of rescaling effective sample size for composition data may be ignoring important processes related to model mis-specification; this is because of a focus primarily on the central tendency of the distribution by definition, not on the structural aspect of the variance across ages within a given year.

R-13. Investigate moving from multinomial towards the new Francis (logistic-normal) procedure for fitting composition data. **Priority: M-H**

This may help resolve some technical aspects of the re-weighting process that is presently conducted for the hoki assessment. Nonetheless, other concerns about model mis-specification (i.e., pooling important process errors as observation errors) would remain.

2.3.3 Priors

R-14. The informative priors used in the assessment should each be reconsidered for possible adjustment or transformation to a non-informative form, in particular to guard against inadvertent double use of data. Additionally, a document should be developed which details how and why the various prior distributions were formulated. **Priority: S-L**

The basis used for specifying priors is important, especially where such priors are influential regarding the value of a quantity of interest for management purposes. Certain impressions about aspects of the resource appear to have been developed over time in the light of experience with earlier analyses, and are now implemented in the form of choices for some of the priors. There is an associated danger that these impressions reflect aspects of the data now input to the assessment, so that a form of double-counting may arise if such priors remain included. A possible instance of this is the specification that certain selectivities must be domed. It is likely that, in the case where an *increase in natural mortality is estimated rather than held fixed, an assumption of flat selectivity for (at least) one survey will need to be made to remove confounding (and this is appropriately implemented for the Cook Strait survey in present assessments). However, other surveys and fisheries should not have domes forced upon them a priori, but rather the data should be left to determine the shape of the selectivity curves (unless unstable estimation performance results).*

A related prior specification that needs documenting is that for the proportion which the East stock contributes to the initial overall abundance. The Panel was advised that early data (e.g. from earlier Japanese surveys) had been used to inform the choice of some priors, rather than being fit directly in the model. This is not necessarily inappropriate, but given the development of the assessment over time, it might be advisable not to incorporate such priors immediately, but rather to first compare model predictions without those priors with these data in a "qualitative consistency check" mode, and to then determine whether and how best to develop priors to reflect such data. Finally, the present upper bound on the maximum exploitation rate (typically about 0.6) should be set higher (say 0.9), unless there is compelling independent information to provide justification for a lower choice.

Documentation needs to include details of the development of the prior distribution for M. In addition, it is unclear how initialisation and the stock-recruitment relationship for the no natal fidelity option have been specified; this should be clarified as it is included as an important sensitivity within the assessment process.

2.3.4 Growth

R-15. Growth assumptions should be examined both for how the fractional year (from the growth curve) is implemented, and for how model-specified estimates compare with actual observed fishery (and survey) estimates of body mass (or length) at age. This should also include an evaluation of L-W relationships (where available) between areas, seasons and over years (presently a fixed L-W relationship is assumed). Should these indicate appreciable annual differences from current assumptions, then year-specific inputs should be used within the assessment. **Priority: S-H**

The sampling from the hoki fishery and surveys is well designed and sufficient in extent to be quite suitable for the age composition information and growth relationships that are used for the assessment. During the week some figures showing patterns in growth (and variability) were discussed. There is some evidence that growth is changing over time, and the suggestion was made that density dependent growth might be occurring. However, within the model, growth is time invariant, albeit stock and sex specific.

2.3.5 Selectivity

R-16. Alternative selectivity forms should be considered (e.g. bicubic splines or a non-parametric form with regularity penalties[1]). Furthermore, the migration "ogives" at age should be specified such that age-age variability is minimised (e.g. condition the ogive to be monotonically increasing as appropriate). **Priority: M-H**

Allowing more flexibility in selectivity (both at age and over time) would fit the well sampled fishery catch composition data more appropriately and may provide more realistic estimates of the fishing pressure on each cohort in each year. The extent that allowing for such flexibility will interact with migration estimates may be problematic but could be evaluated. Currently the selectivity assumptions are quite rigid in their form (logistic and double-normal). One of the most well sampled fisheries (WCSI) allows for variation in the ascending inflection linked to the median date of the fishing season. This has resulted in better fits to these composition data by adding only a single parameter which scales the median date to the inflection point parameter. The Panel raised a concern that the "process" errors in fitting composition data have been ignored by using a separable fishing mortality configuration; i.e. one that decomposes fishing intensity to a single age effect and a separate

[1] Martell, S., & Stewart, I. 2013. Towards defining good practices for modeling time-varying selectivity. *Fisheries Research*, 1–12. doi:10.1016/j.fishres.2013.11.001

(typically effort-related) year component. The assumption is that cohorts migrate to the spawning grounds at a fixed (but estimated) rate each year. It may be that the migration is density dependent, or that vulnerability to fishing varies with year-class strength and densities.

