# NATIONAL MARINE FISHERIES SERVICE REPORT ON GROUNDFISH MANAGEMENT

National Marine Fisheries Service (NMFS) Northwest Region will briefly report on recent regulatory developments relevant to groundfish fisheries and issues of interest to the Council. NMFS Northwest Fisheries Science Center will also briefly report on groundfish-related science and research activities.

**Council Task:**

**Discussion.**

Reference Materials:

1. Agenda Item B.3.b, Attachment 1: A Summary Report from the Stock Assessment Modeling Workshop held October 25-29, 2004 at the Northwest Fisheries Science Center, Seattle, Washington.

Agenda Order:

a. Regulatory Activities — Steve Freese
b. Science Center Activities — Elizabeth Clarke
c. Reports and Comments of Advisory Bodies
d. Public Comment
e. Council Discussion

PFMC
03/18/05

**A Summary Report from the Stock Assessment Modeling Workshop held October 25-29, 2004 at the Northwest Fisheries Science Center Seattle, Washington**

**Northwest Fisheries Science Center, FRAM Division**
**March 16, 2005**

This report summarizes the discussion and outcomes from the West Coast Groundfish Modeling Workshop, held October 25-29, 2004 at the Northwest Fisheries Science Center in Seattle, Washington. This workshop was the third of three "Off-Year" Science Improvement Workshops convened during 2004 for the purpose of preparing for the West Coast groundfish stock assessments to be conducted in 2005. The overall goal of the West Coast Modeling Workshop was for authors to announce and discuss the models they will use in the 2005 West Coast groundfish stock assessments. Specifically, the workshop was convened to examine the performance of stock assessment models, such as Stock Synthesis 2 (SS2), and discuss analytical methods for preparing model inputs, calculating and reporting uncertainty in stock assessments, and species-specific modeling issues.

The meeting was held in the auditorium of the Northwest Fisheries Science Center, 2725 Montlake Blvd. E., Seattle, Washington. Authors of west coast groundfish stock assessments, members of the SSC groundfish subcommittee, members of the Council family, and the public attended the workshop. The workshop agenda and list of participants are included in Appendix I and Appendix II, respectively.

***Session I. Introduction***

Stacey Miller opened the workshop by providing a list of the stock assessment models that assessment authors are expecting to use in the 2005 assessments. While many of the assessments will be conducted using SS2, additional models include other age-structured models also written in ADMB, and potentially WinBugs for data-poor assessments.

Steve Ralston followed with a presentation on the groundfish stock assessment and review process for 2005-2006 (PFMC 2005) and the Stock Assessment Terms of Reference. The Terms of Reference serves as the primary document that outlines the stock assessment and review process as well as the responsibilities of the involved parties, and requirements for the stock assessment documents. Steve provided an overview of the required contents for each section within the stock assessment documents. He also outlined the guidelines for the executive summary, reporting uncertainty, and inclusion of decision tables. It was noted that stock assessment (STAT) teams must produce three versions of the assessment document: 1) draft to be reviewed by STAR Panel two weeks prior to STAR Panel meeting, 2) complete draft after STAR panel to be provided to Council family, and 3) a final report for publication in the Stock Assessment and Evaluation (SAFE) report. The Terms of Reference document can be obtained by contacting either Stacey Miller (Stacey.Miller@noaa.gov) or Steve Ralston (Steve.Ralston@noaa.gov).

An additional document defining the Terms of Reference for rebuilding analyses will also be provided by the Scientific and Statistical Committee (SSC) of the Pacific Fishery Management Council (PFMC). It was noted that, for species that are below the overfished threshold, assessment authors are expected to conduct separate rebuilding analyses following the conclusion of the STAR Panel meeting.

*Recommendations and Action Items*

- A cut-off date for inclusion of new data will be STAR Panel dependent.
- Workshop participants suggested that a document should be produced that records the management process and actions taken after the STAR Panel review. The document could be included in the SAFE document published by the Pacific Fishery Management Council. Workshop participants suggested the Groundfish Management Team of the PFMC could produce this document.
- Workshop participants discussed what criterion should be used for determining that a stock is overfished when conducting a Bayesian assessment. Some participants suggested a stock should be considered overfished when the probability of an overfished condition is greater than 50%. However, no agreement was reached.

