Fisheries Occasional Publication No. 99, 2011

Review of the Western Australian Rock Lobster Stock Assessment

Report to Western Australian Department of Fisheries

by André E. Punt, University of Washington

May 6, 2011

Government of **Western Australia**
Department of **Fisheries**

Western Australian Fisheries
and Marine Research Laboratories

Fish for the future

Department of Fisheries
3rd floor SGIO Atrium
168-170 St George's Terrace
PERTH WA 6000
Telephone: (08) 9482 7333
Facsimile: (08) 9482 7389
Website: www.fish.wa.gov.au
ABN: 55 689 794 771

Published by Department of Fisheries, Perth, Western Australia.
Fisheries Occasional Publication No. 99, August 2011.
ISSN: 1447 - 2058 ISBN: 978-1-921845-22-2

Fisheries Occasional Publication No. 99, 2011

Review of the Western Australian Rock Lobster Stock Assessment

Report to Western Australian Department of Fisheries

by André E. Punt, University of Washington

May 6, 2011

Government of **Western Australia**
Department of **Fisheries**

Fisheries Research Division
Western Australian Fisheries and Marine Research Laboratories
PO Box 20 NORTH BEACH, Western Australia 6920

Fish for the future

Review of the Western Australian Rock Lobster Stock Assessment
Report to Western Australian Department of Fisheries

by

André E. Punt, University of Washington
May 6, 2011

EXECUTIVE SUMMARY

- The structure of the assessment model is appropriate for evaluating stock status relative to limit and threshold reference points for the three Zones. It is also appropriate given the data available on abundance, movement, growth, and catch and population length-compositions.
- The documentation of the model is vastly improved compared to that reviewed in May 2010. However, a number of gaps in the documentation remain (e.g. the basis for many of the pre-specified parameters and priors). I have no reason to believe that the values assumed for these quantities are incorrect, but the lack of justification remains a flaw.
- The assessment document and associated diagnostics focus exclusively on a best (or base/reference) case model, with uncertainty quantified in terms of standard errors for model outputs. However, prior to the assessment being finalized, it will be necessary to assess the sensitivity of the results to varying *inter alia* data weightings, values for pre-specified parameters, and specifications for priors. Likelihood profiles should also be used to explore the sensitivity of key model outputs to the values for parameters which would be expected to impact yield-related outputs (e.g. natural mortality).
- The breeding stock indices in all three Zones are assessed to be larger than the threshold and limit values with high probability. However, the measures of precision provided only pertain to uncertainty related to estimation error and ignore that due to model-structure, data-set choice and the values assumed for pre-specified parameters. Sensitivity tests, and retrospective analyses, should be used to examine the extent to which asymptotic confidence intervals represent the actual level of uncertainty.
- This report highlights a number of other recommendations. These are divided into recommendations that I would expect to see completed before the assessment is finalized and longer-term recommendations.

A. Introduction and structure of the review

In 2000, the West Australian Rock Lobster (WRL) fishery became the first fishery in the world to be certified by the Marine Stewardship Council (MSC). At the annual audit in December 2010, the MSC's independent certifying body, Scientific Certification Systems, set a number of conditions for ongoing certification. The most important condition under Principle 1, Stock Assessment, was to have an international stock assessment and modelling expert undertake a review of the latest version of the WRL Stock Assessment Report that has been used to provide advice for the management of the fishery, and 2010 stock assessment and 2010/11 management settings,

The stock assessment reviewed in this report considers the three Zones (A, B and C), with each Zone divided into a number of regions (3 in Zone A, 4 in Zone B, 4 in Zone C). The management objectives are specified at the Zone-level. The model operates from November 1974 to June 2010, and considers 11 within-year time-steps to capture key aspects of the dynamics of the population (e.g. fishing, movement, growth). The model is effort-conditioned in that effort is assumed to be measured without error and the fit to the historical catches is included in the objective function. The values for the parameters of the model are estimated by maximizing an objective function which consists of the likelihoods associated with a variety of data sources (catch, length-composition, tagging etc.), as well as a set of penalties (priors).

