

**REPORT ON THE WEST COAST GROUND FISH
STOCK ASSESSMENT MODELING WORKSHOP**

OCTOBER 25–29, 2004

SEATTLE, WASHINGTON

Prepared by

**Patrick Cordue
Fisheries Consultant
New Zealand**

for

**University of Miami
Independent System for Peer Review**

12 November 2004

EXECUTIVE SUMMARY

The West Coast Groundfish Stock Assessment Modeling Workshop was held October 25–29, 2004, at the Northwest Fisheries Science Center, in Seattle, Washington. I participated in the workshop as a CIE representative. Background documentation and executable code for the new version of Stock Synthesis (SS2) with an example assessment were supplied before the meeting in sufficient time to allow me to become familiar with SS2.

The workshop was conducted in a constructive and amicable atmosphere, and presentations were generally to a high standard. Most aspects of stock assessment were discussed, with a major focus on SS2 and Bayesian stock assessment. SS2 has most of the features of the age-length version of SS1 but it is implemented in Autodif Model Builder (ADMB) and, therefore, has bootstrapping and Markov chain Monte Carlo (MCMC) capability.

There is an ambitious schedule of stock assessments planned for 2005. The modeling workshop was a well conceived step towards achieving quality assessments. The development of SS2 is an advance. Properly used, it will enhance stock assessment author's prospects of producing scientifically sound assessments which appropriately characterize the level of uncertainty given the available data. A danger for many authors is that they may attempt to do full Bayesian assessments which could be beyond their capabilities given their current experience with SS2 and MCMC. There are relatively few experts available to support the stock assessment authors.

My conclusions with regard to the workshop are:

- The workshop was very successful.
- Excellent presentations and constructive discussion enabled stock assessment authors to increase their knowledge of SS2 and Bayesian assessment methods.
- The discussion and feedback will enable SS2 to be appropriately revised and enhanced.
- There is a necessary move toward a “production line” approach to stock assessment with participants striving to achieve consistency in all aspects of stock assessment (input data preparation, model assumptions, presentation of diagnostics, and description of uncertainty).

My main recommendations are:

- Where possible, point estimates should be presented with credibility intervals from Bayesian posterior distributions.
- As an alternative interval estimates can be constructed using bootstrap distributions.
- Extensive sensitivity runs with regard to the main dimensions of uncertainty should be done.
- The use of MCMC does not preclude the need for sensitivity runs.
- There must be a sequence of runs with explainable effects leading from the previous basecase assessment run to the new basecase assessment run (an “audit trail”).
- Some features of SS2 models do not transfer to the rebuilding software. In the medium to long term there should be an integrated package of assessment and projection/rebuilding software.
- A meta database of groundfish stock assessment data urgently needs to be designed and implemented.

BACKGROUND

The West Coast modeling workshop was held from 25-29 October 2004 at the Northwest Fishery Science Center, in Seattle, Washington. It was the third in a series of workshops in 2004 aimed at improving the quality of 23 groundfish stock assessments planned for 2005. The stated purpose of the workshop was to “discuss and improve the models to be used in the 2005 West Coast groundfish stock assessments.” However, for many stock assessment authors, there was also a significant training component to the workshop, given the major presentations on the new version of Stock Synthesis (SS2) and its capability to perform Bayesian assessments through the use of Markov chain Monte Carlo (MCMC). The workshop also provided an opportunity to discuss and implement common approaches to the extraction and analysis of primary data sources for many groundfish assessments.

REVIEW ACTIVITIES

Meeting Preparation

Prior to the meeting the supplied background material (Appendix 1) was selectively consulted. My main focus was on the documentation of SS2 (Methot & Taylor in prep.), papers on survey data analysis (Helser et al. in press, Hamel & Wallace 2003) and the GAO report (Oleson et al. 2004). The 18 previous groundfish assessments and associated STAR Panel reports were only briefly consulted (as I was already familiar with the approaches taken – primarily SS1 or similar integrated models implemented using Autodif Model Builder (ADMB)).

Meeting Attendance

A narrative of the meeting is given below. The meeting adhered fairly closely to the Agenda (see Appendix 4).

25 October

The meeting was convened at 1pm and was introduced by the Workshop Coordinator Stacey Miller.

The first presentation was by Steve Ralston on the revised “stock assessment terms of reference”. He covered the proposed schedule of seven STAR Panel meetings in 2005, with a possible “mop up” meeting for unresolved issues. He explained that the assessment results would be used in management decisions which would not take effect until 1 January 2007. This was part of the reason for the large number of assessments which were planned for 2005. One significant new clause in the terms of reference attempted to clarify the role of STAR Panels, in that they were to review the STAT Team’s assessment rather than try to impose an assessment of their choice. The analogy of reviewing a manuscript was given. The requirement for a description of the uncertainty associated with the assessment results was discussed. One option was to “bracket” the uncertainty with alternative model runs along the “dimension of greatest uncertainty”. Decision tables and Bayesian assessments were also discussed.

The second session covered trawl survey data inputs. There were presentations by Tom Helser and Ian Stewart.

Tom presented a proposed method to use the AFSC slope trawl survey series (Miller Freeman) together with the NWFSC slope series (multi-vessel). Various GLM models had been used to analyse the combined series (Helser et al. in press) and a method which treated vessel as a random effect was proposed. There was much discussion on the use of “super years” (where multiple years are combined because of

inadequate spatial coverage) and its appropriateness given fish growth and/or migration patterns. Methods of producing age and length frequencies from the surveys were also discussed. The observation error for biomass indices are an output of the GLM analysis, but no method was proposed for estimating error associated with the length or age compositions. This is left to the modeling stage and the specification/tuning of effective sample sizes. I mentioned that, in New Zealand, observation error calculations are routinely done for all data inputs – at the modeling stage, error is taken as a “sum” of observation error and process error. Agreement was reached on the indices that Tom would produce for each of the relevant species being assessed in 2005.