## *Session II. Model Inputs*

### Survey Data

Tom Helser presented a proposal for developing a slope survey analysis for Dover sole, sablefish, shortspine thornyhead and longspine thornyhead, (DTS species) and slope rockfish species. The proposal focused on applying a generalized linear model (GLM), with the specific objectives of 1) examining the AFSC and NWFSC slope survey data to identify a meaningful post-stratification scheme for use with a GLM, and 2) conducting a GLM-based analysis of slope survey data (accounting for spatial/temporal covariates) and generating biomass indices and variances for SS2.

DTS species have formerly been evaluated using the basic survey stratification design of five INPFC areas and two depth zones, 183-566 m, 567-1280 m. However, slope rockfish, many of which occur only within the 183 – 567 m depth strata and are rare by DTS standards, may need to be evaluated for reasonable post-stratification. In particular, any post-stratification analysis should take into account biological features of the species such as gradients in average body weight with depth and latitude, and sufficient samples to obtain reasonable estimates of density and variances. The presented material illustrated gradients in average body weight as grounds for post-stratification for species such as the less frequently caught darkblotched rockfish.

Tom also presented catch data for sablefish and darkblotched rockfish which suggested that the numerous zero hauls and heavy-tailed positive catch distributions may best be modeled using a generalized linear model analysis using the delta distribution. It was proposed that because of the multi-vessel nature of the NWFSC slope survey, a generalized linear mixed model (GLMM) be applied as outlined by Helser et al. (2004)

and that the model's error distributions be evaluated based on the methods of Dick (2004).

Tom's presentation identified a number of challenges to developing slope survey abundance indices, including:

1) Incorporating the AFSC and NWFSC as separate surveys in SS2 models, or combining them (if possible) using a GLMM, treating vessel as random effect;
2) The need to assume that survey selectivity is the same for species of interest when combining surveys. (Alternatively, one could evaluate selectivity empirically, investigating equal q and selectivity within SS2);
3) The incomplete spatial coverage of AFSC slope survey prior to 1997, and the resulting need to combine survey data from multiple years ("super years").

Complete spatial coverage is available for the AFSC slope survey from 1997-2001, (excluding 1998) and for the NWFSC slope survey from 1998-2004. Between 1990 and 1996 (excluding 1994) FRV Miller Freeman covered partial coastal areas. For the DTS species, "super years" were created which combines 1990-1993 and 1995-1996 as the "1992" and "1996" super years, respectively.

Discussion on the GLM-based analysis centered on the need to include the "super years" since, for many species, contrast in the amount of species biomass would be lost if those years were simply omitted. However, an additional variance component would need to be included for the "super years" to account for the incomplete spatial coverage. Also, prior DTS species' assessments included the "super years" in the slope survey biomass series and, as such, any changes would require model sensitivity to a different set of biomass inputs.

Discussion regarding generation of slope survey length/age compositions focused on the appropriate weighting factor in terms of between tow catches and stratum level expansions. The initial proposal was to sum length frequency weighted by predicted tow catch within each stratum. As such, the weighting would be consistent with the model-based approach for generating biomass index. However, the group discussed the necessity of making the expansion consistent with previous methods by weighting each haul by the observed catch.

*Recommendations and Action Items*

GLM-based survey biomass indices:

- Biomass estimates will be generated separately for the AFSC slope and NWFSC slope surveys, and then for a combined AFSC-NWFSC slope survey. For those surveys that include "super years", such as the AFSC and the combined AFSC-NWFSC survey, density estimates by stratum will be generated for "super years." (note: if "super year" is used, analysts will need to account for an additional variance component).
- The generalized linear mixed model (GLMM) analysis (Helser et al. 2004) using the delta approach will be applied to any survey which uses multiple vessels (not

including the multi-vessel shelf survey), and various error distributions will be evaluated using the approach of Dick (2004).
- Using combined surveys, i.e., the AFSC-NWFSC slope survey, implicitly assumes that survey q from the separate AFSC and NWFSC surveys are equal. However, analysts may explore unequal survey q in the SS2 model if separate slope surveys are used.
- If a different slope survey configuration is used from previous assessments, then a sensitivity analysis should be performed to evaluate its affect on model outcomes.
- The GLM-based approach will be fully documented for distribution to STAR panels.

Generating size and age compositions:
- Length and age comps will be decoupled from the GLM analysis using the observed catch weight as a weighting factor for expansion to the stratum level. However, size/age compositions will be developed that are consistent with the stratification from results of post-stratification.
- Owen Hamel will generate the size/age compositions based on the post-stratification and GLM results from Tom Helser.

