Consultant Report on:

STAR Panel for Coastwide Lingcod and Widow Rockfish Assessments

5-9 June 2000, Santa Rosa, California

Paul A. Medley
Sunny View
Main Street
Alne, YO61 1RT
UK
paul.medley@virgin.net

Contents

1 OVERVIEW

The following are some personal observations on future improvements in the assessments for widow rockfish and lingcod and are not necessarily shared by the other panel members. This report does not repeat issues raised in the main STAR panel report, which needs to be consulted for the panel's views on these assessments.

Overall I felt the assessments were well conducted. The STAT teams were very helpful and open in explaining the assessment and the problems they had with it. This led to a constructive review and a number of significant improvements in the assessments.

A number of problems arose which could not be resolved within the time frame of the review. These were mainly problems related to data and its interpretation. While the specific problems are in the main report, I expand some more broadly here.

Some analyses were carried out during the review by panel members to help elucidate the problems. I include a summary of results below from my analysis of a widow rockfish bycatch abundance index.

2 INDICES OF ABUNDANCE

In common with many stock assessments, both lingcod and widow rockfish suffered from having no good indices of abundance. This can be dealt with through both improvements in data collection and improvements in the models linking the index to population size. Data collection, such as acoustic surveys, are a longer term solution. Models, such as generalized linear models (GLMs), can be used to make corrections to raw indices, producing better abundance indices that have a better theoretical basis.

A significant problem with a number of the indices is that they have been affected by management controls making the time series incomplete. The effect of the controls may be removed by models. For example, trip limits on catches (e.g. Oregon logbooks) censor larger values for CPUE as catches larger than the maximum are discarded. Models attempting to remove these effects should be tested using simulated data.

Another example where models could improve the assessment is in the case of excluding zeros. Tows containing zero catches of the species were not included due to the difficulties of modelling zeros. Unfortunately this excludes useful information, since the proportion of zero tows of a trawl is probably related to the species abundance. A model could be developed which makes use of these values.

2.1 Recreational Fisheries

Lingcod is subject to significant fishing mortality from the recreational fishery. In common with many other recreational fisheries, the data provided by the fishery is minimal. This has arisen because the recreational fishery catch was considered insignificant. This is no longer the case in many fisheries, and lack of information considerably adds to uncertainty in the stock assessment. Another fishery which has a similar problem is Florida spiny lobster where recreational catches are unreported and their scale or trend is simply not known. There is a need to improve public awareness and use education targeting the recreational fisheries to obtain widespread co-operation on data collection programmes. Not only would this allow the impact of recreational fishing to be measured and controlled, it may also lead to much improved indices of abundance than those available from commercial fishing.

2.2 Reef by Reef Depletions

In the case of lingcod, recreational fishery data were discussed as a potential index of abundance. The data were available at a resolution which would allow the analysis on a reef-by-reef basis. This has some advantages as the data only have to be complete for individual reefs and not the fishery as a whole. However, the population is unlikely to be closed and so fish movement may have to be accounted for. The advantage of analyses on small time and spatial scales is that they can provide useful information for the whole stock. Estimates of reef populations can either be used as indices of abundance or as estimates of average fish density in their own right. This approach was used spiny lobster[1], where fishing activities over a short time were used as a recruitment index for the whole fishery. However, the recruitment index was dependent on the life history characteristics of spiny lobster. The success of this approach is likely to depend on exploiting life history or behavioural characteristics of lingcod in much the same way.

2.3 Generalised Linear Model of Pacific Hake Widow Rockfish By-catch

The indices in the previous assessment had used the widow rockfish bycatch for the Pacific hake fishery as a proportion of the catch times the annual hake biomass taken from the Pacific hake stock assessment. It had been suggested that the widow rockfish catch per trawl minute would be a better index. Concerns were also expressed regarding the latitudinal distribution of the hake trawls, and their effect on the index. To test these ideas, a preliminary analysis of the hake and widow bycatch data was undertaken using a generalized linear model (GLM) approach. The aim was to see whether an improved index might be available through an analysis.

---

[1] Medley PAH and Ninnes CH (1997) A recruitment index and population model for spiny lobster (*Panulirus argus*) using catch and effort data. *Canadian Journal of Fisheries and Aquatic Sciences* 54: 1414-1421.

2.3.1 Method

To make the analysis tractable, the data was summarised in a matrix consisting of sums of the variables widow catch, hake catch, haul minutes, counts of observations, and the standard deviations of these variables broken down by year, fleet and latitude. The latitude bands were South of 43^{o}N, between 43^{o}N and 46^{o}N (Heceta Bank) and North of 46^{o}N (Astoria to Cape Flattery). Effort was unevenly distributed across bands with the middle band being the most heavily fished (45%), and the south least fished (21%).