This review involved first reading the stock assessment document, a document listing a variety of diagnostic statistics, and the ADMB code on which the assessment is based. Although desirable, I made no attempt to recode the model based on the provided specifications (this would not have been possible because the documentation is incomplete), nor did I evaluate the assessment method by applying it to simulated data sets. Both of these checks should occur once the assessment has "stabilized". My initial review identified a number of queries regarding issues which were unclear to me from the information provided. I therefore contacted Drs Simon de Lestang and Peter Stephenson (WA Fisheries) who responded to my queries by email. The responses to my queries led to revised versions of the document listing diagnostic statistics as well as to Section 5.5 of the assessment report.

This report is structured around the following Terms of Reference:

1. Review the Western Rock Lobster Fishery (WRLF) Stock Assessment Report and Model and comment on its robustness and make recommendations regarding how it could be enhanced.
2. Evaluate and 'audit' the WRLF Stock Assessment Report and Model against the [short-term] recommendations made by the May 2010 Review Panel.
3. Evaluate the 2010 WRL stock assessment / state of the stock and review the WRLF Stock Assessment and Management Settings for 2010/11

Section B of the report pertains to TOR 1 by commenting on the structure of the model, the data utilized, and the diagnostics provided, Section C conducts an 'audit' relative to the short-term recommendations from the May 2010 Review Panel, and Section D comments on the state of the stock relative to the management settings for 2011. In relation to TOR 3, this report does not comment on the appropriateness of the decision rules and reference points generally, but rather focuses on those aspects which relate to the scientific practice of stock assessment. The final section (E) lists a number of minor errors found in the assessment document. Each section is structured by first summarizing a number of 'findings' based on my examination of the available information and then listing a set of recommendations. Those recommendations I would expect to see implemented before the assessment is finalized are indicated by asterisks.

I would like to thank Drs de Lestang and Stephenson for their prompt (and very thorough) responses to my queries for clarification.

B. General comments on the stock assessment

B.1 Model structure and specification.

B.1.1 Findings

The selection of a mathematical model as the basis for a stock assessment involves a trade-off between realism (does the model capture the main features which drive the dynamics of the assessed population) and pragmatism (is the model structured to make use of the available data; are data available which can be used to estimate the values for the parameters of the model; does the model produce the key quantities on which management decision making is based). Most of the models on which assessments for hard-to-age species (crab, rock lobsters, and abalone) are based are length-structured and several explicitly consider the spatial dynamics of the assessed population. The population dynamics model selected to model the dynamics of western rock lobster (sex-, region-, and length-structured) is therefore appropriate and comparable with those used for other similar species. The model is simplified compared to that provided to the May 2010 Review Panel and reasonable amounts of data are available for each region.

The assessment team explored a number of model configurations during the development of the assessment (e.g. switching on various parameters), although the assessment document does not refer to this. A major weakness of the current assessment is that while results are provided for a base-case model, it is not possible at present to fully understand the sensitivity of the results from the assessment to changes to key assumptions.

Movement is a key feature of the assessment model. The model assumes that only "whites" migrate and that the proportion of the migrating whites which go to each region is known (Tables 5.7a,b). I was advised by Dr de Lestang that one of the next steps of model development pertains to estimating the proportion of whites moving to each cell given that they move (these proportions are currently pre-specified; Tables 5.7a,b) and I fully support this development. Although the structure of the model is adequate, it will take some time for the assessment team to fully explore all of the available options. In principle, the migration rates should be year-specific (report page 33). However, in the absence of large quantities of tagging data, attempting to estimate year-specific movement is not justified.

B.1.2. Recommendations

1. The report provides what I would consider a base-case model (best settings for fixed parameters, data set choices and weights). However, I was expecting to see a table listing how sensitive key model outputs were to, for example, changing weights, values for fixed parameters, specifications for priors, etc. While the base-case model is adequate for decision making, including results of sensitivities will help managers understand the uncertainty associated with the model outcomes, in particular because asymptotic confidence intervals often under-estimate the true extent of uncertainty. ***Add a table listing key assumptions, weights and fixed values for model parameters and explore sensitivity to varying these assumptions, weights and values.**

