Selectivity: theory, estimation, and application in fishery stock assessment models

Workshop Series Report 1

June 2013

Edited by Paul Crone[1], Mark Maunder[2], Juan Valero[3], Jenny McDaniel[1], and Brice Semmens[4]

[1]Southwest Fisheries Science Center (SWFSC)
[2] Inter-American Tropical Tuna Commission (IATTC)
[3]Center for the Advancement of Population Assessment Methodology (CAPAM)
[4] Scripps Institution of Oceanography (SIO)

Center for the Advancement of Population Assessment Methodology (CAPAM)

NOAA / IATTC / SIO
8901 La Jolla Shores Dr.
La Jolla, CA 92037, USA

CONTENTS

ACRONYMS

A-SCALA – Age-structured statistical catch-at-length analysis
ADMB – AD Model Builder
AIC – Akaike information criterion
AKFSC – Alaska Fisheries Science Center
ANOVA – Analysis of variance
B_{MSY} – Stock biomass at MSY
B.C. – British Columbia
BIC – Bayesian information criterion
˚C – Degrees Celsius
CAPAM – Center for the Advancement of Population Assessment Methodology
CASAL – C++ algorithmic stock assessment laboratory
CDFW – California Department of Fish and Wildlife
CPUE – Catch per unit effort
CV – Coefficient of variation
DIC – Deviance information criterion
F – Fishing mortality
F_{MSY} – Fishing mortality at MSY
GP – Gaussian process
HQ – Headquarters
IATTC – Inter-American Tropical Tuna Commission
IPHC – International Pacific Halibut Commission
JIMAR – Joint Institute for Marine and Atmospheric Research
M – Natural mortality
MCMC – Markov chain Monte Carlo
MFCL – MULTIFAN-CL
MSE – Management strategy evaluation
MSY – Maximum sustainable yield
NEFSC – Northeast Fisheries Science Center
NMFS – National Marine Fisheries Service
NMFS-AMWG – National Marine Fisheries Service Assessment Methods Working Group
NOAA – National Oceanic and Atmospheric Administration
NRIFSF – National Research Institute of Far Seas Fisheries
NWFSC – Northwest Fisheries Science Center
OBJ – Floating object
OM – Operating model
OSU – Oregon State University
PBF – Pacific bluefin tuna
PIFSC – Pacific Islands Fisheries Science Center
POP – Pacific ocean perch
R_0 – Virgin (unfished) recruitment
RSMAS – Rosenstiel School of Marine and Atmospheric Science
SEFSC – Southeast Fisheries Science Center
SIO – Scripps Institution of Oceanography
SPR – Spawning potential ratio

SS – Stock Synthesis
SS3 – Stock Synthesis 3
SWFSC – Southwest Fisheries Science Center
UCSD – University of California San Diego
UEM – Universidade Estadual de Maringá (Brazil)
UW – University of Washington
VPA – Virtual population analysis
YFT – Yellowfin tuna

Preface

The CAPAM is a collaborative undertaking, jointly supported by the Southwest Fisheries Science Center (SWFSC, NOAA Fisheries), the Inter-American Tropical Tuna Commission (IATTC), and Scripps Institution of Oceanography (SIO, University of California, San Diego). This Report is the first in a *Workshop Series* published by CAPAM, with CAPAM staff serving as Editors. The CAPAM advisory panel, keynote speakers, and various workshop participants provided useful reviews for improving the Report. Workshop presentations and recordings are available online from CAPAM at www.CAPAMresearch.org. Formal papers produced from proceedings of the workshop and other contributions will be included in a special issue publication of the scientific journal *Fisheries Research* (Maunder et al. *In preparation*). This Report summarizes presentations and discussions made during the workshop. As such, it represents the general views expressed, rather than any achieved consensus set of recommendations. A number of important research topics on selectivity are identified to guide further research, along with recommended practices to consider when developing stock assessment models.