As the CPUE rather than catch was being modelled, the best weighting method was the number of observations rather than a weight based on the standard deviation. It was also assumed the errors were normal, so the parameters were estimated using least squares. This is probably reasonable where the number of observations was reasonably large within a cell.

$$CPUE = Exp\left(Year + Lat + Fleet + \beta\left(CPUE_{HAKE} - qB_Y\right)\right)$$

where $CPUE$ = widow catch per minute haul, $Year$ is the year term used as an index, Lat is the latitudinal term for north, middle and south zones, and $Fleet$ is a term to account for the difference between the foreign and joint venture/domestic fleets. There was only one year overlap between the joint venture and domestic fleets which also had close to the same CPUE in that year, a particularly low value compared to all other years for the domestic fleet. There was no overlap between the domestic and foreign fleets. Therefore given only one anomalous year in common between the time series, the difference in CPUE could not be estimated between these and the change had to be treated as a $Year$ effect.

The inclusion of hake CPUE required adjustment to remove the correlation with the hake biomass, using a least squares estimate of catchability (q). By using the residuals rather than raw CPUE we assume that the difference between the two is driven by changes in catchability unrelated in population size. For example, environmental changes may make catch rates for hake lower than expected in any particular year. If the same effects apply to the bycatch, then the anomaly can be removed from the widow rockfish index by including the hake residuals in the model.

2.3.2 Results

The raw data, as widow per minute haul for the central region was used in the base run of the model. The series is flat but increases towards the end of the series (Figure 1). Including the other regions and a term to allow for differences in latitude, a similar pattern occurs in the year terms of the model. However, the terms produce a smoother time series due to the exponent (the model is multiplicative) and taking means across latitude bands. The time series is flat over its whole range. In contrast, including the hake catch rate residuals induces a negative trend because the anomalies for hake have been positive in recent years. This is similar to the trend from the assessment based on logbook CPUE for the bottom trawl data.

The series 1991-present is not necessarily on the same scale as the series prior to 1991. There is only one year overlap between the domestic and joint venture fleets, with the domestic fleet making up the 1990-2000 time series. The year 1990, although consistent with the joint venture fleet, appears very low compared to the remaining time series. The data point consists of very few observations and was clearly at a point where the industry changed, so may be due to changes in reporting, for instance, rather than underlying changes in stock size. Whereas 1991 onwards may be considered a separate index, there appears no significant differences between the foreign and joint venture fleets.

Table 1 Analysis of variance table for GLM fit to widow rockfish bycatch CPUE within the Pacific hake fishery. Approximate log likelihood ratios (F) indicating the relative importance of the different terms.

Term	SS	df	MS	F
Year	3928.5	23	170.81	7.08
Fleet	7.2	1	7.18	0.30
Lat	831.2	2	415.62	17.22
Hake Residuals	389.2	1	389.17	16.13
Error	1158.3	48	24.13	

The difference between the foreign and joint venture fleet appears unimportant (Table 1). The remaining terms for the year, latitude and hake residuals appear to explain significant variation in the index series. A formal statistical test would probably indicate these terms should be included in the index model.

2.3.3 Conclusion

The results would suggest the index requires further research as it can currently be interpreted as a potential upward, stable or downward trend depending on the index model used. The inclusion of the hake catch rates in particular would require further study as no information was available to see how the hake biomass estimates were derived. However, given the nature of bycatch, it would make sense to see whether a correction could be applied to the catchability from the analysis of the target species, independent of the hake population model. Given the lack of good indices for widow rockfish, it is difficult to reject any index out of hand.