2. Estimating growth within the model should be possible in principle. However, Dr de Lestang advised me that attempts to do this led to unrealistic outcomes so the parameters determining growth were pre-specified (Table 5.5.2). Although this is a reasonable approach for the present assessment, the reasons for the unrealistic estimates need to be determined (e.g. are some of the data contradictory). Table 5.5.2 lists the values for the parameters which determine the mean growth increments, with the cv for the growth increment set to 0.02. It is, however, well known that the outcomes from size-structured models are very sensitive to the variance in the growth increment (a key reason for estimating growth within an assessment model). ***Explore sensitivity to the cv of the growth increment before the assessment is finalized.**
3. There are many pre-specified parameters (e.g. Table 5.8). However, while references are provided for how some of these values were obtained (e.g. $M^{I}_{s,r}$, $I^{50}_{s,r}$ and $S^{I}_{s,r}$), this is not case for all of the parameters (some of the parameters of the model, e.g. β_r were not included in version of Table 5.8 sent to me originally). In discussions, with Dr de Lestang, it does appear that there is a logical basis for the values for the pre-specified parameters. However, the assessment document should be updated to ensure that the basis for all of the parameters is clear (if necessary short appendices could be added to the document). ***The final report must list all of the pre-specified values and how each was obtained. Furthermore, the impact of "guestimated' parameters on key model outputs should explored in tests of sensitivity.**
4. The model description is incomplete. I list a few minor issues in Section E below and some major issues here:
    a. The approach used to include the tagging data in the likelihood function is reasonable. However, it is poorly described in the assessment report (Equation 26). As written, this description does not match what is actually done (subroutine MoveTagLike), which involves taking each 'tag group' (animals of a given sex and size released in a given region at a particular time and recaptured at the same time later) and computing the relative probability of being recaptured in each region at the time they were recaptured (i.e. a recapture conditioned likelihood).
    b. The approach used to fit the IBSS data (which should really be referred to as fishery-independent surveys for animals 75mm and larger) differs between the documentation, which indicates that a survey selectivity pattern, $\tilde{S}_L$, is estimated, and the code which assumes selectivity is given by q^{A}_{L} for animals in size-classes 16+ [zero below this].
    c. The rationale for Equation 15 (which appears to relate to residual patterns) needs to be added.
    d. The code includes specifications for the pre-fishery years, but this information is missing from the document.
    e. According to the code $\hat{P}^{\Lambda}_{r,y}$ and $\sigma^{\Lambda}_{r,y}$ are linear functions of time. The justification for this is not provided.
    f. There is an adjustment to 2008 effort, but this is neither documented nor justified.

***The documentation should be checked and updated as needed.**

5. Equation 19 specifies the model-predicted length-composition. However, no account is taken in this equation of fishery selectivity (perhaps it is uniform for all the recruited size-classes).
6. I identified a few (fairly) minor coding errors, as outlined below:
    a. Some of the prior distributions appear to have been coded incorrectly.
    b. σ_R^2 is added to the variance of the puerulus data twice.
7. The model allows for 10% discarding starting in 2010/11. ***While the equation for the discard mortality is correct, the equation for the landed fishing mortality should be multiplied by $1-\varpi_y$ otherwise, allowance for discarding also implies a change in catchability.**
8. The selectivity pattern for the fishery is pre-specified (Equation 16a,b). Some of the residual patterns could be improved by estimating rather than pre-specifying selectivity.
9. The model is considerably simpler than earlier versions, which has enhanced the ability of the assessment team to implement and test the model. However, there are still very many (46) size-classes. Further analyses should be conducted to assess whether it is possible to use fewer size-classes (perhaps 2 mm size-classes for undersized animals near the size limit and wider size-classes for the larger animals which are relatively uncommon in catches and the population).

B.2 Data utilized

B.2.1 Findings

A large number of data sets are available for western rock lobster. A major concern of the May 2010 Review Panel was that the use of some of these data sets were inconsistent (e.g. used multiple times or analyzed using a variety of inconsistent assumptions). The current assessment is a vast improvement in this regard. The major sources of data and my understanding of their use in the assessment are:

- Puerulus data – these data are used as indices of recruitment.
- Catch monitoring:
    - Fishery-dependent breeding stock indices – not used when fitting the model, but considered as a diagnostic check on the results.
    - Under-sized abundance – not used when the fitting model owing to a lack of data to estimate the relevant selectivity pattern.
    - Length-composition data – used when fitting the model as the length-frequencies associated with lobsters >75 mm.
- Independent breeding survey:
    - Fishery-independent Breeding Stock Indices (IBSS) – not used directly in the assessment; rather the data from this survey are used to construct the CPUE for lobsters $\geq$ 75mm.
    - IBSS length-composition data – used when fitting the model as the length-frequencies associated with lobsters ≥ 75mm.
- Volunteer research log-books
    - Catch and effort data – these data are used to pro-rata the compulsory catch and effort data to depth and half-month time step.
    - Catch rates of spawners, setose and the numbers of females returned to sea – not used in the stock assessment model
    - Catch of undersize lobsters – not used in the stock assessment model
    - Daily index of swell height – not used in the stock assessment model

    - Octopus catches – not used in the stock assessment model
- Compulsory catch and effort data:
    - Catch – used when fitting the model
    - Effort – used to determine fishing mortality
- Recreational fishery surveys:
    - Estimates of recreational catch – not used in the stock assessment model

B.2.2 Recommendations

1. ***Include a table in the assessment which lists *all* of the data sources and how each is used (or not used).**
2. Include diagnostic statistics for the GLM analyses used to analyze spatial data.
3. While a relatively minor component of the total catch, it is nevertheless important to extend the model to allow the data from the recreational fishery to be included in the assessment. The projection component of the model will also need to be modified to account for changes in recreational fishing practices into the future.
4. Consideration should be given to implementing a model in which the catches (rather than effort) are assumed to be known exactly.

B.3 Diagnostic statistics

B.3.1 Findings

A number of diagnostic statistics were included in the supplementary material provided to me. The model is effort-conditioned and is fitted to a number of data sets: (a) catches by time-step and region (b) puerulus indices of deviations in recruitment from mean recruitment by region, (c) IBSS indices by region, (d) tag-recapture data, and (e) length-compositions from the fishery and the IBSS survey.

The fits to the various data sources are generally very good. I have the following observations:

1. The fits to the catch data are good. However, there is some evidence for model-misspecification (periods of positive then negative residuals), especially for region 2. If the residuals are serially correlated, consideration should be given to estimating annual effort deviations which should reduce this problem.
2. The fits to the puerulus indices are extremely good. In fact, the fits seem to be much better than would be expected given the sampling standard deviations and the assumed level of extra variation. The accuracy with which puerulus recruitment can predict future catches indicates that they are generally good indices on incoming year class strength (de Lestang et al., 2009). Either the puerulus indices are remarkably good indices of incoming year-classes (unlikely) or these data are dominating the fit the other data sources.
3. Although there is a relationship between the IBSS data and the model-predictions, it is not clear how good this relationship is.
4. The assessment report notes that the IBSS data and modeled levels of egg production exhibit very similar trends (page 150), but this is not clear from the trajectories provided (especially for regions 1 and 4).
5. The diagnostics showing the ability to mimic the tagging data are important but unclear. The colors on the plots do not match those in the captions. Also, the proportions should all add to 1, but this appears not to be the case for the red circles. Finally, some measure of the expected precision of the proportion

recaptured by region should be added to the plots to enable a more thorough evaluation to be conducted.

6. The model generally fits the length-frequencies adequately. However, there are some combinations of sex, time-step and region for which the fits are relatively poor (e.g. the fits to the length-frequency data for time-step 10 for sex 1). Also, the proliferation of negative residuals for the first size-class for the commercial length-frequency data highlights the need to consider estimating fishery selectivity (see recommendation B.8 above).
7. The assessment document notes that the model-estimated trends in efficiency are similar to those inferred from other information (page 149). However, it is not clear how this comparison was made. Also, the change in efficiency in regions 10 and 11 is very large and attempts should be made to compare this with independent estimates of efficiency increase.