2.3.6 Spawning and homing migrations:

R-17. Adjust the no natal fidelity runs to allow for a proportion of the West stock to move to the East stock, and *vice versa*, each year (the current assumption is that any fish moving to the West stock at first spawning stays there, and similarly for fish moving to the East stock). Increase this proportion until the fit cannot reproduce the differential larger reduction of West stock during period of high catches in the West stock area in the early 1990s. Effecting this may require improvements to the dispersal options currently available in CASAL. **Priority: S-L**

Over time it is important both to broaden the range of model structures considered, and to seek objective analyses that assist determination of the appropriate relative weights to assign amongst such structures. The extents of natal and of spawning fidelity are important to identify, and provide a good starting point for this general exercise. The correlation of the recruitment residuals for the East and West stocks is about 0.7. This suggests favouring the no natal fidelity over the natal fidelity model (or that both stocks are subject to similar environmental conditions). It seems that the latter has been the favoured approach previously because mixing of the two stocks was thought to be incompatible with the abundance trends in the West stock area over the period of very high catches in the 1990s. The purpose of this exercise is to test whether this incompatibility claim is correct; if not, greater weight should be accorded to the no natal fidelity scenario.

R-18. Consider adding annual variability in the proportion that move from the SA to the WCSI spawning area each year and couple this with a prior penalty in order to improve fits to the fishery composition data. For estimation issues, it may be useful to assume a constant age-age trend (in the proportion migrating to spawning grounds) and apply, for example, a time-varying component to the ogive asymptote and/or the inflection point. **Priority: M-L**

This may improve the lack of fit in highly observed and sampled fishery from the WCSI spawning grounds and elsewhere—see recommendation R-12 with respect to the reweighting of catch composition aliasing for effects such as essentially unobserved changes in the migration behaviour. In particular, this may improve the fits (relative to bootstrap estimates of input effective sample sizes) to the WCSI age composition data.

R-19. Additional updated information on the proportion of hoki that spawn is needed. **Priority: S-H**

Since the proportion spawning is largely based on limited observations of histological samples spread over three years during the 1990s, additional samples are needed to determine whether changes have occurred and whether inter-annual variability may affect the model specification process. These data are clearly informative, given that removing them does alter the estimates for the West stock appreciably, so it is appropriate to consider more data relating to this process (building on the work done more recently by S. Parker).

R-20. Evaluate the extent to which movement estimates are confounded with selectivity, M, and YCS. **Priority: M-H**

The likelihood surface for some key parameters would be informative; in particular, showing the extent to which a profile over fixed plausible (and smooth with age) migration rates affects different data components would help to explain (and perhaps communicate) how movement assumptions interact in this complex model configuration. Additionally, investigation of the correlation matrix (be it from the MCMC or from the MPD) would provide a first look at any potential linkages in this region of parameter space.

R-21. Provide diagnostic checks on fits to sex ratio. **Priority: S-H**

The model is fitting proportions at age and sex in each year; this vector sums to one, implying that sex ratios are included within the objective function. Available data (Fig. 1) show strong patterns for which it might be useful to show residuals since growth and natural mortality are sex specific. Should *these residuals appear problematic, further consideration about model assumptions and configurations would be required.*

2.3.7 Natural mortality

R-22. Consider alternative specification (e.g. the Siler or the Lorenzen form) for the natural mortality at age. **Priority: S-L**

A shape that avoids the very low MPD estimates of M (for females in particular) should be considered.[2] *Analysts should check the bounds applied at present, and investigate why high values (>2.0) were observed for some MPD fits as there is concern that this may provide unanticipated interactions with YCS.*