**Observer Data**

*Discard estimates and data*

Preliminary discussion of discard data and methodologies for estimating species discard occurred during the West Coast Groundfish Data workshop, held July 2004 in Seattle, WA. Participants from the Data workshop requested that Jim Hastie evaluate stratification alternatives to develop annual estimates of discard by INPFC area and evaluate the availability of average weights from the West Coast Groundfish Observer Program (WCGOP) prior to the modeling workshop in order to facilitate additional discussion on the topic. In addition, it was requested that the NWFSC explore the potential for making historical observer data, Pikitch et al. (1988) and Oregon Department of Fish and Wildlife's (ODFW) Enhanced Data Collection Program (EDCP) data, available to assessment authors for exploratory analysis. Stacey Miller was requested to assemble a compilation of historical discard assumptions used in the most recent assessments. Three presentations were made during this session to report progress and outcomes from the requested analyses and data availability.

*Report on NWFSC WCGOP Discard estimates for 2000-03 - NWFSC*

Owen Hamel provided sample results from Jim Hastie's estimation of discard for selected species using the "simple" methodology agreed on during the data workshop in July, 2004. Following the same stratification used for the observer data, estimated total discard poundage and average weights of measured discarded species for observed tows during 2002 and 2003 were presented. The estimated total discard tonnage for all observed and unobserved tows was also presented. It was noted that the zeros in the bycatch ratio tables were a result of closed areas, species distribution, and sampling vagaries. Average weights of discarded fish were not presented during the workshop but will be produced by Jim Hastie.

*Recommendations and Action Items*

- Assessment authors accepted the current level of stratification, by both depth and area, in the discard analysis. In addition, Jim Hastie will provide weighted coast-wide totals and the number of positive tows and the total number of observed tows going into analysis. Authors are requested to contact him if they desire different stratification than what was presented at the workshop.
- Jim Hastie will summarize the tow-by-tow average individual weight of discarded fish for each stratum. A complete listing of the average weights of discarded fish from every tow can be provided, upon request.

*Report on availability of historical observer data*

Points of contact for the historical observer data sets were contacted and asked if the discard data could be made available to all stock assessment authors. Dr. Ellen Pikitch, point of contact for the Pikitch et al. (1988) discard data is contacting colleagues of the project. Mark Salens (ODFW), point of contact for the EDCP data, will be making the EDCP data available to the public.

*Report on Compilation of Historical Discard Analyses and Assumptions*

Stacey Miller presented a compilation of historical use of discards in stock assessments. A PDF of the powerpoint presentation is available at the following ftp site: ftp://ftp.afsc.noaa.gov/nwfscguest/WCG_StockAssessment/ModelingWorkshop/

## *Session III. Stock Assessment Models*

### Stock Synthesis 2 (SS2) and performance testing using simulated data

The Stock Synthesis 2 (SS2) assessment model has been developed to meet the needs of west coast groundfish assessments. It is a length- and age-structured model that incorporates nearly all the features and flexible setup of Stock Synthesis 1 (SS1), which was developed for the west coast sablefish assessment in 1988 and subsequently used for many west coast and Alaska groundfish assessments. SS2 surpasses SS1 by including the apportionment of the population into growth morphs to provide the capability of length-survivorship, and it can partition the population into discrete, intermixing regions. It is coded in ADMB, which provides faster execution and the integration of powerful variance estimation procedures. The model can estimate growth parameters while taking into account the effects of size-selective sampling and ageing imprecision on the size and age data. SS2 estimates annual recruitment as deviations from a spawner-recruitment curve, thus integrates estimation of the spawner-recruitment function. The model provides a procedure for allowing any parameter to vary over time or be a function of a time series of environmental data. SS2 includes specification of the prior probability distribution for each parameter and is capable of a full Bayesian analysis using the Monte Carlo Markov Chain procedure. SS2 is also capable of modeling several fishing fleets and surveys in a length- and age-structured configuration.

The first stage of performance testing involved analysis of simulated data which had characteristics similar to those used in many west coast groundfish assessments. The data simulator feature of SS2 was used to generate 20 data sets using a parametric bootstrap

feature. When SS2 analyzed each of these 20 data sets, the mean estimate of the underlying population was essentially identical to the population from which the data were generated. In addition, the variability of the estimated population between these 20 model runs was essentially identical to the parametric estimate of variability generated by each model run. Thus, the model has demonstrated the basic capability to estimate the abundance and productivity of a harvested fish stock from data available for some west coast groundfish.