B.3.2 Recommendations

1. ***Enhance the plots of observed versus model predicted catches by including the estimated levels of uncertainty about the data, and by reporting the results of runs tests**.
2. ***Augment the plots of residuals versus season by checking for lack of homescedactity and adding a loess smoother to better assess possible trends in residuals. There would be value in plotting the residuals against the fitted values.**
3. ***Multiply the model-predictions by survey catchability and show a 1-1 line to improve the ability to evaluate the model fit to these data. Also, show the fits to the IBSS indices as time-series.**
4. ***Conduct sensitivity tests in which the IBSS data are down-weighted or simply ignored**.
5. ***Improve the diagnostics used to summarize the ability to mimic the tagging data.**
6. ***Conduct sensitivity tests to explore the impact of changes to the weights assigned to the puerulus indices to examine whether changing these weights impacts the fits to the catch and length-composition data.**
7. The puerulus indices appear to be correlated with temperature. There may be value in considering a temperature covariate in the future version of the model.
8. The growth model (see Table 5.5.2 for parameter values) is based on fits to tag-recapture data. However, very little information is provided regarding the fit of this model. ***Table 5.5.2 should be annotated by measures of precision and the observed and model-predicted growth increments reported. The results of the fits should be used to justify the assumed cv for growth increment of 0.02 (page 138)**.
9. The proportion of whites is assumed to change linearly over time (page 135). However, no data are presented to justify this assumption.
10. The data on undersize animals could be added to the likelihood function if selectivity is estimated.
11. ***Likelihood profiles should be used to explore the sensitivity of key model outputs to the values for parameters which would be expected to impact yield-related outputs (e.g. natural mortality).**
12. ***Retrospective analyses (leaving out one then two, etc. years of data and then re-running the assessment) should be conducted.** These allow the

stability of the results to be assessed, along with a measure of the forecast skill of the model (by projecting forward with the known effort and seeing whether reality was in the confidence intervals).

C. Responses to the recommendations from the 2010 Review Panel

I requested that the assessment team provide a brief summary of progress against each of the short-term recommendations made by the 2010 Review Panel. The responses from Dr Simon de Lestang are listed in italics below each recommendation. In general progress against the 2010 recommendations is adequate. Where I have further comment on one of these recommendations, it is listed in bold below the response from the assessment team. The 2010 Review Panel also identified a number of 'long-term' recommendations. It would not be expected that many of these would be addressed by now and I have not commented on them.

1. The assessment team should critically examine the new data inputs and ADMB code for the new ITQ model to verify that the intended revised model structure is correctly implemented. This is likely to be facilitated by revising the mathematical description that summarises the equations (Appendix G) that are implemented in the revised ADMB code. Errors that are detected should be corrected.
   This was done. Both Peter and I stepped through all functions in the code 1. to understand exactly what the code was doing and 2. to determine whether what the code was doing was appropriate. A number of small errors were detected and corrected. One significant problem was identified. The initial conditions in the new model did not appear to be working well for areas with little/no effort in the mid-late 1970s (i.e. the start of the model). A significant amount of time (months) has been put into recoding the initial conditions.
   Further checking of the code is still required (as I note above). The code would also benefit from having sections annotated by the equations in the report – as I note above, the documentation of the model is perhaps the one of largest issues of concern with the assessment. Annotating the code with the equations will also help to ensure that the model is fully and adequately documented in the assessment report. An attempt should also be made to remove "legacy" code (generally code used when developing testing the program).
2. The assessment team should confirm that, when using different initial values of parameters, the new ITQ model converges to the same parameter estimates, e.g. through a jitter analysis. Details of the analysis and the results should be documented. Problems in obtaining convergence should be addressed.
   Jitter analysis was not conducted until it was felt that the model (including adjusted initial conditions) was functioning properly. Starting parameters of the new ITQ model have been systematically adjusted and their impact on the model convergence recorded. The model appeared fairly robust to the jitter analysis. The details of this have not been documented because model development has continued (eg. new migration code). Details from a jitter analysis will be conducted and results documented on the version of the model that is to be used for the coming quota setting process / MSC review. This model version will include suggested changes from this (your current) review.
   The assessment report should include a summary of the outcome of the 'jitter tests'. For example, a plot could show the value of the objective function at its minimum along with key model outputs as y-values against the number of the

jitter test. This will allow 'major' sensitivity to initial values to be distinguished from 'minor' sensitivities.