Background

The Center for the Advancement of Population Assessment Methodology (CAPAM, www.CAPAMresearch.org) hosted a workshop on *Selectivity: theory, estimation, and application in fishery stock assessment models* from March 11-14, 2013 at the Southwest Fisheries Science Center (SWFSC) in La Jolla, CA, USA. The four-day meeting inaugurated what will be a longer-term *Workshop Series* that focuses on the broader goal of developing guidance for *Good Practices in Stock Assessment Modeling*. This workshop was sponsored by the National Marine Fisheries Service Assessment Methods Working Group (NMFS-AMWG) and chaired by Mark Maunder (IATTC). A diverse body participated in the workshop, including 65 scientists from federal, state, and international fishery institutions, 21 researchers that contributed recent analysis and case studies pertaining to selectivity, and four keynote presenters. Keynote speakers provided reviews for major sub-topics under selectivity: underlying processes (David Sampson, Oregon State University); specification and estimation (James Ianelli, Alaska Fisheries Science Center); model selection and evaluation (André Punt, University of Washington); and impacts on management (Doug Butterworth, University of Cape Town).

The workshop was structured in a manner that allowed both novice practitioners and experienced analysts to gain insight into selectivity properties and parameterizations involved in developing robust stock assessment models. Each sub-topic comprised a review and several research presentations, followed by group discussion that addressed focus questions and outlined priorities for future research. Additionally, two sessions provided hands-on training for modeling selectivity in various settings. The first was based on the widely-used stock assessment framework and software Stock Synthesis (SS, Methot and Wetzel 2013), with Ian Taylor (Northwest Fisheries Science Center, NWFSC) presenting a tutorial on how to implement the various selectivity options available in SS, and Hui-Hua Lee (Joint Institute for Marine and Atmospheric Research, JIMAR) and Juan Valero (CAPAM) providing accompanying simulation methods and software available for SS to test model assumptions and evaluate model misspecification. In the second training session, Steve Martell (International Pacific Halibut Commission), Mathew Supernaw (Southeast Fisheries Science Center, SEFSC), and Athol Whitten (University of Washington) presented methods for developing software libraries using the open-source software platform AD Model Builder (Fournier et al. 2012), focusing on selectivity examples.

Summary

Major findings and areas of future research generated from group discussions during the workshop are presented below. See Appendix C for focus questions that were used during each session and generally serve as the basis for the following summaries.

<u>Contact selectivity and availability</u>

Fishery-specific selectivity as implemented in contemporary, fully-integrated stock assessments is intended to represent the combined factors that affect fish vulnerability. This includes both contact selectivity (the probability a fish is captured when it encounters the fishing gear) and availability (the probability that a fish is in the area where and when the fishery occurs). Although underlying processes and gear experiments provide insight on the expected shape of selectivity curves, the combination of spatial processes, both in terms of the biology of the fish (e.g., migration, spatial structure, and habitat occupied) and fishing intensity, can alter size- or age-specific selectivity forms. For example, asymptotic selectivity may be assumed based simply on the characteristics of the gear employed, but dome-shape selectivity may occur due to differences in spatial and temporal availability of the fish. Cases characterized by spatially-variable selectivity confound standard approaches of interpreting fishing mortality rates in concert with selectivity. For example, the assumption that all fully-selected fish can be caught with ‘infinite’ fishing intensity is problematic in cases where the fleet is not homogeneous spatially and/or temporally, but rather operates in specific locales and/or at particular times of the year, or not at all in certain areas (e.g., marine reserves).

<u>General selectivity specification and estimation</u>

In assessment models, selectivity is the phenomenon that relates the population’s size and age composition to the size and age composition of the fish observed by a fishery or survey. Selectivity is influenced by fishing gear characteristics, fish behavior, and spatial heterogeneity in the distribution of different sizes/ages of fish and the spatial distribution of the sampling. The potential complexity of these factors means a particular form of selectivity is difficult to define and estimate reliably. In general, the group recommended that the selectivity parameters be estimated (not assumed) within fully-integrated stock assessment models (e.g., Stock Synthesis, CASAL, and MULTIFAN-CL). Assuming fixed selectivity will affect model fits and potentially, compromise estimates of key stock parameters (e.g., growth, natural mortality, and recruitment). Consequently, estimating selectivity accounts for the uncertainty of this process and provides the ability to evaluate interactions among the different data sources and parameters of interest. However, fixing selectivity or mirroring selectivity from one fishery to another can be useful for diagnostic purposes and during model development. Fixing selectivity (or mirroring) can also be helpful in cases where the fishery data (age or length compositions) are limited spatially or temporally.