2.3.8 Recruitment

R-23. Evaluate likelihood profiles within the assessment process. For example, preliminary investigations profiling B_{mean} suggested that one of the strongest components affecting the MPD (and presumably the posterior distribution obtained from the MCMC) was the penalty (Fig. 2). **Priority: S-H**

Apparently this component is arising primarily from the recruitment penalty and there is concern that this may not have been anticipated behaviour. The potentially complex linkage between B_{mean} *and* B_0*, and whether such an effect persists for* B_0 *as well, should also be explored.*

R-24. Attempt to improve the estimate of the most recent recruitment by means of a better relative weighting of the information provided by the Chatham Rise survey and the expectation from the stock-recruitment relationship. Given the amount of age data and length of some of the survey time series, this process might be assisted by estimating σ_R and using this to inform the prior on YCS instead of fixing it at a value that appears to be much larger than the empirical estimates obtained from the estimated year-class strengths coming from the assessment. **Priority: S-H**

Within the assessment there are a number of issues that need to be addressed with respect to how recruitment is both estimated and a priori modelled—particularly in the most recent years.

It is clear that a (potentially) strong or weak year class with very few observations can result in large or small yet uncertain year-class strength estimates that become very influential in subsequent projections. An example of this, which is of some importance from a management perspective, is provided by the high estimate for the most recent recruitment in the 2013 assessment which has a large impact on projections. A more reliable estimate of the most recent recruitment than available from a single survey follows from shrinking such an estimate towards the value expected under the stock-recruitment relationship. It is important to weight these two sources of information appropriately in the shrinkage process, and the Panel was concerned that the current assessment underweights the expectation from the stock-recruitment relationship through use of a wide prior for σ_R*. Whatever the value of* σ_R*, there might be merit in placing a more restrictive prior on the most recent recruitments that are to be estimated in the model, where that prior is perhaps based on an estimate of* σ_R *from the model fit itself.*

Estimating σ_R *may present some methodological difficulties when using a mix of MPD and MCMC estimation techniques. This parameter is not estimable in the MPD framework used to tune the data weighting prior to the MCMC runs, but would be estimable in an MCMC sense. This is because estimation makes the model a random-effect model in the MPD sense or perhaps a Bayesian hierarchical model in the MCMC sense. Given the way the current MCMC algorithm works in CASAL, which depends on the covariance matrix from the MPD fit, one can envisage potential*

---

[2] E.g., Brodziak, J., J. Ianelli, K. Lorenzen, and R.D. Methot Jr. (eds). 2011. Estimating natural mortality in stock assessment applications. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-F/SPO-119, 38 p.

problems in the MCMC estimation phase as it is quite likely that the distribution obtained from the MPD fit will not be representative of the likely distribution of σ_R*. This is a problem in the wider stock assessment sense and is perhaps already being explored in the CASAL update work that is ongoing, but is something to consider given the importance of this parameter for both the assessment and the projections.*

R-25. Attempt to develop environmental recruitment relationships for strategic planning. **Priority: M-L**

In the near term, in particular, it seems unlikely that environmental indices, which are themselves difficult to characterise, will yield useful information with which to inform future recruitment. However, if longer-term trends that contribute to variability can be identified, this may prove useful within a management strategy evaluation context (i.e. for developing plausible operating models). Nevertheless, the poor record of success elsewhere with attempts to develop such relationships cautions against investing too many resources towards such attempts.

2.3.9 Projections

R-26. Future development of the assessment should explore how to include covariance between the East and West stock year-class strengths. **Priority: S-H**

This would better reflect the relative status of future trends between East and West stocks. There is an appreciable correlation (~0.7) between the year-class strength estimates from the assessment that is not currently being reflected in the projections as they are sampled independently of each other. Furthermore this exercise should continue to include some configuration options that incorporate the auto-correlated nature of the recruits as may prove appropriate.

3 Recommendations summary

The following lists the recommendations by priority.