Ian discussed the triennial trawl survey time series. There was need for post-stratification for some species so that length frequencies could be obtained in all strata. The post stratification had little effect on estimated biomass or variance.

26 October

The meeting resumed at 8.30 am with a continuation of the session on model inputs. Stacey summarized the discard assumptions which had been made in previous assessments. Availability of data from the West Coast groundfish Observer Program was then discussed. Agreement was reached on the form of data summaries which would be made available to assessment authors. No estimates of precision are available. Availability of early discard data time series was discussed. “Pikitch” data were soon to be available (pending data releases from participants) but no progress on “EDCP” data – but this was thought to be publicly available anyway. Discussion on how assessment authors should “fill in the gaps” between early data and Observer data was put-off until Friday.

The third session was concerned with SS2, its capabilities and some limited performance testing. Rick Methot gave excellent presentations on various aspects of SS2 at various levels of detail. He also went through the example assessment which had been supplied to participants before the meeting (see Appendix 1). SS2 does not include all of the options available in SS1 which is both good and bad. Some streamlining of options was perhaps overdue, but some previous assessments may have used options in SS1 which are not available in SS2. The change from “continuous F” (Baranov) to instantaneous removals (half natural mortality, whole catch, half natural mortality) should not make much difference – other than some speed increase in execution. The effect of length based selectivities was illustrated with respect to “growth morphs” – strong fishery length selection can have a major impact on population composition which will not be modeled unless growth morphs are used. The simulation testing was just a simple bootstrap on a “friendly” data set. This was very much a first step, but the performance of SS2 was fine.

27 October

Session 3 resumed at 8.30 am with two examples of transitioning from SS1 to SS2. Rick presented a successful transition for canary rockfish: there were some differences in biomass trajectory, but they were explained by different mean weight at age. The second presenter, Han-Lin Lai, had not been successful in his attempt to duplicate with SS2 the previous SS1 assessment of petrale sole. There was a difficulty because a selectivity parameterization used in the SS1 assessment was not available in SS2. This accounted for some differences, but there were some unexplained differences. After discussion it was recommended that SS2 be given the capability to run with specified selection at length so that any SS1 selection pattern could be exactly specified in an SS2 run. This does not enable a check that SS1 and SS2 deliver the same estimates with the same data, but it at least allows the estimates from SS1 to be input into SS2 to make sure that the same biomass trajectories, etc, are achieved.

Alec MacCall gave a presentation on his suggested approach for vermilion rockfish; which appears to be more than one species. He proposed a Bayesian meta analysis to estimate steepness and year class strength effects, which would then be used to drive a stock reduction analysis via a delay difference biomass model.

Xi He gave a presentation on a simulation study very loosely based on the current widow rockfish assessment. The results were extremely surprising showing poor performance for estimators using 30 or 60 years of absolute biomass indices each with a 10% coefficient of variation (c.v.). The meeting was not convinced that the results were correct and rejected the claim that conclusions with regard to the current widow model could be drawn from the study (since no absolute biomass estimates are available).

Ian Stewart gave an excellent presentation on MCMC model diagnostics and illustrated the use of the diagnostics for the example assessment. The simplest approach to deal with parameters which appear to be stopping an MCMC converging properly is to specify them rather than estimate them. However, if this is done for “too many” parameters, then the model associated with an apparently converged MCMC run may be quite different from the model which the assessment author actually wanted to use – and will underestimate uncertainty relative to that model.

The day concluded with a brief presentation on weighting model inputs. Alec MacCall suggested a pragmatic approach to establishing effective sample sizes for length or age composition data which the meeting agreed should be used in the short-term in the absence of anything better. Data weighting is an area of active research and there is no agreement on the best method for the relative weighting of data sets or for establishing the overall level of input data precision (i.e., observation error + process error).

28 October

The morning started with discussion of prior distributions including which parameters should have informative priors and which parameters should be estimated. I described practice in New Zealand with recent moves to estimate M . Although it seems desirable to incorporate uncertainty about M in posterior distributions the estimates of M are often higher than expected. I also recommended that the prior beliefs about derived parameters and “risks” be established by running an MCMC with no data except the catch history. It is important to establish what the starting beliefs are to see if the data actually have any impact.

There was discussion of informed priors for surveys qs . I described the approach in New Zealand of specifying a range for each factor (Cordue 1996) and then sampling from a uniform distribution on each factor to obtain a prior (approximately lognormal) from the product of all factors. An alternative was suggested being an update of a previous meta-analysis. However, it was also suggested that such an approach could, to some extent, be using data which later goes into the assessment.

A discussion of the species groups for Friday’s break-out working groups concluded that the existing groups needed to be altered. A rearrangement was made on the basis of common assessment data sets.

The discussion on selection of phases for key parameters concluded that an SS2 feature should be added to allow emphasis factors to be phase specific. This allows an optimum path to be followed to the minimum; emphasis factors must all be returned to 1 for the final phase. A deadline for changes to the version of SS2 to be used in the 2005 assessments was suggested, or perhaps more than one deadline, given the staggered nature of assessments. It was agreed that a proliferation of SS2 versions must be avoided.

The method of treating discards in SS2 was described by Rick Methot. There was discussion on the two types of discards: those driven by size, and those driven by trip limits. Unless trips are explicitly modeled, it is hard to distinguish. A consistent approach to specifying discard rates and using discard data was recommended. To that end, it was suggested that a “bycatch coordinator” be appointed.

There was brief discussion on the use of “minimum estimates” and how they could be used. An obvious approach is to put an informed prior on the q , but there may be little value with only one point in a time series (and little knowledge about q). Using it outside the assessment to judge the plausibility of the result seems less desirable.