Selectivity misspecification can impact estimates of management quantities (e.g., MSY, biomass, and depletion). However, few studies have tested for the consequences of model-specification errors for commonly used parametric, non-parametric, and semi-parametric approaches. Some stock assessment data indicate the need for flexible selectivity patterns that differ substantially from what can be estimated using common functional forms (e.g., logistic or 'double-normal' selectivity patterns). Studies also indicate that patterns are likely to change over time to a greater extent than has typically been assumed in stock assessments. The group noted that research on methods for specifying selectivity patterns (e.g., objective ways to set penalty weights for non-parametric approaches) and evaluating the consequences of model misspecification should be a priority for future work.

Asymptotic or dome-shape selectivity

It is common practice in stock assessments to assume asymptotic selectivity for at least one fishery or survey to stabilize parameter estimation. If dome-shape selectivity is estimated for all gears, a 'cryptic' biomass phenomenon may arise, which may translate to population estimates of older fish that are not proportional to those observed through sampling efforts. Assuming dome-shape selectivity for all fisheries and surveys is inherently confounded with assumptions surrounding natural mortality and typically will increase the uncertainty in abundance estimates. For these cases, if one assumes that selectivity for one fleet (fishery or survey) is asymptotic, then estimates will likely be more precautionary (but generally producing poorer fits to the data). Ideally, information should be available to provide an objective stance for specifying at least one gear type having asymptotic selectivity; otherwise, such action would be subjective and necessarily affect accurate estimation of uncertainty. Such variance estimates are needed to apply precautionary buffers (e.g., US National Standard 1 guidelines) based on stock assessment uncertainty. Importantly, these buffers are intended to be applied to assessment models that are 'risk neutral' (i.e., that avoid precautionary assumptions, such as asymptotic selectivity). Recent research indicates that some degree of dome-shape selectivity is to be expected in many situations, due largely to incomplete mixing of individuals and spatial heterogeneity in fishing intensity. Further simulation studies could be used to determine the management implications and overall model performance of assuming (correctly or incorrectly) at least one fishery has asymptotic selectivity.

Size- or age-based selectivity

Choosing whether to model selectivity as a function of size or age should depend on the population dynamics, fishery characteristics, availability (and type) of composition data, and performance variables of interest. As with all data included in a stock assessment, age and/or size compositions should be scrutinized beforehand. For example, analysts should evaluate the error (precision and bias) associated with age-determination methods and examine the consistency between the length and age compositions. Comparisons of size-at-age between fleets can help determine if selectivity is most appropriately modeled as size-based, rather than age-based. This choice can affect estimates of management quantities, particularly due to biases in observations of size-at-age when size-based selection occurs. Simulation analysis is underway to evaluate the effects of assumed and potentially, misspecified selectivity on derived quantities used for management; a high priority research area noted the group noted as in need of that this area of research should be prioritized in future work.

Most contemporary, age-structured stock assessment models are based on the assumption that the size-at-age distribution reverts back to a normal distribution at the start of each time-step and does not change over time. Such an assumption may produce biased estimates of management quantities, particularly in cases with strong size-selective mortality, and may warrant a fully length-structured modeling effort or the use of growth ‘platoons’ or ‘morphs’ as implemented in *Stock Synthesis*. Finally, future research is needed to better understand critical interactions between selectivity type (e.g., age- and/or length-based), available data (e.g., age and/or length), and relevant population processes (e.g., growth, natural mortality, and recruitment) modeled in stock assessments.

Fleets as a proxies for spatial processes

Assessment scientists commonly specify fishing fleets with different selectivities as a proxy for different spatial distributions of the fish and fishery. In these cases, the catches, indices of abundance (CPUE), and biological-composition data are partitioned by area. The appeal of this method is that a simpler model can be used to capture some of the complexities of a spatially-structured situation. However, some studies indicate that this approach may result in biased estimates of management quantities. General guidelines are needed for identifying cases where using fleets as proxies for spatial processes may be appropriate and applications where they should be avoided.