Priority	Recommendations
U	**R-2, R-7**
S-H	**R-6**, **R-11**, **R-15**, **R-19**, **R-21, R-23, R-24**, **R-26**
S-L	**R-3**. **R-5, R-8, R-9, R-10, R-14, R-17**, **R-23**
M-H	**R-1, R-12, R-13, R-16, R-20**
M-L	**R-4**, **R-18**, **R-25**

4 Figures

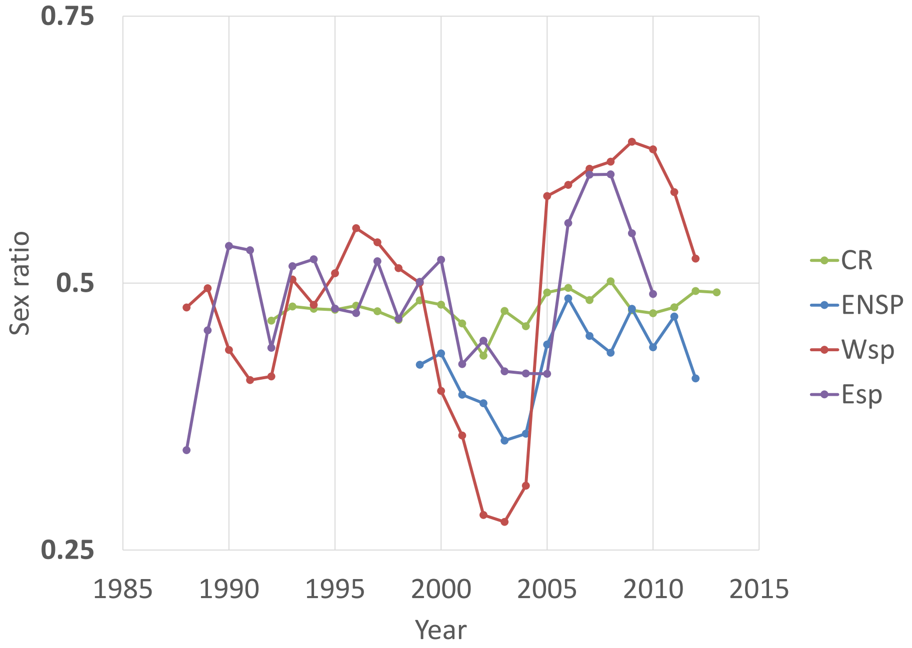


Figure 1. Sex ratio (proportion male) for longer time-series information for hoki.

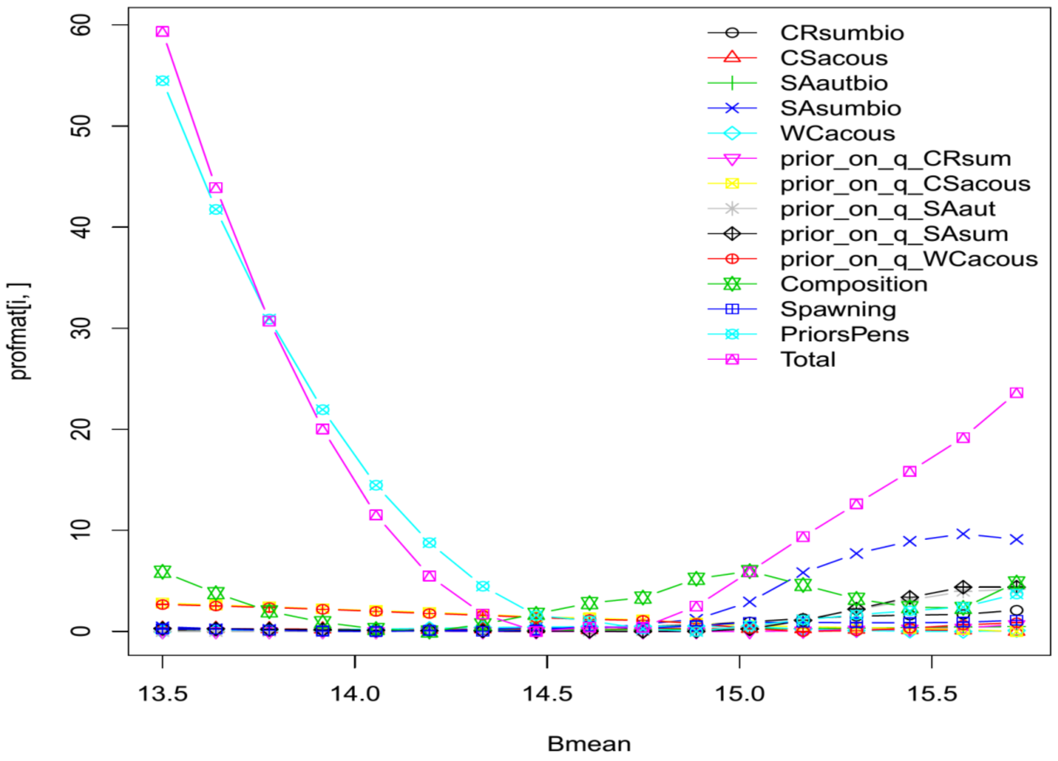


Figure 2. Example likelihood profile.