There was a brief discussion on power terms in the relationship between observed and expected values. For CPUE a default value was discussed – 1 seems inappropriate in many cases. For the juvenile midwater

trawl survey, it was suggested that a power term was justified because density dependent mortality most likely occurs after the trawl survey.

In the afternoon, Andre Punt gave presentations and led discussion on quantifying uncertainty and rebuilding projections. If MCMC were done, then some diagnostics would be necessary to convince the STAR Panel that the runs had probably converged. In the absence of MCMC, the suggested fall-back position was sensitivities on the “main dimensions” of uncertainty. There was lots of discussion but no general agreement on the best way to present uncertainty in “decision tables”. Obvious dimensions of uncertainty were highlighted: M , σ_R , and h . It was suggested that F_{MSY} and MSY be added to the SS2 reports so that sensitivities would be more obvious.

The rebuilding software (“Puntizer”) was described. There are some inconsistencies between SS2 and the “Puntizer” (e.g., some selectivities have different parametric forms). There is the ability to redefine B_0 by choosing a range of years to represent mean recruitment. This should be discouraged as all the estimates in SS2 are conditioned on the estimated B_0 . A fully integrated estimation and projection package is obviously desirable, but there has already been a lot of “investment” in the “Puntizer”, and there has been legal acceptance of the results.

The main workshop concluded with a presentation and discussion of the workshop recommendations. This was done fairly quickly, and the recommendations were obscured by the summary of what was discussed, rather than being given as a specific list. The exception to this was the suggested changes to SS2. Rick Methot produced a prioritized list which was slightly modified after discussion.

29 October

There was a morning session only, primarily for stock assessment authors. I attended the first group in the main auditorium discussing cabezon, cowcod, scorpion, gopher, kelp greenling, vermillion, and yelloweye. There was very valuable discussion, especially for new assessment authors, about the availability of and common pitfalls for some data sets. Although there was a recommendation from the Data Workshop to go to PACFIN and RECFIN for commercial and recreational data, it was acknowledged (within this break-out group) that for “marginal species” there was good reason for looking at State databases. In terms of obtaining detailed, but short-term data sets which could be useful (or even invaluable) for “ground truthing”, names of various contact people were supplied to relevant authors (depending on the spatial distribution of their species). There is an obvious need for a meta-database. There is also a need for some standardization of data management procedures. Two worrying examples were mentioned. First, a large number of market samples from south California had just been “discovered”, which will double the amount of available market samples. Second, although there is a good Californian otolith database, many otoliths which are supplied and aged are never returned (otoliths can therefore go missing – e.g., otoliths recently being found in someone’s garage).

There was some discussion about how to describe uncertainty if MCMC runs were not achieved or were only obtained after a model had been substantially altered to achieve apparent convergence. I suggested that bootstrap distributions might be used (this feature is already in SS2, because it is an ADMB option). The question arose whether the bootstrap distributions could be used in the rebuilding analysis – Andre Punt was to be asked. Alternative models to SS2 were raised. It is inadvisable to use simple models which make strong assumptions, but there are perhaps some alternative models where strong assumptions are not made which could be tried for some species.

Conduct of the Meeting

The meeting was held in a constructive and amicable atmosphere. Presentations were generally of a high standard and discussion well focused. The only deficiency in the meeting process was in the formulation of workshop recommendations. There was an attempt in the last main session to do a “workshop wrap-up” ,

including recommendations. Unfortunately, recommendations had not been identified during previous sessions except by the rapporteurs in their notes – so the listed recommendations were those which the rapporteurs took from the discussion, rather than those which the meeting might have agreed to by consensus. Also, recommendations were not explicitly identified during the wrap-up session, so again there may be some interpretation by the authors of the workshop report as to what the workshop actually recommended.

SUMMARY OF FINDINGS

There is an ambitious schedule of stock assessments planned for 2005. The modeling workshop was a well conceived step towards achieving quality assessments. The development of SS2 is an advance. Properly used, it will enhance stock assessment authors prospects of producing scientifically sound assessments which appropriately characterize the level of uncertainty given the available data. A danger for many authors is that they may attempt to do full Bayesian assessments which could be beyond their capabilities, given their current experience with SS2 and MCMC. There are relatively few experts available to support the stock assessment authors. If SS2 proves reliable and the transition from SS1 not too difficult, then there will be moderate demands on Rick Methot's time. If SS2 is less than robust and/or the transitions from SS1 assessments prove difficult, then demands on Rick's time could be extreme.

Strengths and Weaknesses in Proposed Assessment Methods

The proposed schedule of stock assessments in 2005 is very ambitious. In order to conduct these many assessments in one year, there is the need to move into a "production line" stock assessment mode. That is, there must be a "production process" where team members perform specialized activities which move the "product" (stock assessment) towards final completion. There was evidence that this approach was being adopted with regard to input data preparation (e.g., standardized trawl survey indices being produced by a common method for all relevant species being assessed). Also, during the workshop, there was a consistent theme of trying to achieve consistency in the methods used to overcome common problems (e.g., the suggestion to appoint a "bycatch coordinator" to deal consistently with discard issues). Of course, the use of SS2 for most species is also an important part of the "production line". In terms of reducing errors in stock assessment models it is crucial to have one main stock assessment package which has undergone extensive and careful testing.

The unfortunate aspect of the schedule is that SS2 has not undergone extensive testing as it has just been developed. Prior to the assessments, it will also undergo a moderate amount of revision as specific features required for the assessments are implemented. There is also the added burden on stock assessment authors in that they will feel compelled to attempt full Bayesian assessments, despite the fact that the concepts are new to many of them, as is the software and MCMC production methods and diagnostics. I was assured at the workshop that full Bayesian assessments were not "expected" of all authors. However, it is hard to see someone reporting to a STAR Panel without an attempt to use this most fashionable of methods.