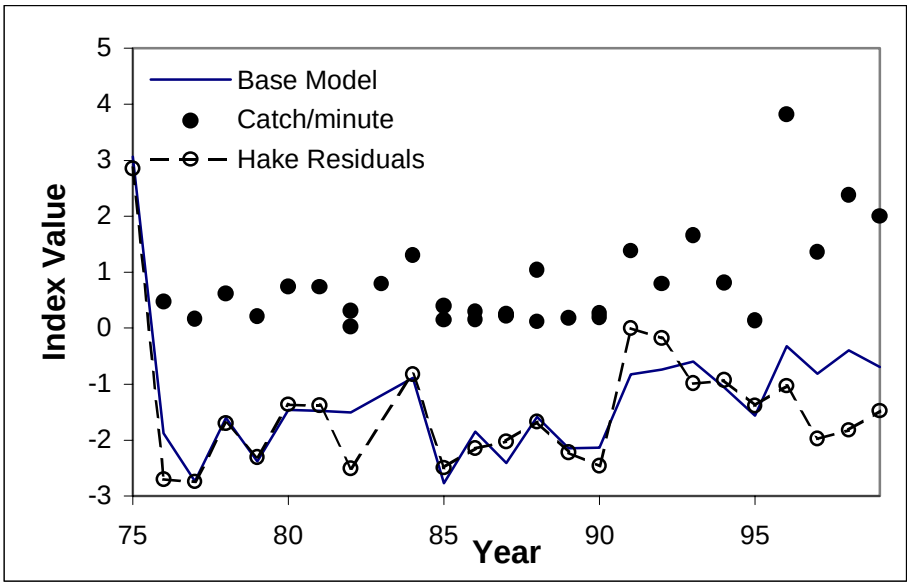

Figure 1 Bycatch of widow rockfish per minute haul for the hake fishery. The Catch/minute is the raw data for the middle latitude band which has the most complete series. The early years 1975-1988 consist of the foreign fleet, the joint venture fleet covers years 1981-1990 and the domestic fleet 1990-2000. The Base Model and Hake Residuals are the year terms for a GLM fitted to the bycatch per minute data using all latitude bands. When the hake CPUE residuals are used in the model, there is a negative trend in the series 1991-1999 as the Hake residuals data (not displayed) has a positive trend over this period.

3 GENERAL COMMENTS

3.1 STAR Panel Review System

The panel review system makes an interesting comparison with working group system used in Europe and other parts of the world.[2] Both possess strengths and weaknesses, and in many ways are very similar. The main aim of both systems is to reduce the chance of errors, increase the quality of assessment and share the responsibility of subjective decisions. The approach differs in that whereas working groups tend to consist of the same people all of whom are involved in the analysis, the review system brings in outside independent experts who were previously not involved in the assessment.

An advantage of working groups is that the members become familiar with the stock and the assessment, and can build effectively on previous work. In contrast, STAR Panel team is not familiar with the fishery or the stock and

[2] This opinion is based on only my experience of this STAR panel meeting and discussions with panel members of other meetings.

therefore a lot of background may be missing preventing the team from identifying some problems with the assessment. The STAR Panel was able to identify simple mistakes, such as coding errors, but unable to resolve more complex issues regarding data and data models.

A disadvantage of working groups is that assessments can become stale with little new input. The assessments tend to become dominated by a few long-running members, new members comments may carry little weight, so alternative approaches to the assessment are not explored. The review approach artificially increases the power of the newcomers to the assessment, forcing the assessment team to react to their comments. This forces the assessment scientists to defend their analysis in a way, which strengthens it and forces it to continually question assumptions. Independent review also ensures transparency in presentation, so the STAT team has to justify their reasoning and include additional studies that test their models. This avoids the assessment settling on procedures based on historical precedent rather than scientific methods.

Perhaps the greatest disadvantage of the review system is its tendency to polarise meeting between a defensive STAT team and attacking panel. If these roles are taken too far, the tendency will be for the STAT team to implicitly cover up problems, rather than bring issues into the open. Openness is the key to success of the review system as it is not possible for the panel to become independently familiar with the stock assessment over such a short time. As it was, the panel was very dependent on issues raised by the STAT team themselves at least at the beginning of the meeting, and it was only when the panel members became more familiar with the assessment towards the end that they were able to identify problems independently. A closely related problem is that there is a tendency for the responsibility of the success or failure of the assessment to lie with the STAT team (essentially one person), even if the problems are beyond their control.

Although the approaches of working groups and review appear very different, in practice they are probably fairly similar. In good working groups, assumptions are tested and new avenues explored. Usually a working group report, at least, is subject to independent review. The STAR panel at this meeting tried to be constructive, the meeting was not heavily polarised and significant improvements in the assessment resulted. In working groups, the onus is usually on one or at most two members to do the modelling, and in practice they take on a similar role to the STAT team. Although the independence of the panel led to questions being asked of some of the data and methods applied, where these were based on historical precedents (e.g. previous unpublished scientific work), neither the STAT team nor panel were able to obtain satisfactory answers due to time and other constraints.

Under some circumstances, notably for complex, contentious stock assessments, it may be useful to try to combine and enhance the strengths of both approaches. This would require the panel to get involved in analyses, rather than just request them. Clearly, this would require more time and higher costs for the review, but would allow the panel to apply their expertise directly rather than rely on the general advice they might give. This would also allow

data to be shared between separate organisations which otherwise may have limited contact. For example, such a meeting might be used to bring two related stock assessments together, such as assessments for a stock shared across the Canadian border.

3.2 Meeting Facilities

The meeting facilities were poor. There was no access to photocopiers, printers or computer facilities, which are usually considered necessary for these sorts of assessments. The meetings were all held in a single room, despite there being two assessments, so there was no extra space for the STAT teams to work on the assignments given them.