Appendix 1. Terms of Reference

Ministry for Primary Industries
Manatū Ahu Matua

Ministry for Primary Industries Terms of Reference for an independent review of the hoki stock assessment model

1. Background

Hoki is New Zealand's largest finfish fishery with a TACC of 150,000 t from 1 October 2013. Although mostly managed as a single stock, hoki are assessed as two stocks, western and eastern. Figure 1 shows the historical catches reported from each stock.

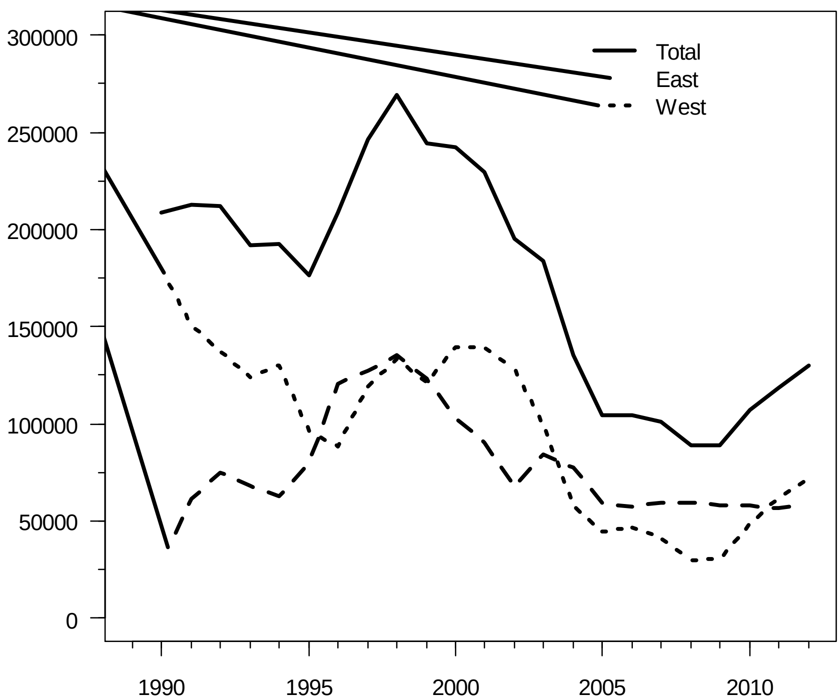


Figure 1: Eastern and western components of the HOK 1 landings since 1988-89.

The management approach for hoki is assessment-based and leads to regular TAC/TACC reviews based on the results of an annual stock assessment. The stock assessment is informed, in turn, by annual monitoring of stock biomass. The assessment provides estimates of stock status relative to the reference points described in the hoki harvest strategy.

The harvest strategy for hoki (both stocks) is described in the table below.

Table 1: Harvest strategy for hoki.

Harvest strategy components	Management response
Management target range of 35-50% B_0	Catch levels are set to maintain stock size within the target range
Soft limit of 20% B_0	The limit below which a formal time-constrained rebuilding plan will be implemented
Hard limit of 10% B_0	The limit below which the hoki fishery will be considered for closure
Rebuild strategy	Catch limit to be set so that the fishery will deliver half the rate of rebuild that would occur in the absence of fishing
Harvest control rule	There is currently no defined harvest control rule for hoki. Currently catch limits are determined by the results of five year projections of stock status relative to the above biological reference points using the results from the hoki stock assessment

Stock structure

Morphometric and ageing studies have demonstrated that there are at least two sub-populations of hoki. Hoki off the west coast of the North and South Islands and the area south of New Zealand (including Puysegur, Snares and the Southern Plateau) are assumed to be one stock unit (the "western stock"). Mature hoki in the area of the east coast South Island, Mernoo Bank, Chatham Rise, Cook Strait and the east coast North Island up to North Cape are assumed to be the other stock unit (the "eastern stock"). Immature fish of both stocks are found mixed on the main nursery ground, the Chatham rise. The genetic relationship between the two stocks is poorly determined.