*General Recommendations and Action Items:*

- Authors who used (or inherited) an SS1 assessment will be expected to transition to SS2
- Groundfish assessments in 2005 should move towards a Bayesian approach to quantifying uncertainty by means of probability distribution for quantities of interest to fishery management.
- SS2 is capable of full Bayesian analysis. Although SS2 offers the tools necessary for a formal Bayesian approach, it may not be possible to conduct an analysis that estimates all possible parameters using specific, informative priors and that integrates over the feasible parameter space to provide a formal posterior distribution of the management quantities of interest. Nevertheless, such a comprehensive analysis is the goal of every assessment by SS2 or other methods.
- A user manual and documentation will be available for the 2005 assessments, but the GUI will not be available this cycle

*Model Enhancements*

The following enhancements were discussed at the workshop. It was agreed that all changes are not expected to be implemented in the January 2005 model version.

- Add prior type as element to each parameter set-up line
- Create Beta distribution as a prior type
- Phase-specific emphasis levels
- Extra variance term for indices
- Iterative re-weighting options for both indices and compositions
- Improvements to selectivity options
- Read and hardwire a specified set of selectivity values
- Change discard approach so that tuning is to the landed catch and discard is added to this amount
- Move maturity and weight-length from data file to control file
- Estimate Fmsy and do projections
- Output for rebuilding program
- Mean-median bias considerations for fitting of indices
- Decouple sigmaR from the penalty on recruitment deviations
- More options for initial age comp
- Blocks of years for S-R deviations
- Random perturbations to initial parameter values
- Specified seasonality for survey timing and for spawning
- Likelihood profiling

**Transitioning from SS1 to SS2**

The goal of this session was to compare SS1 and SS2 model results in an effort to: 1) verify that the basic model equations would give concordant results based on the same data sources, 2) identify important features of SS1 not yet included in SS2, and 3) identify new features or changes in SS2 that authors familiar with the older version should be aware of. Rick Methot and Han-Lin Lai presented preliminary comparisons of the results from parallel stock assessments (canary rockfish and petrale sole) conducted using SS1 and SS2; both assessments expect to use SS2 for the 2005 assessment.

*Example 1. Canary rockfish*

The canary rockfish example generally showed close correspondence in fits to observed data, management quantities and overall population trend.

A number of features that had been used in the SS1 model but are not available in SS2 were identified and added to the list of desired additions before the final version is complete. Sex-dependent double logistic selectivity was not yet implemented in SS2. There was debate regarding the estimation of selectivity parameters and this was identified as an area for further updates to SS2.

It was suggested that it might be important for some assessments to allow recruitment under initial conditions to deviate from $R_0$. This was not available in SS2, nor was the estimation of initial age-structure. Rick suggested starting the model farther back in time if equilibrium conditions were not expected in the first year of the model. Increasing flexibility for initial conditions was noted as an area for future development.

In the example using canary rockfish, an emphasis value for the recruitment lambda less than one was used. This was the topic of some debate, including concern regarding stochastic projection into the future and the link between the applied $\lambda$ and assumed $\sigma_R$. Workshop participants agreed to leave this issue for the future development of SS2.

There was debate regarding bias correction to the SR relationship when little constraint was imposed on recruitments. Alternate approaches included allowing shrinkage to the mean of the SR curve to be independent of $\sigma_R$ with a prior on year-class strength or changing the framework for estimation of recruitment deviations to a mixed- or random-effects implementation. Again, participants agreed that this issue required further research before better approaches would be available in SS2.

Changes to input data formats, control files quantities and model calculations resulted in small changes to a number of the results. Other differences were observed where dynamics had changed from purely age-based to length-based or a combination of age- and length-based in SS2. There were a number of small changes that authors should be aware of that were not identified as likely candidates for inclusion into SS2.

Likelihood components associated with length-frequency data differed; this may have been due to a simpler structure implemented for the emphasis coefficients ($\lambda$'s) in SS2,

one lambda for each type of data and fleet. The problem of treating lengths and ages as independent likelihood components when they are often based on the same individual fish has not changed between models. One suggested approach was to make the sum of emphasis weights ($\lambda$) to age and length composition data equal to 1.

Small constants are added to composition data in SS2 that were not present in SS1 and may account for some small difference in results. Mean weight at age is calculated from weight at length in SS2 rather than input directly.

The adjustment in SS1 for growth of individuals in the accumulator age for the population is not implemented in SS2. No changes in selectivity, natural mortality, growth, etc. occur for the individuals once they reach this group.