One benefit of Bayesian assessments is the easily interpretable measures of uncertainty through "probability intervals" or "credibility intervals" (note, they are *not* "confidence intervals" – that is a frequentist concept). The use of MCMC to obtain posterior distributions and, hence, "credibility intervals", is the best method available to "clearly present the uncertainties in the assessments" (in response to the GAO recommendation, Oleson et al. 2004). However, there are technical difficulties which can prevent an MCMC run from demonstrably achieving converged posterior distributions. There is a danger for some assessments that authors will substantially alter their initial models as they strive to achieve MCMC convergence. There is also a concern that STAR Panels may only get to review the initial models (i.e., the "mode of posterior" (MPD) estimates) because of the time it can take to do MCMC runs. Authors new to MCMC should plan, in the first instance, to provide STAR Panels with bootstrap distributions, and only move to MCMC if time permits.

Stock Synthesis 2

The new version of Stock Synthesis incorporates most of the main features of the length version of SS1. It has the interesting innovation of “growth morphs”, a generalization of sex specific growth. There is also a reporting feature which gives an indication of total annual estimated fishing pressure by presenting the associated spawning biomass per recruit (SSR) as a percentage of virgin SSR (e.g., you can see at a glance whether the total catch in any year – possibly over multiple fisheries – is fishing above or below $F_{40\%}$).

While the SS2 software is perhaps at a “beta level”, the technical documentation (Methot & Taylor in prep.) is very much a rough draft (pre “alpha”). There were several errors in the documentation which made it difficult reading in parts. I have supplied Rick Methot with a list of the errors I detected and some suggestions for revision. The formulation of the SS2 likelihoods contains a minor technical error in that indices with lognormal error structure are “median unbiased” rather than mean unbiased (Appendix 2). Rick has undertaken to change the equations so that the indices are mean unbiased.

The issue of specifying or estimating the standard deviation of log recruitment (σ_R) was discussed. The current formulation requires that σ_R is specified as a parameter in the prior on recruitment deviations. It is up to the assessment author to check that the specified value is consistent with the standard deviation of the *estimated* recruitment deviations. I assume that if they are very different then a re-estimation procedure is manually iterated until the values are consistent. An alternate, more natural approach is to allow the specification of a prior on the recruitment multipliers (year class strengths) which is independent of σ_R and to estimate σ_R as a derived parameter (i.e., as the standard deviation of the year class strengths). This has the advantage of producing a posterior for σ_R and avoids possible iteration. The priors on year class strength could be informative or uninformative as desired (see Appendix 3 for the natural uninformative priors). A criticism of this approach is that it is not strictly technically correct as the year class strengths are better modeled as a random effect (although this cannot currently be done in SS2). However, it is no worse than the many other technical accommodations in SS2 and other modeling software (e.g., composition data often will not follow a multinomial distribution, and trawl survey catchability varies from year to year and technically would be best modeled as a random effect). The decoupling of σ_R from the prior on recruitment multipliers is on Rick’s “to do” list for SS2 modifications.

SS2 can fit to length data and estimate growth parameters for multiple growth morphs. However, for some species, there may be large cohort specific variation in growth (with possible year effects – e.g., good growing years for all cohorts) which would make use of length data in SS2 problematic (i.e., having to estimate growth for many cohorts with time varying blocks as well). In such cases, it might be better to convert the length data to age data outside of SS2. This can be done using a purpose written estimation model which allows the estimation of cohort specific growth and year effects (e.g., Hicks et al. 2002).

MCMC and ADMB

SS2 is written using ADMB, and it utilizes the built-in MCMC routines. Unfortunately, ADMB requires that a positive definite Hessian matrix be obtained at the MPD estimate. This is because the Hessian is used in the random steps when all parameters are updated simultaneously during the MCMC. This is inconvenient as there are many other ways to do MCMC updating, some of which might be preferable in some circumstances (e.g., see Brooks 1998). It is desirable to have more options for MCMC in SS2, but this requires changes to ADMB, which is proprietary software (the author of ADMB has been amenable to user suggestions in the past and this could perhaps be pursued).

CONCLUSIONS AND RECOMMENDATIONS

My conclusions with regard to the workshop are:

- The workshop was very successful.
- Excellent presentations and constructive discussion enabled stock assessment authors to become much more familiar with SS2 and Bayesian assessment methods.
- The discussion and feedback will enable SS2 to be appropriately revised and enhanced.
- There is a necessary move toward a “production line” approach to stock assessment with participants striving to achieve consistency in all aspects of stock assessment (input data preparation, model assumptions, presentation of diagnostics, and description of uncertainty).

I have recommendations in a number of areas given below.

Assessment uncertainty:

- Where possible, point estimates should be presented with credibility intervals from Bayesian posterior distributions.
- An alternative method of constructing interval estimates is to use bootstrap distributions.
- Extensive sensitivity runs with regard to the main dimensions of uncertainty should be done.
- The use of MCMC does not preclude the need for sensitivity runs.
- MCMC sensitivities should be undertaken to explore the effect of alternative priors, data weightings, and specified values.
- The prior assumptions with regard to important derived parameters should be obtained by doing an MCMC run with no data other than the catch history.

Assessment continuity:

- There must be a sequence of runs with explainable effects leading from the previous basecase assessment run to the new basecase assessment run (an “audit trail”).
- In the case of a previous SS1 assessment this requires a comparable SS2 assessment run.
 - At a minimum the SS2 model must be able to reproduce the SS1 results when run with the SS1 estimates.
 - Ideally, the same or similar parameterization in SS2 will enable the same estimates to be obtained as in the SS1 run (with the same data and specifications).
 - Minor differences are of no consequence, but major differences must be explainable.