Productivity

Hoki growth rates are known for both stocks from otolith readings. Routine catch at age data are collected from the four main commercial trawl fisheries and from two wide-area trawl surveys. Year class strength for each stock is estimated within the assessment models.

Stock monitoring – historical data series

Abundance biomass estimates have been obtained from fisheries-independent wide area trawl surveys of the Chatham Rise (1992-2013) and Southern Plateau (1991-93 and 2000-12) with catch at age data collected from both the commercial trawl fishery and research surveys.

In addition abundance estimates are available from acoustic surveys conducted in the spawning aggregations in Cook Strait and Pegasus areas (eastern stock) and off the west coast of the South Island (western stocks).

CPUE indices from commercial trawl fisheries are standardized each year but are not currently used as an input to the stock assessment for hoki as the trends in the indices do not match the trawl survey data.

Stock assessment – model structure and assumptions

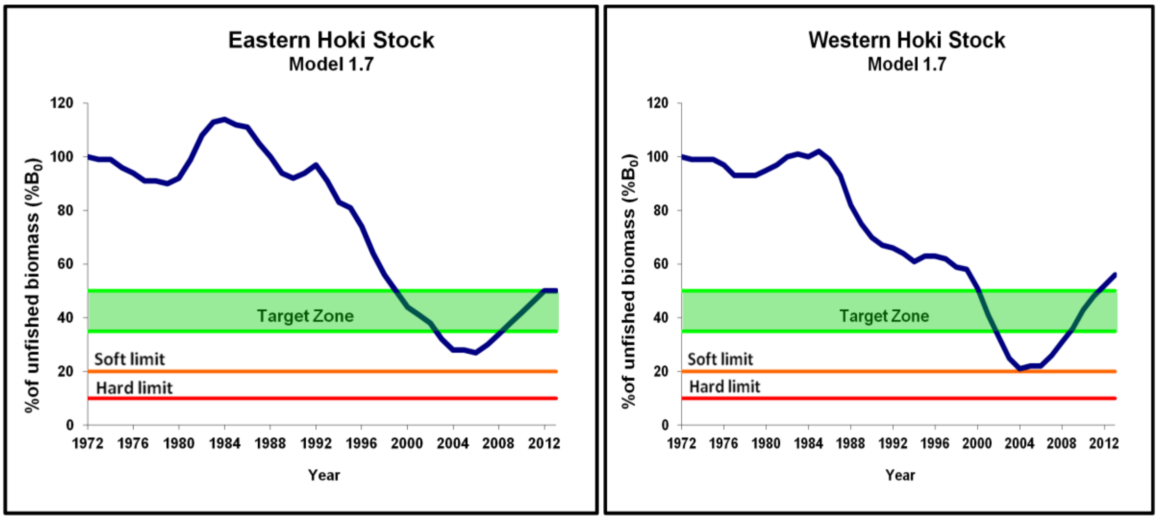
Two stocks are assessed. Fish from the eastern stock spawn in Cook Strait and have their home grounds on the Chatham Rise; the western stock spawns on the west coast South Island and have their home grounds in the Sub-Antarctic. Soon after being spawned, all juveniles move to the Chatham Rise. In the assessment two alternative assumptions concerning natal fidelity are modelled. One assumption is that adults show natal fidelity – that is, they spawn on the same ground where they themselves were spawned. Under this assumption, the stock to which a fish belongs is determined at birth. At some time before age 8 all Western fish migrate to their home ground in the Sub-Antarctic. The alternative assumption, used first in 2006, is that there is no natal fidelity (but all fish once they have spawned in a given area return there for future spawnings; i.e. adult fidelity).

The model partition divides the population into two sexes, 17 age groups (1 to 17+), four areas corresponding to the four fisheries (Chatham Rise, Cook Strait (eastern stock spawning fishery), Sub-Antarctic, and West Coast South Island (western spawning fishery)), and two stocks (Eastern and Western). In the model the non-spawning fishery is split into two parts, separated by the migration of fish from the Chatham Rise to the Sub-Antarctic, giving a total of six fisheries in the model.