Time-invariant or time-varying selectivity

Selectivity is a function of fishing and biological processes. Consequently, at the population level, it is unlikely to be homogeneous over space and time. This suggests that time-varying selectivity should be assumed for most fisheries. However, estimating selectivity changes over time can be difficult and will rely on good data and a clear understanding of the fishery characteristics. For example, in some cases, the fishery may target abundant cohorts over time (e.g., species that aggregate by size, such as sardine, hake, tuna, and pollock). Ignoring temporal changes in selectivity can produce biased estimates of management quantities and underestimate uncertainty. Introducing time-varying selectivity in cases where it is negligible will effectively down-weight the catch-composition data and likely result in reduced precision. The group discussed whether time-varying selectivity for fisheries should be a default assumption and noted doing so would likely minimize bias, but increase variance, particularly if other biological-composition data (e.g., from surveys) are assumed to have constant selectivity. The principle of parsimony was emphasized by the group, with multiple diagnostic examinations (based on residuals, profiles, retrospective patterns) recommended before increasing model complexity.

Traditional VPA-like approaches are based on aggregating catch from all fisheries and assuming that the estimates of catch-at-age accurately reflect the removals from the population. Selectivity becomes a derived quantity from the calculated fishing mortality rates-at-age within each year, and can vary substantially across ages and years. This general method requires complete and reliable age-composition data for all time periods, fleets, and surveys, which is often not possible for stocks of interest. Furthermore, an important assumption in VPA-like approaches is that selectivity of the oldest age can be related to the selectivity of younger ages each year, often by assuming the two oldest ages have the same selectivity. Results can be sensitive to this assumption, especially in stocks exhibiting low total mortality. Integrated analysis that includes more structure in time-varying selectivity has been used to overcome poor or missing catch-at-age data for specific periods, but such an approach generally assumes that variation in selectivity is similar over the entire modeled period, which may be inappropriate for many applications. A hybrid approach whereby multiple fisheries are used with a combination of time-varying and constant selectivity may be more appropriate. Alternatively, the biological-composition data for some fisheries that have time-varying selectivity could be ignored and a representative selectivity used so fish are removed from the population at roughly the correct ages/lengths. Additionally, down-weighting these biological data is another way to account for time-varying selectivity, but this may lead to (indirect) estimation of selectivity parameters based on relatively uninformative data. The choice among these methods should take into account the quality and temporal coverage of the fishery composition data, as well as the availability of accurate composition data from surveys. Recently, a triple 'separability' approach has been proposed to account for cohort effects, but work is still underway and this method is considered preliminary at this time. In summary, much more research is needed to investigate the appropriateness of time-varying selectivity, estimation of the associated smoothness parameters, and ultimately, the merits and drawbacks of a default assumption of time-varying selectivity as a good practice in stock assessment modeling efforts.

Poor composition data

Composition data (age or length) that strongly influence results from a stock assessment (e.g., as indicated in a profile of likelihood components over fixed values of virgin/unfished recruitment) may be due to how selectivity is parameterized. The effect of composition data on the fit to population indices of abundance should be carefully examined in stock assessments. For example, model runs with the composition data excluded or down-weighted (e.g., by reducing the input sample size; also see Francis 2011), or modeling selectivity in a more detailed manner (e.g., time-varying fishery selectivity) should be considered. These scenarios should be evaluated to see if residual patterns on abundance indices improve and if the model fits are consistent with the assumed variances of the indices.

Management strategy evaluations

Management strategy evaluation (MSE) represents a formal approach for testing the respective roles of assessment models in fisheries management. A MSE is a framework to evaluate catch-determination methods over a broad range of plausible 'states of nature.' The evaluations typically include examining the types of data that are collected and methods used to analyze the data all the way through to management actions that impact actual catches. Alternative selectivity options during the estimation/catch specification stage should be tested using other plausible 'states of nature' from the operating model. One important test would be the robustness of age- or length-based assumptions for selectivity. As noted above, other assumptions to test would include parametric or non-parametric estimation, dome-shaped vs. asymptotic patterns, and time-varying vs. constant selectivity. Such an MSE would provide confidence that selectivity specification is robust in terms of management outcomes and risk. Note that as for any simulation method whereby the estimation model assumes (inappropriately, to some degree) the same model structure that is used to simulate the data (i.e., the 'true' operating model was known), results will necessarily be overly optimistic, given additional uncertainties in natural settings would likely lead to more variable and less robust findings. However, a necessary first-step in this overall diagnostic approach involving simulation is to control for as many factors as possible and minimize potential confounding, interacting, and correlations between model parameter estimates, while focusing on the measurement variables of interest. Ultimately, this simulation design provides meaningful results for examining the quality (precision and bias) of derived quantities used for providing management advice.