Software:

- A large number of assessments depend on the availability, reliability, and maintenance of ADMB, SS2, and the “Puntimizer” (rebuilding software).
- Some features of SS2 models do not transfer to the “Puntimizer”. In the medium to long term there needs to be an integrated package of assessment and projection/rebuilding software.
- Consideration needs to be given as to how the current dependence on the continued availability and good health of three individuals can be reduced (e.g., understudies for Rick Methot and Andre Punt; access to ADMB source code).

Data:

- A meta database of groundfish stock assessment data urgently needs to be designed and implemented.

REFERENCES

(see Appendix 1 for further references)

- Brooks, S.P. 1998. Markov chain Monte Carlo method and its application. *The Statistician* 47: 69–100.
- Cordue, P.L. 1996. A model based method for bounding virgin biomass using a catch history, relative abundance indices, and ancillary information. New Zealand Fisheries Assessment Research Document 96/8. 48 p
- Hicks, A.C., Cordue, P.L., and Bull, B. 2002: Estimating proportions at age and sex in the commercial catch of hoki (*Macruronus novaezelandiae*) using length frequency data. *New Zealand Fisheries Assessment Report 2002/43*. 51 p.

APPENDIX 1: MATERIAL PROVIDED

- Barnes, T. et al. 1999. Star Panel Report on the petrale sole assessment. 7 p.
- Brodziak, J. et al. 1998. Star Panel Report on the blackgill rockfish (*Sebastes melanostomus*) assessment. 8 p.
- Butler, J.L., Jacobson, L.D., Barnes, J.T., Moser, H.G., Collins, R. 1999. Stock assessment of cowcod. 117 p.
- Butler, J., Jacobson, L. and Barnes, J.T. *Stock Assessment for Blackgill Rockfish*. In Pacific Fishery Management Council. 1998. Appendix: Status of the Pacific Coast Groundfish Fishery Through 1998 and Recommended Acceptable Biological Catches for 1999: Stock Assessment and Fishery Evaluation. Pacific Fishery Management Council, 2130 SW Fifth Avenue, Suite 224, Portland, Oregon, 97201.
- Carlile, D. et al. 2002. Star Panel Report on the yelloweye rockfish assessment. 4 p.
- Cope, J.M., Piner, K., Minte-Vera, C.V., Punt, A.E. 2003. Status and Future Prospects for the Cabezon (*Scorpaenichthys marmoratus*) as assessed in 2003. 146 p.
- Conser, R. et al. 2003. Star Panel Report on the Pacific Ocean perch assessment. 7 p.
- Conser, R. et al. 2003. Star Panel Report on the widow rockfish assessment. 5 p.
- Butler, J.L., Barnes, T., Crone, P., Conser, R. 2003. Cowcod Rebuilding Review. 9 p.
- Hamel, O.S., Stewart, I.J., Punt, A.E. 2003. Status and Future Prospects for the Pacific Ocean Perch Resource in Waters off Washington and Oregon as Assessed in 2003. 123 p.
- Hamel, O.S. & Wallace, J.R. 2003. Calculation of summary statistics for the Pacific west coast upper continental slope trawl survey of groundfish: resources off Washington, Oregon and California. Fishery Resource Analysis & Monitoring Division Contribution N. FRAM-03/04-2. 20 p.
- Hastie, J. (draft). Estimation of discard for use in 2005 groundfish stock assessments: A presentation of sample results for the October 25-29, 2004 Modeling Workshop. 10 p.
- Helser, T.E. et al. 2003. Star Panel Report on the bocaccio assessment. 11 p.
- Helser, T.E., Punt, A.E., Methot, R.D. (in press). A generalized linear mixed model analysis of a multi-vessel fishery resource survey. Fisheries Research. 14 p.
- He, X., Ralston, S.V., MacCall, A.D., Pearson, D.E., Dick, E.J. 2003. Status of the widow rockfish resource in 2003. 137 p.
- Jagiello, T.H., Wallace, F.R., Cheng, Y.W. 2003. Assessment of Lingcod (*Ophiodon elongatus*) for the Pacific Fishery Management Council in 2003. 151 p.
- Lai, H.L. et al. 2003. Star Panel Report on the cabezon stock assessment. 6 p.
- Lai, H.L. et al. 2003. Star Panel Report on the lingcod stock assessment. 7 p.
- Lai, H.L., Tagart, J.V., Ianelli, J.N., Wallace, F. 2003. Status of the yellowtail rockfish resource in 2003. 101 p.
- MacCall, A.D. 2003: Status of bocaccio off California in 2003. 60 p.
- Methot, R.D. 2000. Technical description of the stock synthesis assessment program. U.S. Dept. Commer., NOAA Tech. Memo. NMFS-NWFSC-43, 46 p.
- Methot, R.D. SS2 executable, sample files and spreadsheets.
- Methot, R.D. (draft). Overview of SS2 for Oct 25-29 Modeling Workshop. 1 p.
- Methot, R.D. (draft). SS2 user manual. 17 p.
- Methot, R.D. et al. 1999. Star Panel Report on the cowcod rockfish assessment. 3 p.
- Methot, R.D. & Piner, K. 2001. Status of the Canary Rockfish Resource off California, Oregon and Washington in 2001. 133 p.
- Methot, R.D. & Taylor, I.G. (in prep.). Technical description of the Stock Synthesis II assessment program. Draft, October 11, 2004. 28 p.
- Methot, R.D., Wallace, F., Piner, K. 2002. Status of yelloweye rockfish off the U.S. West Coast in 2002. 75 p.
- Oleson, K.W. et al. 2004. Pacific groundfish: Continued efforts needed to improve reliability of stock assessments. United States General Accounting Office Report to Congressional requesters. 47 p.
- Rogers, J.B. 2003. Darkblotched Rockfish (*Sebastes crameri*) 2003 stock status and rebuilding update. 55 p.
- Piner, K., Methot, R.D. 2001. Stock Status of shortspine thornyhead off the Pacific West Coast of the United States 2001. 161 p.