3. Computer software should be developed to automate the production of detailed and complete model and diagnostic outputs, e.g. tables of indicator variables and reference points, plots of predicted versus observed values, results of residual analysis, bubble plots of Pearson residuals for length composition data, etc. Such software will assist in reporting the details of the assessment and the results and will facilitate evaluation of the model's integrity.
   R code has been developed to automate the production of detailed diagnostics and model outputs.
4. The assessment team should identify key indicator variables and reference points that should be considered when comparing the results of sensitivity tests and which will be required by the decision rules and by managers and fishers seeking to assess the management implications of model results. These indicator variables will include results of model projections. If necessary, the ADMB code should be modified to calculate and output the required statistics.
   The R diagnostic code produces key indicator variables and reference points used to compare the results of different model runs. Indicator variables include estimates of residual biomass and breeding stock with threshold values being based on the mean 1980s values of breeding stock in each region
   A table summarizing the values for these key indicator variables for different model runs (e.g. weights for data sources and values for pre-specified inputs) should be included in the assessment report.
5. The revised model should be run, with the results being accepted as the current base case. Details of the analysis, the parameter estimates, asymptotic standard errors, parameter correlation matrix, and detailed diagnostic outputs should be documented and examined. Errors detected in the results should be resolved by correcting the code and the model description. The assumptions of the model may need to be modified and a new base case generated if the diagnostics identify a major structural uncertainty with the model. Sensitivity analyses identified below may assist in determining how the model may need to be restructured.
   This has been the process conducted by Peter, Norm Hall and myself since the last review.
6. The assessment team should develop an explicit list of key uncertainties in model assumptions that reflect uncertainties in model structure and that should be considered as alternative cases. A critical assessment of the diagnostic outputs for the base case model is likely to assist in identifying aspects of the model requiring exploration though sensitivity tests.
   An explicit list of uncertainties has not been developed as the model is still being re-structured.
   While it is reasonable not to have a complete set of runs at present, this will be a key task prior to completion of the assessment. At present, it is not possible to fully understand the sensitivity of the model outputs to plausible variations to some of the specifications of the assessment.
7. The model should be re-run to explore the sensitivity of results to each of the alternative model cases. Results of the alternative models should be compared with the results from the base case model. The contributions of components to the likelihood should be used to develop likelihood profiles for selected key input parameters, such as natural mortality.
   As the base case model is yet to be finalised, this is not possible.

See comment above.

8. Results of the above analyses should be critically assessed to determine whether the results from the base case model are appropriate for determining the status of the fishery and advising on the appropriateness of alternative management strategies. The results of the diagnostic plots will be crucial for this assessment, and may identify deficiencies in model structure that must be addressed before model results can be considered reliable (see Point 5, above). The results from the sensitivity runs should be used to provide information on the uncertainty of the results associated with model structure and data inputs if the base-case model is considered acceptable. The results of this assessment should be documented.
   As the base case model is yet to be finalised, this is not possible.
9. The current value of a target reference point, based on current and projected costs and prices and using an egg production or harvest rate proxy to an "equilibrium" MEY level, should be determined for use in the decision rules to be applied to the fishery.
   An CRC project is providing the funding to determine target reference points. This work is underway.
10. The results of the base case model, and the alternative models, should be considered in the context of the decision rules for the fishery to determine whether the current management regime adopted for the fishery is adequate to maintain the egg production of the stock at a level consistent with the requirements of the decision rules, and likely to be consistent with regulations required to attain the target reference point.
    The results from a range of model runs will be considered in the context of the decision rules.
    See the recommendation regarding projections in the absence of exploitation under D.2 below.
11. A more detailed and clearer description of the model should be published.
    A more detailed and clearer description of the model has been published on the Depts website within the Draft Stock Assessment Report. This report will be finalised following your current review (your suggested changes included). This will then be "locked down" and published as research Stock Assessment Report # 217.
    As noted above, the model description is markedly improved compared to that reviewed in May 2010. However, some errors were found during this review and the basis for many of the pre-specified parameters and the priors is not clear from the documentation provided.
12. To reduce model dimensionality, the model should be further revised to merge the regions that, on the basis of correlations among regions of, for example, catch-rates and puerulus settlement data, appear to be candidates for combination. Model results before and after merging regions should be compared to determine whether the simplified model improves the robustness of the results and to confirm that the model revision has had no unintended consequences.
    The dimensionality of the model has been reduced based on correlations between regions and between timesteps. Sixteen regions have not been reduced to 11 regions and 24 timesteps have now been reduced to 11 timesteps. This has dramatically increased the running speed of the model while not negatively affecting the model fit.
    Analyses should be conducted to assess whether it is possible to implement the model with fewer size classes.

13. A version control system, such as TortoiseSVN, should be implemented.
    A version control system has been implemented.