Standardizing selectivity in concert with CPUE estimation

Stock assessment models provide estimates of fishing mortality that ultimately depend on assumptions underlying the catch-at-age estimates. Ensuring that removals occur at the correct ages depends on selectivity assumptions and the information available from catch-composition (age or length) data. With adequate age-composition data, assuming the appropriate level of flexibility in time-varying selectivity specifications will help ensure that the catch-at-age and fishing mortality estimates are unbiased, but may mask critical information in the data regarding the values for model parameters. In comparison, assuming constant selectivity will lead to the maximum amount of information being retained about the values for model parameters from the catch-composition time series, but could lead to bias if time-varying selectivity is ignored. The CPUE data are typically standardized for purposes of making catchability (the proportionality constant between CPUE and abundance) as constant as possible across the modeled time period. These standardizations rarely account for the changes (due to selectivity) in the composition data. This may cause an inconsistency in how CPUE indices are included within an assessment compared to how selectivity curves are applied and fit to composition data. In theory, selectivity could be standardized for the same factors as for CPUE by dividing the fleet into multiple fleets, each with their own selectivity (catch, biological-composition data, and nominal CPUE), based on the factors used to standardize the CPUE (e.g., year, season, and spatial unit). In practice, this approach would become intractable in cases with too many factors of interest or if further partitioning of fleets leads to inadequate sample sizes for developing reliable length or age compositions. Sharing selectivity parameters among stocks using a meta-analytical approach may allow for more robust selectivity estimation in these cases. Such an approach is probably more complicated than necessary, but more research is needed to gain insight surrounding selectivity parameterization for standardized CPUE-based indices of abundance.

Survey selectivity

Survey selectivity needs special consideration both outside and inside a stock assessment, given that ideally, the survey should be designed in a manner consistent with the assumption that selectivity is constant over time and asymptotic. Standardized survey gear and protocol may approximately achieve constant ‘contact selectivity’ (see Contact selectivity and availability above), but constant availability is unlikely a typical phenomenon for many fish stocks, particularly for highly mobile species. Therefore, an assumption of constant, asymptotic selectivity for surveys should be carefully scrutinized. Reliable, standardized, and representative age- and/or size-composition data from surveys should allow for robust stock assessments, even when fishery selectivity is assumed to be time-varying or the catch-composition data from the fisheries are down-weighted.

Model selection

Although model selection procedures are widely used when developing statistical models, the overall process is not straightforward, given that most stock assessments include multiple sources of data and complex likelihood specifications. Common statistical tests and evaluations for formal model selection are inappropriate in most stock assessment settings, due largely to difficulty in correctly specifying likelihood functions, sample sizes and variances, and random effects. When presented with alternative selectivity parameterizations, the analyst should consider a broad range of potential diagnostics, including examination of residual and retrospective patterns, profiling over the selectivity parameter space to evaluate influences on other model estimates, and consideration of gear characteristics and fleet behavior. Classical model selection criteria may be considered as another diagnostic tool, but should not be over-interpreted or relied on exclusively due to the difficulty of correctly specifying likelihoods. Future research is needed to evaluate the utility of a generic selectivity form that is robust, broadly applicable in a variety of fishery applications, and implicitly assists in model selection (e.g., estimation of the smoothness parameters of a nonparametric selectivity curve).

Diagnostics

Identifying selectivity misspecification as part of overall model diagnostics is essential for constructing reliable stock assessments. However, in common with many components of stock assessment models, there is little guidance and few objective criteria to diagnose selectivity misspecification. Ultimately, criteria should also provide information for determining possible solutions to the misspecification. The merits and drawbacks of employing classical model selection tools in stock assessments, such as retrospective analysis, residual analysis, and formal goodness of fit information criteria (e.g., Akaike information criterion-AIC, Deviance information criterion-DIC, and Bayesian information criterion-BIC) need further examination before good practice recommendations can be developed. Additionally, it was generally agreed in the workshop that simulation analysis, cross-validation studies, and profiles showing likelihood components (see Poor composition data above) are promising areas for refining model specification practices. Developing good diagnostics for evaluating selectivity and other critical stock parameters involved in modeling fish populations was identified as one of the highest research priorities for future work.

References

M. N. Maunder, P. R. Crone, B. X. Semmens, and J. L. Valero (editors). *In preparation*. Selectivity: theory, estimation and application in fishery stock assessment models. Fisheries. Research.