- Ralston, S. et al. 2002. Review of the updated 2002 sablefish stock assessment. 7 p.
- Ralston, S. et al. 2003. Star Lite Panel Report on the darkblotched rockfish (*Sebastes crameri*), yellowtail rockfish (*Sebastes flavidus*), and cowcod (*Sebastes levis*) stock assessments. 9 p.
- Sampson, D.B. & Wood, C. 2001. Stock Status of Dover Sole off the U.S. West Coast in 2000. 119 p.
- Sampson, D.B. & Lee, Y.W. 1999. An Assessment of the Stocks of Petrale Sole off Washington, Oregon and Northern California in 1998. 99 p.
- Schirripa, M.J. 2002. Status of the Sablefish Resource off the Continental U.S. Pacific Coast in 2002. 63 p.
- Stauffer, G. et al. 2002. Star Panel Report on the canary rockfish assessment. 10 p.
- Tagart, J. et al. 2001. Start Panel Report on the dover sole assessment. 7 p.
- Tagart, J. et al. 2001. Start Panel Report on the shortspine thornyhead rockfish assessment. 6 p.
- Yin, Y. & Sampson, D.B. 2004. Bias and precision estimates from an age-structure stock assessment program in relation to stock and data characteristics. *North American Journal of Fisheries Management* 24: 865–879.

APPENDIX 2: LIKELIHOOD AND LOGNORMAL ERRORS

The presentation of error structures in the draft SS2 documentation (Methot & Taylor in prep.) is not ideal as the error structures are not explicitly presented in the equations. There is a statement of the expected values in the “Observation Model” section followed by the log-likelihood in the “Statistical Model” section. This does allow the explicit error structure to be deduced, but it would be better to present the formal equations. When this is done, it is actually seen that the likelihood being used for lognormal indices is inconsistent with the usual notion of an index.

To illustrate, consider a biomass index X_i :

$$X_i = qB_i\varepsilon_i$$

where B_i is the biomass (in year i), q is the proportionality constant, and ε_i is the error (in year i). Suppose that the errors are lognormal: $\log(\varepsilon_i) \sim N(\mu_i, \sigma_i^2)$. It then follows that,

$$\log(X_i) \sim N(\log(qB_i) + \mu_i, \sigma_i^2)$$

and the negative log-likelihood (ignoring constants) is

$$\frac{1}{2} \sum_i \left[\log(\sigma_i^2) + \frac{(\log(X_i) - \log(qB_i) - \mu_i)^2}{\sigma_i^2} \right]$$

If the variances are assumed known, then the first term in the square brackets in the above equation can be ignored. The likelihoods presented in Methot & Taylor (in prep.) are consistent with the assumption, in every year, that $\mu_i = 0$. However, under this assumption it follows that:

$$E(X_i) = qB_i E(\varepsilon_i) = qB_i e^{\frac{\sigma_i^2}{2}} = qB_i \sqrt{cv_i^2 + 1}$$

where cv_i is the specified c.v. in year i .

When the c.v.s are relatively small (< 0.35), there is a very small bias in the indices. However, by definition, they are no longer indices in the usual sense. The assumption in the draft document is consistent with “median” unbiased indices, in that there is a 50% probability that an index will be above or below the true value (qB_i). This would be acceptable if the random variables in question could be expected to have this property. However, this would not generally be true and it would be preferable to use “mean” unbiased indices:

$$E(X_i) = qB_i E(\varepsilon_i) = qB_i$$

This requires $\log(\varepsilon_i) \sim N(-0.5\sigma_i^2, \sigma_i^2)$ and for known variance the negative log-likelihood (ignoring constants) is:

$$\frac{1}{2} \sum_i \left[\frac{(\log(X_i) - \log(qB_i) + \frac{\sigma_i^2}{2})^2}{\sigma_i^2} \right]$$

When the likelihood is expressed as a function of q and differentiated one can derive a formula for the q which minimizes the negative log-likelihood for given biomass:

$$\frac{dL(q)}{dq} = \frac{-1}{q} \sum_i \left[\frac{\log(X_i) - \log(qB_i) + \frac{\sigma_i^2}{2}}{\sigma_i^2} \right] = 0$$

Which implies:

$$\log(q) = \frac{\frac{n}{2} + \sum_i \frac{\log\left(\frac{X_i}{B_i}\right)}{\sigma_i^2}}{\sum_i \frac{1}{\sigma_i^2}}$$

which is analogous to the formula given in Methot & Taylor (in prep.) for “median” unbiased lognormal indices but has the extra term $n/2$.

APPENDIX 3: ALTERNATIVE PRIOR FOR YEAR CLASS STRENGTH

The current SS2 formulation has lognormal priors on recruitment deviations about a “bias corrected” stock-recruitment relationship (Methot & Taylor in prep.). The formulation requires that σ_R (the s.d. of the log-deviations) is specified despite there generally being information in the data with regard to recruitment variation. It could be useful to allow alternative priors to be specified for the recruitment deviations, or the year class strengths (being the product of the bias correction and the recruitment deviations) and to estimate σ_R as a derived parameter (i.e., being the s.d. of the estimated year class strengths).

Since YCS is a multiplier, the natural uninformed prior for year class strength (YCS) is a uniform on $\log(\text{YCS})$ with $E(\text{YCS}) = 1$. A method for specifying this type of prior is given below.