*Example 2. Petrale sole*

The results were qualitatively similar in the Petrale sole example, but differences were noted in most of the model outputs. Changes in structure from SS1 to SS2 and inability to input exactly the same quantities in the same formats appeared to be the source of much of the change in results. This example illustrated that a combination of relatively small differences could have a large effect on overall results.

In an effort to simulate similar initial conditions, the SS2 model included an extended time series back to 1965, versus 1977 for SS1, but early numbers at age did not match. For testing purposes, it would be useful to be able to input and fix the actual selectivity values from SS1 rather than approximating or estimating the parameters.

An ageing-error matrix (e.g. surface vs. break-and-burn) cannot be directly input to SS2. Only the mean observed age and CV for each true age can be input to SS2. This new approach can still account for biased ageing methods and changes in variability with age.

Due to changes in specification of the SR relationship, the SR parameters are different between SS1 and SS2. Additionally, in SS2 the recruitment is always to age-0 rather than other ages used in SS1.

*Recommendations and Action Items:*

- The newly developed SS2 model was shown to be able to reproduce assessment results very similar to those obtained using SS1 when based on the same data. Based on this outcome, it was decided that the 2005 groundfish assessments would use SS2 as the preferred model, with exceptions for some "data poor" situations and some assessments based on age-structured models implemented using ADMB.
- Potential bugs, additional needs and changes should be reported when assessment authors are using SS2.
- Many areas were identified where increased flexibility could be added.
- There are changes to many technical details which users of SS2 should become familiar with when transitioning from SS1 to SS2.

**Models other than SS2 that will be used in 2005 assessments**

Alec MacCall introduced a preliminary approach that he is considering to use for the upcoming vermilion rockfish assessment using Bayesian stock-reduction-analysis (SRA) implemented in the WinBUGS software. Alec provided background on the vermilion rockfish assessment, noting the existence of new genetic evidence that there may be more than one species currently included in the vermilion data sources. There are landings data, as well as a recreational CPUE series available for this species, but it is 'data-poor' compared to other west coast species, and any work must assume that the vermilion 'complex' behaves as one stock.

He proposed to use a delay difference approach based on the SR relationship. By using a Bayesian approach, use of informative priors can help with estimation problems in SRA. These priors will come from Martin Dorn's meta-analysis of stock-recruit functions for rockfish, modified by He et al.'s (in review) derivation of an evolutionary-based prior reducing the probability of steepness values very near 0.2. He intends to use WinBUGS for this analysis.

There was lively discussion of the pros and cons of informative priors and the SRA approach in general. This led to a debate regarding the issues inherent in assessing some species and about developing a common approach to data-poor species; this topic was unresolved. It was noted that Alec will be presenting his approach to the SSC at an upcoming meeting.

Xi He presented the results of a simulation experiment designed to explore the performance of an ADMB-based stock assessment model used for widow rockfish under various assumptions about recruitment variability and prior probability distributions for the steepness parameter of the stock-recruitment function. Xi's presentation described a simulation study based on the Widow rockfish assessment model used in 2003. Widow recruitment inferred from the model, appears to be highly variable, with low steepness (~0.2) and low recruitments in 1990s. The general approach was to explore why estimates of recruitment from the assessment were low and identify potential bias caused by the modeling approach used. He described the process of data-generation and subsequent fitting under a range of assumptions about the S-R function. He presented a summary of the performance of various estimated quantities including current depletion and S-R parameters.

Xi also introduced recent work developing an informative prior on steepness based on the evolutionary persistence principle; this prior has very little density for $h < 0.25$. This research is currently in review at Fishery Bulletin. There was substantial discussion regarding the support for this prior and its use in upcoming assessments. There was also some uncertainty expressed about the conditions for data generation and performance of the simulation under 'very good' data scenarios.

*Recommendations and Action Items*:

- Stock assessment authors should be aware of the level of uncertainty in estimates of modeled recruitments and stock-recruitment parameters.

- WinBUGS stock reduction model seems a viable option for some data poor situations.

**MCMC diagnostics**

Ian Stewart presented an overview of model diagnostics and techniques applicable to use of SS2 as a Bayesian model. He also presented an applied example of these approaches using the Simple2 files distributed to authors with the most recent version of SS2.

Ian described the implementation of Markov-Chain Monte-Carlo (MCMC) and details specific to AD Model Builder (ADMB) that are unlike some other platforms for Bayesian analysis. He proposed that many potential convergence problems could be identified before beginning MCMC runs. Key points included examination of the correlation file for very high or low values which can reveal poor estimation of the Hessian matrix and selectivity parameters are often problem parameters.