Fournier, D. A., H. J. Skaug, J. Ancheta, J. Ianelli, A. Magnusson, M. N. Maunder, A. Nielsen, and J. Sibert. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optim. Methods Softw. 27:233-249. Visit the ADMB project website: www.ADMB-project.org.

Francis, R. I. C. C. 2011. Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences 68:1,124-1,138.

Methot, R. D., and C. Wetzel. 2013. Stock Synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research 142:86-99.

APPENDIX A: Agenda

Date and Time	**Presentation**	**Presenter**
11 March (*Monday*)		
8:30 am – 5:00 pm	SS session	Taylor/Lee/Valero
12 March (*Tuesday*)		
8:00 am – 8:30 am	Welcome/Workshop Overview	Methot/Semmens/Crone
8:30 am – 9:30 am	Presentation - A1	Sampson (Keynote)
9:30 am – 10:00 am	A2	Somerton
10:00 am – 10:30 am	Break	
10:30 am – 11:00 am	A3	Schueller
11:00 am – 11:30 am	A4	Song
11:30 am – 12:00 pm	A5	Hurtado-Ferro
12:00 pm – 1:30 pm	Lunch	
1:30 pm – 2:00 pm	A6	Walter
2:00 pm – 3:00 pm	Group discussion- A	
3:00 pm – 3:30 pm	Break	
3:30 pm – 4:30 pm	B1	Ianelli (Keynote)
4:30 pm – 5:00 pm	B2	Iwata
5:30 pm – 7:30 pm	Evening social	
13 March (*Wednesday*)		
8:00 am – 9:00 am	C1	Punt (Keynote)
9:00 am – 9:30 am	C2	Hanselman
9:30 am – 10:00 am	C3	Ichinokawa
10:00 am – 10:30 am	Break	
10:30 am – 11:00 am	C4	Teo
11:00 am – 12:00 pm	Group discussion- C	
12:00 pm – 1:30 pm	Lunch	
1:30 pm – 2:00 pm	B3	Courtney
2:00 pm – 2:30 pm	B4	Crone
2:30 pm – 3:00 pm	B5	Owashi
3:00 pm – 3:30 pm	Break	
3:30 pm – 4:00 pm	B6	Aires-da-Silva
4:00 pm – 4:30 pm	B7	Lee
4:30 pm – 5:00 pm	B8	Thorson
5:30 pm– 8:30 pm	ADMB session Martell/Whitten/Supernaw	

14 March (*Thursday*)

8:00 am – 8:30 am	B9	Martell
8:30 am – 9:00 am	B10	Kinzey
9:00 am – 10:00 am	Group discussion-B	
10:00 am –10:30 am	Break	
10:30 am – 11:30 am	D1	Butterworth (Keynote)
11:30 am – 12:00 pm	D2	Wang
12:00 pm – 1:30 pm	Lunch	
1:30 pm – 2:00 pm	D3	Okamura
2:00 pm – 2:30 pm	D4	Sharma
2:30 pm – 3:00 pm	D5	Stewart
3:00 pm – 3:30 pm	Break	
3:30 pm – 4:30 pm	Group discussion - D	
4:30 pm – 5:00 pm	Closing remarks/adjourn	Maunder

Selectivity workshop – Major sub-topics and keynote speakers

A. Underlying processes (D. Sampson)
   - Characteristics of the gear (e.g., mesh size)
   - Behavior of the fish (e.g., seasonal movement)
   - Spatial structure of the population (e.g., availability/vulnerability)

B. Specification and estimation (J. Ianelli)
   - Functional forms
   - Interactions with related parameters
   - Estimating smoothness parameters
   - Time varying (time blocks, temporal deviates, VPA-like)
   - Size vs. age

C. Model selection and evaluation (A. Punt)
   - Bootstrap methods (error estimation and data set construction)
   - Hypothesis tests
   - Simulation analysis
   - Convergence issues
   - Diagnostics

D. Impacts on management (D. Butterworth)
   - Robustness
   - Management strategy evaluations
   - Biological reference points

APPENDIX B: Presentation abstracts

A. Underlying processes

A1. Title of Presentation: Fishery selection and its relevance to stock assessment and fishery management.