2013 stock assessment results

The hoki stock assessment traces the history of stock size (biomass) from a largely unfished state in 1972. International best practice suggests that stocks like hoki should be fished down to a level of about 35–40% B_0. New Zealand uses a target of 35–50% B_0 for hoki. The upper panel (model run 1.7) shows that the eastern stock of hoki has never been below the soft limit, having remained far above the biomass limits that would signify an overfished stock.

The lower panel shows that the western hoki stock began to decline below target levels around the year 2001, but then increased in abundance (when the TACC was cut) and has been fully recovered since 2011.



The current hoki stock assessment model is the result of a long process of evolution through the Working Group process. Changes in model structure and various assumptions have been used over time to improve the fit to the abundance indices including:

- Upweighting of trawl survey data to improve fits (pre 2012)
- Allowing changes in catchability in the Sub-Antarctic research trawl survey (either years 1994-98 or 2005-12) (new in 2012 assessment)
- Natal fidelity or not

- Lack of older fish in the catch at age and research survey data (model solutions have been either killing them off with age varying mortality or hiding them with domed selectivity)
- Recent or long-term recruitment values used for projections
- Fixing parameters at their MPD values in the MCMC runs
- Relative data weighting

These decisions should be considered in the review of the current stock assessment model.

2. Terms of Reference

- An independent panel comprising Drs Doug Butterworth, Jim Ianelli and Richard Hillary will be convened. All the members of the panel have extensive scientific expertise in stock assessment using Bayesian methods.

- Panel members must have no connection with the original work and must declare any actual or possible conflicts of interest that might affect their ability to come to an objective view of the model structure and results and any alternative approaches.

- The lead researcher who completed the last stock assessment, Andy McKenzie, will assist with the review and the Fisheries Stock Assessment Manager, the Chair of the Deepwater Working Group and the Principal Advisor Fisheries Science will be available if required.

- The primary objective for the expert panel is to provide advice to the Ministry for Primary Industries on the quality and reliability of the hoki stock assessment model used to determine stock status for the hoki stocks. Specifically:
    - *Inputs:* the extent to which the data, modelling assumptions, structure, priors, and penalties are appropriate (including assessment of both biological and fishery components)
    - *Implementation*: whether the modelling has been implemented using best practice methods or, where other approaches or approximations have been used, the extent to which these could bias the results
    - *Sensitivities*: the choice and execution of sensitivity runs and the extent to which these can be relied upon to describe the full uncertainty in the modelling results (including an assessment of the likely sensitivity of the results and conclusions to modelling choices not formally assessed through sensitivity trials)
    - *Reliability*: noting that information to inform fisheries management is often uncertain, the extent to which the results can be relied upon as a basis for setting the catch limit for hoki
    - *Improvements to modelling*: any improvements that could (or must) be made to the modelling to increase the reliability of the results for future management decision making – including the potential use of alternative models or model structures
    - *Improvements to data and research*: other key areas of research or data collection that could decrease uncertainty or increase the utility of the modelling for future management decision making

- The expert panel will summarise their findings and any recommendations in a report to the Principal Advisor Fisheries Science, Ministry for Primary Industries. Where consensus cannot be reached by the external reviewers, any differences of opinion should be recorded.

3. Background documents

The following documents will be provided:

- Final reports and published papers describing hoki biology, stock structure and research surveys directed at measuring the relative abundance of hoki stocks
- Description of the commercial fishery and catch at age sampling programmes
- Final reports describing the most recent and previous model runs and results
- Notes and presentations from Hoki WG meetings where modeling results were discussed
- Reports from a previous review of an earlier (1999) version of the model
- Relevant reports, journal articles, papers, or other documents relating to hoki

4. Format for review

The format for the review will be a workshop involving the independent external reviewers, key players and other interested parties in Wellington, New Zealand to discuss the model and results in detail over a period of 5 days. The review will start with a number of presentations to ensure a common understanding of the work (about 1.5 days), and be followed by a period of contemplation, focused discussions with the lead researcher or other parties (at the panel's discretion), and drafting of a report containing conclusions and recommendations (2–3 days). Additional model runs or analyses may be requested by the panel. The review panel will present a draft version of their findings to interested parties on the last day to receive feedback and suggested corrections on matters of fact. The review panel may, at their discretion, reflect such feedback in their report. The aim would be to have a near-final version of the report by the end of the week.