The presentation was divided into two types of convergence diagnostics: qualitative and quantitative. It was noted that all diagnostics can point to problems with convergence, but not prove convergence. A main point was that the use of several different diagnostic tests is preferred over any single approach to assess MCMC convergence.

Ian described qualitative analysis of the trace plot (iteration vs. parameter value), examining for both autocorrelation and trend. Running means and cumulative percentiles can be useful as well as density plots to identify substantial departures from multivariate normal. Autocorrelation at lag- plots as well as cross-correlation plots were used to locate high correlations among parameters that could be causing convergence problems. The following quantitative diagnostic statistics were described: Geweke, Time-series methods including effective sample size and naïve vs. corrected chain standard deviations, Heidelberger and Welch, Raftery and Lewis and single-chain Gelman. Ian raised the question of how many runs are enough and there was some discussion of the pros and cons of multiple short chains vs. longer runs. When using multiple chains, the Gelman and Rubin statistic can be added to the list of criteria.

Presentation of MCMC convergence diagnostics was divided into 'key' model parameters and derived quantities of interest, usually a small number, and grouped parameters such as time-series deviations, biomass, and recruitments. Multi-panel displays of key parameters were used to carefully explore convergence, while histograms of convergence criteria were introduced to quickly summarize and present large numbers of values simultaneously. Key parameter plots included trace, density, running mean and percentiles, and autocorrelation plots as well as some suggested summary statistics including median, 5 and 95% quantiles, AC lag 1, Effective N, Geweke and Heidelberger, and Welch statistics. Plotting time series of full posterior densities is an effective method for visualizing the uncertainty around biomass or recruitments over time.

The second presentation used the SIMPLE2 data provided in the October release of the SS2 software. It was noted that the behavior of parameters will be different in each application, and the example was for general illustration not identification of specific difficulties with SS2. The approaches introduced in the first talk were applied to this example, with multiple MCMC chains performed iteratively until convergence appeared to have been reached. Results were summarized using the graphical tools introduced earlier.

*Recommendations and Action Items:*

- Bayesian analysis is not required by the SSC but STATs should conduct Bayesian analyses, including sensitivity to parameter priors, to the extent practicable
- Model runs with zero emphasis on parameter priors are encouraged, to distinguish information in the data from information provided by the priors. Model runs using only parameter priors and landings may also be informative in this regard.
- Detailed reports on convergence of key parameters are recommended; summary of convergence for other parameters and derived quantities is an efficient way to present results.
- Examination of correlations, residuals, and proximity to parameter boundaries needs to be investigated with or without MCMC analysis.
- STAR panels should recognize the time requirements of Bayesian analysis and realize that requests for additional MCMC runs during the STAR Panel meeting may not be feasible

### *Session IV. Modeling Issues and Considerations*

This session of the workshop consisted of round table discussion of issues to consider during modeling exercises. A number of topics were discussed, some of which were touched upon in earlier workshop discussion. This is an attempt to summarize the discussion and outcomes of the various discussion topics.

**Tuning effective sample sizes and survey error levels, methods for calculating and weighting CV and additional variance components for area-swept biomass indices**

Alec MacCall addressed issues regarding the specification of effective sample sizes. Actual sample size or estimated error variability often misrepresents the precision of composition or index data. As a consequence, the use of actual sample sizes can lead to other sources of information being largely ignored by a model. However, truncation is an ad-hoc method of compensation, which risks over-emphasizing small samples and under-emphasizing large ones. Directly calculated error variances in abundance indices are often too small, and generally are replaced with mean-squared-errors of the index to the fitted values. A similar adjustment of age and length compositions uses effective sample size, which is the implied sample size if the error variability about the fitted values were due only to sampling variability. In early uses of the effective sample size approach, actual sample sizes were replaced with a constant representing the overall average effective sample size. However, there is often a relationship between actual and effective sample size that allows for replacing the actual value with a predicted effective sample size. This can be achieved via a linear or non-linear regression of effective on actual sample size. An iterative procedure is needed to estimate effective sample size and

mean-squared-error because of interactions; one recalculation is usually sufficient unless there is strong disagreement among data sources within the model.