3.3 Specific Recommendations

To improve the speed and accuracy of review, it would be valuable, as far as possible, to standardise the format and information provided. To some extent, this would be additional guidelines to those already given for the stock assessment report. Standardised diagnostics, such as residual plots, parameter correlation matrix, retrospective analysis, CPUE projections compared to observed values for the last two years of the time series could all be required as standard output. A demonstration that the model fits simulated data where assumptions are not violated would greatly increase confidence in the model since it would show there were no basic errors, such as in the coding or model structure. Guidelines for the model code for AD modeller, spreadsheets and other software would also make reviewing easier as reviewers could become familiar with layouts, variable names and code structure helping them to understand the models and identify errors. Summaries of the overall model, such as a flow diagram, may also help. The format of the widow rockfish and lingcod was adequate, but there is no guarantee similar format would be followed by other assessments.

Where possible, simple models (e.g. biomass dynamics models, simpler forms of SPA) should be fitted and presented alongside complex final assessment models. The models for widow rockfish and lingcod were complicated with over 100 parameters in all cases. It is very difficult to understand how the fitting process behaves and how much results depend upon the choice of model. Simpler models may indicate whether overall qualitative results are robust to structural errors. They may also be useful in identifying the most parsimonious model; for example a tuned SPA might be used to indicate the kind of selection curve most likely to fit the data.

The number of assessments to review should be strictly limited. In practice, the STAR panel was required to review three assessments. This is reflected in the time that could be spent on each assessment and therefore the depth of review. This was not too big a problem in this case, but if the widow rockfish had been split into two assessments, it would have been difficult to complete a proper review within the schedule.

For shared stocks, it would be useful to invite attendance from institutions from the other countries concerned. In the case of widow rockfish and lingcod,

Canadian fisheries scientists should be invited at least to observe if not participate.

The effect of the ability of stock assessment to monitor the fishery needs to be consider when deciding on management controls. A number of controls, such as trip limits, have compromised abundance indices, making it difficult to carry out assessments and to monitor the effect of the control on the stock.

The management council needs to provide guidance on the structure of the decision table they require for each assessment.

The meeting should always include computer, printer and photocopier facilities.

4 ATTENDANCE AT THE REVIEW MEETING

STAR Panel Members

Dr. David Sampson, Oregon State University, STAR Chair

Dr. Paul Medley, Independent Assessment Scientist, United Kingdom, Reviewer

Dr. Russell Millar, University of Auckland, New Zealand, Reviewer

Dr. Kevin Piner, NMFS Northwest Fisheries Science Center, Reviewer

Dr. Ray Conser, NMFS Southwest Marine Science Center, SSC Representative

Mr. Dave Thomas, California Department of Fish and Game, GMT representative

Mr. Pete Leipzig, Fishermen's Marketing Association, GAP representative

STAT Team Members

Lingcod Team –

Dr. Thomas Jagielo, Washington Department of Fish and Wildlife

Ms. Deborah Wilson-Vandenberg, California Department of Fish and Game

Widow Rockfish Team –

Dr. Erik Williams, NMFS Southwest Fisheries Science Center

Dr. Alec MacCall, NMFS Southwest Fisheries Science Center

5 TERMS OF REFERENCES

The consultant will participate in the Stock Assessment and Review (STAR) Panel of the Pacific Fishery Management Council (PFMC) in Santa Rosa, California from June 5-9, 2000. The STAR panel will review assessments for two species, lingcod and widow rockfish, during this meeting.

The consultant is expected to participate actively in the panel, offering advice and constructive criticism of the assessments. The consultant is also expected to assist in the preparation of two panel reports, documenting the technical quality and completeness of these assessments. Finally, the consultant is expected to provide an additional written report describing the consultant's review activities and an assessment of the review. Areas of importance include on how the procedures of the review activities may be improved, panel composition, roles and operation, duration, quality and comprehensiveness of reviews and panel reports, and other panel-related criticisms that the consultant may offer.

Specific

1) Read and become familiar with the assessment reports, and any anonymous reviews of the previous assessment, provided in advance to the consultant. These documents will be provided directly to the consultant by PFMC.

2) Participate actively in the discussions during the STAR panel meeting.

3) Offer constructive criticisms on the procedural and technical aspects of all the assessments, in accordance with the terms of reference.

4) Under the leadership of the STAR panel chair, assist in the drafting of the STAR panel reports (one per species).

5) No later than July 30, 2000, the consultant will submit a written report of his review activities and assessment of the STAR process. The consultant will send the report to David Die, UM/RSMAS, 4600 Rickenbacker Causeway, Miami, FL 33149 (email: ddie@rsmas.miami.edu).