5. Timetable

The workshop is set down for 17–21 February 2014 and will be held in the Main Conference Room, National Institute for Water and Atmospheric Research (NIWA), Greta Point, Wellington, New Zealand. It is anticipated that Pamela Mace, MPI, will chair the open sessions.

Monday 17 Feb	Presentations on the biology, commercial fishery and stock assessment modeling of hoki	Open session
Tuesday 18 Feb a.m.	Presentations conclude	Open session
Tuesday 18 Feb p.m.	Panel confers with individual researchers	Panel's discretion
Wednesday 19 Feb	Panel confers with individuals or works alone	Panel's discretion
Thursday 20 Feb	Panel works on review	Closed session
Friday 21 Feb a.m.	Panel presents draft findings	Open session
Friday 21 Feb p.m.	Panel concludes review	Closed session

It is anticipated that the review can be concluded by 5 pm on Friday 21 February, although final drafting of the report may take place over subsequent days.

Appendix 2. Ministry for Primary Industries Agenda for an independent review of the hoki stock assessment model

Main Conference Centre, NIWA, Greta Point, Evans Bay Parade, Wellington

Chair (of open sessions 1–6 and 10): Pamela Mace, DDI (04) 819 4266, email Pamela.Mace@mpi.govt.nz

AGENDA

Monday 17 February 2014 (starting 09:30)

1. Introductions, apologies, general arrangements for the review – Pamela Mace

   Presentations:

2. Hoki biology –Richard O'Driscoll, Mary Livingston, Rosemary Hurst, Sira Ballara and Peter Horn

3. Hoki: Age determination and catch sampling – Peter Horn

4. Catches, size and age structure from the hoki fishery – Sira Ballara

5. Hoki surveys – Richard O'Driscoll, Neil Bagley, Darren Stevens, Sira Ballara and Peter Horn

Tuesday 18 February 2014 (starting 09:30 or earlier as agreed)

   Presentations continue:

6. Stock assessment modeling (including data inputs, assumptions, priors, bounds, fits, diagnostics, sensitivities, MCMCs, projections, and criteria for assessing performance) – Andy McKenzie

   Panel in session

7. Panel discussions with presenters or others at their discretion

Wednesday 19 February 2014

   Panel in session

8. Panel discussions with presenters or others at their discretion

Thursday 20 February 2014

   Panel in session

9. Panel probably in closed session (but may have discussions with presenters or others at their discretion)

Friday 21 February 2014 (starting 09:30 or earlier as agreed)

Conclusions and recommendations

10. Panel presents findings and conclusions to interested parties for general impressions and corrections on matters of fact (open session)

11. Panel concludes deliberations (closed session)

It is anticipated that the open session on Friday 21 February will be completed by 1 pm.

Appendix 3. List of Participants in Open Sessions

Independent Expert Review Panel

Doug Butterworth	University of Cape Town, South Africa
Richard Hillary	CSIRO, Australia
Jim Ianelli	NOAA Fisheries, USA

Participants

Pamela Mace (Chair)	MPI
Neil Bagley	NIWA
Sira Ballara	NIWA
Tiffany Bock	MPI
George Clement	Deepwater Group
Patrick Cordue	Innovative Solutions Limited
Paul Crozier	WWF New Zealand
Ian Doonan	NIWA
Alistair Dunn	NIWA
Charles Edwards	NIWA
Jack Fenaughty	Sanford Ltd.
Dan Fu	NIWA
Vivian Haist	Consultant
Peter Horn	NIWA
Rosemary Hurst	NIWA
Aaron Irving	Deepwater Group
Andy McKenzie	NIWA
David Middleton	Seafood New Zealand
Richard O'Driscoll	NIWA
Vicky Reeve	MPI
Marie-Julie Roux	NIWA
Paul Starr	Deepwater Group
Kevin Stokes	Consultant
Kevin Sullivan	MPI
Geoff Tingley	MPI
D'Arcy Webber	VUW