**Selection of prior distributions**

The development of prior distributions for parameters is undoubtedly the most controversial aspect of any Baysian analysis. It was therefore recommended that considerable care should be taken to document, to the fullest extent practical, the basis for the various prior distributions. As a programming rule, every estimated variable in SS2 must have associated with it an initial estimate, a prior, a range of possible values, and an associated distribution type. The degree of informativeness associated with any prior (*X*) can range from maximally informative (i.e. fixed, where *X* = initial, a meaningless range and prior, and SD = 0) to maximally uninformative (*X* is within a large, yet definable range with a uniform distribution by a SD = + infinity). Varying degrees of informativeness are then defined by adjusting either the range and/or the standard deviation accordingly. In the present version of SS2, all prior distributions are modeled as normal, Gaussian distributions around the user defined mean and standard deviation. In this way, the user can assume a uniform distribution by choosing an adequately large standard deviation, even though technically the choice of distribution is normal. Discussions were generated around incorporating distribution options other than normal.

*Recommendations and Action Items:*

- "Expert opinion" is the most widely used and accepted means to arrive at a value for a prior.
- Informative priors should be used on variables whose values can be estimated outside the model with a reasonable degree of certainty (e.g. growth parameters)
- Some variables that were deemed appropriate candidates for an informative prior were stock-recruitment steepness, recruitment deviations, survey catchability (Q), and natural mortality (M).
- Values from reliable and pertinent meta-analyses are also good candidates for informative priors
- Informative priors should not be used on variables based solely on the purpose of eliminating erratic behavior of that variable

**Selecting phases for estimation of key parameters**

In general, five phases should be appropriate for modeling most stock assessment situations. Variables that are critical to setting the overall population size (e.g. Ro, Q, etc.) should be estimated in Phase 1 of the fitting procedure. Time varying variables, such as those used in selectivity should be estimated in the final phase.

**Inclusion and estimation of spawner-recruitment curve in assessment model**

Stock assessment authors will provide details of what they did in terms of estimating recruitment and SR relationship.

*Recommendations and Action Items:*

- Stock assessment authors should be aware of the level of uncertainty in estimates of S/R relationships.
- One alternative is to provide alternatives based on different values / priors for S/R
- There needs to be some constraint on recruitment deviations
- The specified sigma-R should be checked for consistency with estimated recruitment deviations
- Options in SS2 for decoupling sigma-R from the penalty on recruitment deviations could be explored

**Consistent approach to invoking time-varying fishery selectivity**

Inconsistent use of time blocks to describe time-varying fishery selectivity can lead to divergent assumptions among assessments. This can be especially troublesome for assessments being evaluated within the same STAR Panel. In an effort to avoid possible complications during the review process, it is in the best interest of assessment authors to maintain a consistent set of assumptions used to describe changes in fishery selectivity based on the behavior of the fleet. Authors should maintain a consistent approach to invoking time-varying selectivity, especially within STAR Panels and among species that have been fished historically similar.

**Recreational CPUE linearity**

Recreational CPUE data are often transformed before being included in stock assessments. The square root transformation is the default, though other transformations should be used if indicated. Non-target species will often have different exponents than target species. In all cases, sensitivity to non-linear transformation vs. no transformation should be tested. Knowledge and understanding of regulation changes, gear changes and natural fluctuations can all be critical in interpreting recreational CPUE data. Any of these may be cause for truncating or splitting a time series to deal with inconsistencies. There can also be significant differences among areas, so if location of fishing changes over time, then either some sort of standardization by area, or splitting of time series may be necessary. Species composition data can help reveal targeting, which may affect how data are treated.

**Juvenile surveys and non-linear relationships**

Steve Ralston presented results from the mid-water juvenile trawl survey where a power function was used to transform the juvenile survey data CV to match the CV at the time of recruitment to the fishery. There was some concern expressed about the paucity of data with which to estimate this relationship, the possibility of inherently more observation error in a juvenile survey, and fitting the exponent parameter outside the assessment model. The group concluded that it is reasonable to transform the juvenile survey index because of compensatory mortality between the time the survey is conducted and when fish recruit to the fishery. However, that transformation should be done within the assessment model, not externally, and it should not be based upon matching variance or CVs. Workshops participants recommended that assessment authors should be explicit about the transformation method used.

**Handling discards in stock assessment models**
Discard rates and compositions will be estimated using observer program data for recent years. However, previous discard rates and compositions may be more difficult to estimate. Most often, discard arises for reasons related to fish size or trip-limit attainment Changes in market forces, gear usage, fishery regulations, and the size composition of a stock can result in significant inter-annual variability in the total amount and size composition of discards. Workshop participants concluded that there should be consistent approaches within species groups and STAR panels. Methods for melding the different historical discard data should be considered within these groups. It was also suggested that future research should attempt to identify the effects that management changes have had on discard over time.