D. Management settings for 2010/11

D.1 Findings

The [current] proposed management objective is "Ensure that the egg production in each Zone of the fishery remains above its threshold level and the probability of still being above this level in five years time is at least 75%", where the thresholds are based on the estimated egg production in the early 1980s for Zones B and C and the mid-1980s for Zone A. These levels are considered to be approximately 20% of the unfished level. However, this inference is not based on the current assessment model. As noted below, the level of egg production corresponding to the threshold and limit levels can be explored using the new assessment model by projecting it forward under zero catch.

The basic structure of the decision rule framework is consistent with world's best practice because it includes 'indicators', and 'reference values', and because the decision rule structure accounts for uncertainty. However, the decision rule framework is not a formal management strategy because while general direction is provided regarding what would occur if stock status was in various categories (e.g. below the limit reference point), the specific management actions that would be implemented are not fully specified. The management rules would need to be formulated more clearly (e.g. for example, what "No management action is required" means on page 172 is undefined – it would seem to be "Keep the TACC at its current level and do not modify input controls") for the performance of the system to be evaluated using Management Strategy Evaluation. As implied by the May 2010 Review Panel, it not currently clear how the various components of the framework interact and the net outcome of applying this framework.

I support the use of the assessment model to produce the measures on which the decision rules are based because this model integrates all [most] of the available information, it allows quantities of management interest (e.g. the size of the breeding stock) to be forecasted, and it should produce a more precise outcomes (less likely to trigger reference points and hence management action due to sampling error).

A system of target reference points based on the breeding stock abundance corresponding to Maximum Economic Yield (MEY) is being contemplated. However, insufficient information was available for me to comment on the scientific issues associated with such reference points. It was somewhat unclear on how the MEY target (usually defined as a long-term biomass) would depend on "abundance" (except in the broadest sense) (page 167).

The model-derived breeding stock indices all currently exceed the threshold (and limit) reference points with high probability and it is projected that this will remain the case for the foreseeable future under the current catch of 5,500t.

D.2 Recommendations

1. ***The plots of residual biomass and egg production are meaningful. However, they should be accompanied by plots of residual biomass and egg production if there had been no fishing at all. This would allow the impact of fishing on the population (and the state of the population corresponding to the limit reference point) to be assessed.**
2. ***Harvest rate is presented, but it was not clear how harvest rate is defined, especially when presented at the Zone level.**

3. ***The fishery is now managed on the basis of output controls. The projections should have included estimates of catch-rate in addition to harvest rate and egg production (with associated uncertainty). Predicted catch-rates would have also enhanced the information available to decision makers.**
4. The decision rules will need to be fully-specified to enable them to be evaluated using Management Strategy Evaluation.

E. Minor Corrections

1. λ_r (Page 138) is the annual proportion of 'whites' which migrate from region r.
2. $\Lambda^s_{r,y,r,l}$ (page 138) should be $\Lambda^s_{r,y,t,l}$.
3. $v_{r',r}$ is the proportion of the 'whites' which migrate from region r' which migrate to region r.
4. There should be a summation over L in Equations 18 and 20.
5. The IBSS index referred to in Equations 23-25 is not really an index of breeding animals, but rather an index of animals 75 mm and higher. The text should be modified accordingly.
6. Efficiency increase does not depend of lobster colour so $\theta^W_{r,y,t}$ and $\theta^R_{r,y,t}$ in Equation 9 should be replaced by $\theta_{r,y,t}$.
7. The definition of $T_{r,y,t}$ should mention that temperature is constrained to be never less than 18.3^{0}C.
8. $\overline{R}_r$ in Equation 26 should be R_r (although Equation 26 does not match the code, which involves the puerulus counts being normalized).
9. Equation 6 refers to M^τ, but needs to refer to M^τ multiplied by the length of the time-step
10. Equation 12. $\theta^P_{\tilde{r}}$ should be $\theta^P_{\tilde{r},y}$. Also, $\theta^T_{\tilde{r}}$ and $\theta^Z_{\tilde{r}}$ should be $\theta^S_{\tilde{r}}$ and $\theta^D_{\tilde{r}}$
11. Equation 9. ϖ should have a y-subscript.
12. Equation 15 should be checked to ensure that it matches the code. There appear to be differences between what is specified here and what appears in the code.