Let $Y = \log(X) \sim U(a, b) : E(X) = 1$. The specified expectation requires:

$$E(X) = \frac{e^b - e^a}{b - a} = 1$$

The problem is to find bounds on YCS, e^a, e^b which are sensible and also satisfy the above equation. The bounds should be wide because we are looking for an uninformed prior. There is no analytical solution to the above equation for a given upper (or lower) bound. However, for given b, e^b the following equation quickly converges to a solution (with starting value $a_0 = 0$):

$$a_{n+1} = e^{a_n} - e^b + b$$

A sample table of solutions is given below:

a	b	e^a	e^b
-7.70	2.30	4.54×10^{-4}	10
-4.19	1.79	1.51×10^{-2}	6
-3.36	1.61	3.49×10^{-2}	5

The pdf for X is:

$$f_X(x) = \frac{1}{(b-a)x} \quad \text{for } e^a \leq x \leq e^b$$

If X_1, \dots, X_n (being n YCS) are given identical independent priors as above, then the negative log likelihood (ignoring constants) is:

$$\sum_i \log(X_i)$$

Because of this, MPD estimates will tend to e^a if there is little or no information for an estimated YCS in the data. However, for such cases in MCMC runs the posterior will tend to the prior which sensibly has a mean of 1. If these priors were to be used for MPD estimates then it might be sensible to impose a penalty encouraging the estimated YCS to average to 1.

APPENDIX 4: WEST COAST GROUND FISH STOCK ASSESSMENT MODELING WORKSHOP AGENDA

MONDAY, OCTOBER 25, 2004

NWFSC AUDITORIUM

Session 1. Introduction

- 1:00 p.m. Welcome – Stacey Miller
- Review list of models authors plan to use
- 1:15 p.m. Stock Assessment Terms of Reference – Steve Ralston

Session 2. Model Inputs

Facilitator: Ian Stewart

Rapporteur: Melissa Haltuch

- 2:30 p.m. Survey Data
- Generating Biomass Indices
 - Progress report on GLM Analysis using AFSC and NWFSC Slope Surveys for DTS and slope species. Report on exploring the error models for slope species –Tom Helser
 - Report on exploratory work toward differentiating trawlable and untrawlable areas for survey biomass expansions –Tom Helser
 - Building Age and Length Comps*
 - Discuss use of and/or smoothing length-age transition matrices when lacking ages or have non-representative ages
- 3:15 p.m. Break
- 3:30 p.m. Survey Data Discussion (Continued or move to the next topic)
- 5:00 p.m. Wrap up for the day

TUESDAY, OCTOBER 26, 2004

NWFSC AUDITORIUM

Session 2. Model Inputs Continued

Facilitator: Ian Stewart

Rapporteur: John Wallace

- 8:30 a.m. Reports on Observer Data
- Report on NWFSC WCGOP Discard estimates for 2000-03 - NWFSC
 - Report on availability of length frequency and average weights from observer data - NWFSC
 - Report on availability of historical discard data (Pikitch and EDCP data) to assessment authors – NWFSC
 - Report on compilation of historical discard analyses and assumptions used in most recent stock assessments - NWFSC
- 9:45 a.m. Break

Session 3. Stock Assessment Models

Facilitator: Stacey Miller

Rapporteur: Tom Helser

- 10:00 a.m. Presentation on Stock Synthesis 2 (SS2) and performance testing using simulated data - Rick Methot
- 12:00 p.m. Lunch (Pre-ordered box lunches available)
- 1:00 p.m. Discussion of SS2 and performance testing (Continued)
- 3:15 p.m. Break
- 3:30 p.m. Discussion of SS2 (Continued or move to the next topic)
- 5:00 p.m. Wrap up for the day

WEDNESDAY, OCTOBER 27, 2004**NWFSC AUDITORIUM****Session 3. Stock Assessment Models Continued***Facilitator: Stacey Miller**Rapporteur: Gavin Fay*

- 8:30 a.m. Transitioning from SS1 to SS2 - Examples and discussion
- Canary rockfish - Rick Methot
 - Petrale sole – Han-Lin Lai
- 10:15 a.m. Break
- 10:30 a.m. Models other than SS2 that will be used in 2005 assessments
- WinBUG for vermilion rockfish – Alec MacCall
 - Underestimate recruitment potential in fishing-down situations? A simulation study – Xi He
- 12:00 p.m. Lunch (Pre-ordered box lunches available)

*Facilitator: Alec MacCall**Rapporteur: Jason Cope*

- 1:00 p.m. Model Diagnostics
- MCMC diagnostics - Ian Stewart and Andre Punt
 - Example – Ian Stewart and Rick Methot
- 3:00 p.m. Break

Session 4. Modeling Issues and Considerations*Facilitator: Michael Schirripa**Rapporteur: Owen Hamel*

- 3:15 p.m. Discussion Topics
- How to weight model inputs
 - Tuning “effective sample sizes” and survey error levels – Alec MacCall
 - Methods for weighting CV and additional variance components for area-swept biomass indices
- 5:00 p.m. Wrap up for the day

THURSDAY, OCTOBER 28, 2004**370W **NOTE ROOM CHANGE******Session 4. Modeling Issues and Considerations Continued***Facilitator: Michael Schirripa**Rapporteur: Owen Hamel*

- 8:30 a.m. Discussion Topics
- Selection of prior distributions
 - Selecting phases for estimation of key parameters
 - Handling discard in stock assessment models
- 10:15 a.m. Break
- 10:30 a.m. Discussion Topics
- Recreational CPUE linearity
 - Juvenile surveys and non-linear relationships
 - Inclusion and estimation of spawner-recruitment curve in assmt. models
 - Consistent approach to invoking time-varying fishery selectivity
 - Can estimated minimum count or biomass derived from in situ observational data be included in model as input data?
- 12:00 p.m. Lunch (Pre-ordered box lunches available)

Session 4. Modeling Issues and Considerations Continued*Facilitator: Andre Punt**Rapporteur: Han-Lin Lai*

- 1:00 p.m. Discussion Topics
- Quantifying and reporting uncertainty
 - MCMC, sensitivity analysis, guidelines for decision tables
 - Rebuilding projections and forecasting - Rick Methot and Andre Punt
- 4:00 p.m. Workshop Wrap-Up and Recommendations - Michael Schirripa

FRIDAY, OCTOBER 29, 2004

MULTIPLE ROOMS

Session 5. Break Out Working Groups

8:30 a.m. Break out groups for assessment authors - All assessment authors are strongly encouraged to attend the break-out working groups to discuss data and/or modeling issues that are specific to species groups.