**Use of minimum count or biomass from in situ observational data**
No place currently exists to input minimum count or biomass data into SS2. It was generally agreed upon that authors should consider such data outside the model as a means of testing whether model outputs are reasonable. These data, where available, may be useful for expanding up to a particular zone (habitat/depth/area), but are less likely to be useful on a coastwide, or even area wide, basis.

**Quantifying uncertainty**
Assessment authors using SS2 for a species previously assessed using SS1 should report and compare results from SS2 runs using data used in the previous assessment. Any changes in the model from the previous assessment should be reported in the assessment document. Sensitivity analysis should include variation in M, sigma-r and steepness (h). Comparison of the .cor file from an ADMB run and posterior correlations may be useful in some cases.

**Decision Tables**
It was recommended that assessment authors arrive at the STAR panel with thoughts on factors to be included in the decision tables. Decision tables should look at reasonable variation in states of nature above and below the preferred model settings. Thus when constructing a decision table, the model results should be profiled over at least 2 parameters, including those considered "dominant". Decision tables should include projected yields (ABC and OY), spawning biomasses, and stock depletion levels for each year, as well as MSY, Fmsy and the exploitation rate at MSY. Only reasonably likely states of nature rather than catastrophic events should be considered.

**Rebuilding projections and forecasting**
Rebuilding analysis is required to be conducted on all species found to be in an overfished state. The current rebuilding program is advantageous as it provides the outputs required by the Council, but other methods are not ruled out.

## REFERENCES

Dick, E.J. 2004. Beyond 'lognornmal versus gamma': discrimination among error distributions for generalized linear models. Fisheries Research. Vol. 70 (2-3): 351-366.

Helser T.E., A.E. Punt, and R.D. Methot. 2004. A generalized linear mixed model analysis of a multi-vessel fishery resource survey. Fisheries Research. Vol. 70 (2-3): 251-264

Pikitch, E.K., D.L. Erickson, and J.R. Wallace. 1988. An evaluation of the effectiveness of trip limits as a management tool. NWAFC Processed Report 88-27.

Scientific and Statistical Committee (SSC). 2005. Groundfish stock assessment and review process for 2005-2006. Pacific Fishery Management Council.

## APPENDIX I. MODELING WORKSHOP AGENDA

**West Coast Groundfish 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 – Tom Helser
- 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)

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 compilation of historical discard analyses and assumptions used in most recent stock assessments – Stacey Miller
- Report on NWFSC WCGOP Discard estimates for 2000-03 and the availability of length frequency and average weights from WCGOP data - Owen Hamel
- Report on availability of historical discard data (Pikitch and EDCP data) to assessment authors – John Wallace and Michael Schirripa

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)

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
- Approaches to weighting 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)

**Thursday, October 28, 2004** **370W **Note Room Change****

***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 II. WORKSHOP PARTICIPANTS

John Brandon (UW)
Corrina Chase (UW)
Jason Cope (UW/NWFSC)
Ray Conser (SWFSC, SSC)
Patrick Cordue (CIE)
Shannon Davis (The Research Group)
E.J. Dick (SWFSC)
Martin Dorn (AFSC, SSC)
Gavin Fay (UW)
John Field (SWFSC)
Owen Hamel (NWFSC)
Melissa Haltuch (UW/NWFSC)
Xi He (SWFSC)
Tom Helser (NWFSC)
Annette Hoffmann (WDFW)
Arlene Hruby (NWFSC)
Tom Jagielo (WDFW, SSC)
Meisha Key (CDFG)
Doug Kinzey (UW)
Neil Klaer (CSIRO)
Han-Lin Lai (NWFSC)
Alec MacCall (SWFSC)
Mark Maunder (IATTC)
Rick Methot (NWFSC)
Scott Meyer (ADFG)
Stacey Miller (NWFSC)
Kevin Piner (SWFSC)
Andre Punt (UW, SSC)
Steve Ralston (SWFSC, SSC)
Jean Rogers (NWFSC)
David Sampson (OSU, SSC)
Michael Schirripa (NWFSC)
Tony Smith (CSIRO)
Ian Stewart (NWFSC)
Ian Taylor (UW)
Theresa Tsou (WDFW)
Farron Wallace (WDFW)
John Wallace (WDFW)