Presenter: Dave Sampson (Keynote speaker)

Authors: D. Sampson

Topic: A

Abstract: Fishery selection (selectivity for short) is the term often used to describe the phenomenon whereby a fish stock experiences vulnerability to fishing that is size- or age-specific. Selectivity operates both at a local scale, as in the direct interactions of individual fish with the fishing gear (gear-selection), and at a stock-wide scale (population-selection), as evidenced by the differential rates of fishing mortality-at-age that are generally observed in stock assessment results. All age-structured stock assessment models have some form of fishery selection to modulate the impact of fishing mortality on differing age-classes, but from a stock assessment viewpoint, selection coefficients are nuisance parameters rather than a focus of attention. We begin with an overview of the three main processes that contribute to and influence fishery selection: (1) physical sorting by the fishing gear or differential responses of the fish to the gear produce the phenomenon of gear-selection; (2) differing selection properties of the different types of fishing gear (e.g., trawl versus longline) in turn generate a composite population-level selection curve that is a weighted average of the different kinds of gear-selection; and (3) when the fish are not well mixed spatially, then the spatial distribution of fishing also affects population-selectivity. A fourth special case arises in species that experience gauntlet fishing as they migrate seasonally. Following the review of the processes underlying selection we explore some of the population-selection curves that have been found in a variety of fisheries. The curves exhibit a wide range of shapes and considerable temporal variability. We conclude with a spatial model for fishery age-selectivity and an exploration of some of its properties. A three-region spreadsheet version of the model is used to demonstrate that the common management reference points MSY, B_{MSY}, and F_{MSY} are functions of both gear-selection and the spatial distribution of fishing, which implies that changes in the spatial aspects of fishing are an additional dimension of uncertainty in our fishery management targets.

A2. Title of Presentation: Review of experimental estimation of survey catchability with a focus on yellowfin sole and snow crab in the eastern Bering Sea.

Presenter: David Somerton

Authors: D. Somerton

Topic: A

Abstract: The experimental methods that have been used to estimate the sampling efficiency of bottom trawls can be grouped into those focused on the various components of the trawl capture process and those focused on the use of an alternate sampling device capable of estimating absolute abundance. The efficiency of the 83-112 Eastern trawl for yellowfin sole was estimated by conducting separate field experiments to estimate herding by the bridles and escapement under the footrope then combining the two estimates in a trawl efficiency model. The efficiency of snow crab was estimated by conducting a side-by-side trawling experiment in which a trawl designed to capture all crabs in its path was towed beside survey vessels conducting normal survey tows. Survey catchability was then estimated as a catch-weighted mean over the entire survey area. The results of the experiment indicated that: 1) trawl efficiency varied spatially with depth and sediment characteristics and 2) the resulting catchability function was clearly not a logistic function of crab size.

A3. Title of Presentation: Determining relative selectivity of the gulf menhaden commercial fishery and fishery independent gill net data

Presenter: Amy Schueller

Authors: A. Schueller

Topic: A

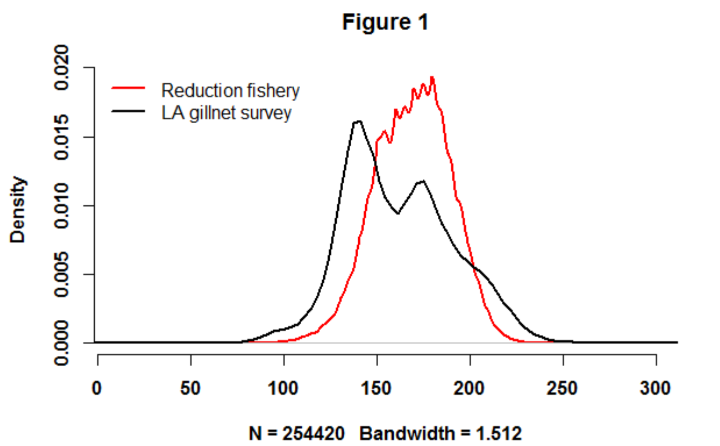
Gulf menhaden are a schooling forage fish that are harvested by one of the largest commercial fisheries by volume in the United States. Purse-seine boats encircle schools, often aided by spotter pilots. The fish are then processed, or "reduced" into fish meal and fish oil and eventually incorporated into products such as poultry and aquaculture feeds and fish oil supplements. The reduction fishery has been routinely sampled since 1964 with dock side port agents taking 10-fish samples, recording length and weight, and removing scales for ageing. Therefore, high quality, long-term data are available to characterize the reduction fishery over five decades. For past stock assessments, the selectivity of the reduction fishery has been assumed to be logistic or flat-topped.