- Petrale sole, English sole, Starry flounder - Room 366 W
- Sablefish, Dover sole, Shortspine thornyhead, Longspine thornyhead, POP, Darkblotched, Blackgill - Room 370 W
- Cowcod, Cabezon, California Scorpionfish, Gopher, Kelp Greenling – Auditorium 1
- Canary, Boccacio, Vermilion, Lingcod, Widow, Yelloweye, Yellowtail – Auditorium 2

12:00 p.m. Workshop Concludes

APPENDIX 5: STATEMENT OF WORK

Consulting Agreement Between The University of Miami and Patrick Cordue

OCTOBER 6, 2004

General

The Pacific Fishery Management Council is in the process of initiating a two-year management cycle for West Coast groundfish. This two-year cycle is comprised of a year in which improvements can be made in assessment data and methodologies followed by a year in which assessments are completed and reviewed. During 2005, assessments are planned for 23 groundfish species. This represents a substantial increase in the number of assessments that historically have been completed and reviewed in a single year. In order to ensure that this demanding schedule yields the best possible science, three workshops are being convened during 2004. The West Coast Groundfish Modeling Workshop is the third workshop in the series of "Off-Year Science Improvement Workshops".

In addition to this being the first year in which over 20 species are assessed, a new version of Stock Synthesis, Stock Synthesis 2 (SS2), will be used by many of the West Coast Groundfish assessment authors. SS2, created by Dr. Richard Methot, is an integrated analysis model that includes age and size-based dynamics as well as diffusion between geographic areas. The model is coded in ADMB and incorporates and expands the feature set of length-Synthesis used for several West Coast groundfish assessments. There is no preset limit on the number of fisheries or surveys, and model parameters can be a function of environmental data and/or have a random walk over time. The population is modeled as a set of phenotypic morphs, each with an assigned gender and unique growth and natural mortality parameters. Numbers-at-age for each morph are tracked independently, so that size-specific fishing mortality will have a differential effect on the survivorship of each morph. Recruitment is apportioned among areas and growth morphs. Different morphs can be assigned to different areas to mimic a cline in size-at-age while also allowing diffusion of each morph along the cline.

The purpose of the modeling workshop is to discuss and improve the models to be used in the 2005 West Coast groundfish stock assessments. Specifically, the workshop will be convened to examine the performance of stock assessment models that will be used in 2005 assessments, as well as discuss authors' progress using these models. Additional topics to be discussed during the workshop include analytical methods for preparing model inputs, calculating and reporting uncertainty in stock assessments, and species-specific modeling issues. The workshop will be open to the public and will be held at the Northwest Fisheries Science Center in Seattle, Washington. Significant attendance is expected by West Coast stock assessment authors, members of the SSC groundfish subcommittee, members of the Pacific Fishery Management Council community, and the public.

The consultant will serve as a participant in this modeling workshop. This independent reviewer will improve the quality and credibility of the assessments by providing additional expertise. The consultant should have hands-on experience in population dynamics and conducting stock assessments of groundfish. Specific experience in the integrated analysis type of modeling approach, using ADMB, age-and size-structured models, use of MCMC to develop confidence intervals, and use of Generalized Linear Models to process survey and logbook data for use in assessment models is desirable.

The consultant will be provided with the following background materials prior to the workshop:

- Most recent assessment reports for 18 groundfish species that have previously been assessed
- Documentation of new version of stock synthesis
- GAO Report

Specific

The consultant's duties shall not exceed a maximum total of 14 days: several days prior to the meeting for document review; the four-day meeting; and several days following the meeting to complete the written report. The report is to be based on the consultant's findings, and no consensus report shall be accepted.

The consultant's tasks consist of the following:

- 1) Become familiar with the new version of stock synthesis (Stock Synthesis 2).
- 2) Participate in the stock assessment modeling workshop in Seattle, Washington during October 25-29, 2004.
- 3) Assist in developing a common approach to writing the stock assessment reports.
- 4) Address the following issues in a written report:
 - Make recommendations on the best methods to address the following GAO report recommendation: "Require that stock assessment reports clearly present the uncertainties in the assessments, such as the margin of error associated with population estimates."
 - Recommend model configurations or formulations as appropriate during workshop and within report.
- 5) No later than November 12, 2004, submit a written report consisting of the findings, analysis, and conclusions (see Annex I for further details), addressed to the "University of Miami Independent System for Peer Review," and sent to Dr. David Die, via e-mail to ddie@rsmas.miami.edu, and to Mr. Manoj Shivlani, via e-mail to mshivlani@rsmas.miami.edu.

Signed _____

Date _____

ANNEX I: REPORT GENERATION AND PROCEDURAL ITEMS

1. The report should be prefaced with an executive summary of findings and/or recommendations.
2. The main body of the report should consist of a background, description of review activities, summary of findings, conclusions/recommendations, and references.
3. The report should also include as separate appendices the bibliography of all materials provided and a copy of the statement of work.