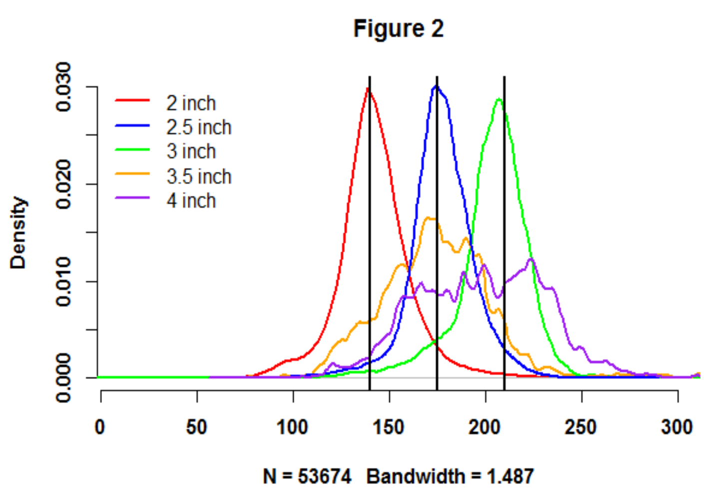
For the latest assessment, fishery-independent gill net data from Louisiana have been considered for creation of an index of adult abundance. Gill nets are fished as strike nets in Louisiana waters and are experimental nets consisting of 5 panels each with a different mesh size. Gill nets have typically been assumed to have dome-shaped selectivity. Unfortunately, age data are unavailable from the gill net survey, although fish lengths are measured.

In order to address the difference in selectivity between the two gear types– purse seine versus gill net, length data were compared between the reduction fishery and gill net survey. The data showed that the gill net survey collects a broader range of sizes than the reduction fishery (Figure 1; red: reduction fishery; black: gill net survey). In fact, modes on the length distribution of the gill net samples suggest ages 1 through 3 are readily identifiable. When average length-at-age from the reduction fishery was compared to the assumed age classes in the gill net survey, it appeared that the fishery harvests large age-1, age-2, and small age-3 gulf menhaden. However, upon further inspection of the length compositions by mesh size, it appears that the "age classes" in the gill net survey are likely relicts of the selectivity of individual mesh sizes in the gill net panels, rather than true age classes (Figure 2).

Thus, the length composition data indicate that the selectivity of the reduction fishery should likely be less than the gill net index for larger and smaller sizes. If the gill net index has flat-topped selectivity that would mean that the reduction fishery should have dome-shaped selectivity and should have a steeper ascending limb than the gill net index. Generally there is no way of defining the functional form of selectivity with certainty; however, the length composition data do provide information on the relative selectivities between the reduction fishery and gill net index.

A potential biological explanation for dome-shaped selectivity for the reduction fishery might be related to the schooling behavior of the gulf menhaden. Gulf menhaden generally school by size. Because of this, the reduction fishery may tend to harvest optimum school sizes with respect to time, effort, oil yield, and other harvest factors. The median size of schools harvested in recent years has ranged 18-23 t. Schools of the oldest and presumably the largest gulf menhaden may be smaller than the optimal school size for harvest. On the other hand, schools of small and younger gulf menhaden, usually observed in large schools and during fall as they exit estuarine waters, are generally avoided by the commercial fishery because of their low oil and protein yields and tendency to "gill" in the meshes of the purse seines.





A4. Title of Presentation: The length structure of bigeye tuna and yellowfin tuna catch at different depth layers and temperature ranges: an application to the longline fisheries in the waters near Gilbert Islands

Presenter: Liming Song

Authors: L. Song and J. Yang

Topic: A

Abstract: Bigeye tuna (*Thunnus obesus*) and yellowfin tuna (*Thunnus albacores*) are the main catch species of longline tuna fisheries in the world. Although tuna longline CPUEs are often standardized by depth or temperature to adjust for the change in depth of longlines, the selectivity by depth or temperature is not changed in the stock assessment. The aim of this study is to evaluate selectivity by depth and temperature to determine if either needs to be considered in the stock assessment. The fishery and environmental data collected from 80 survey sites in waters near Gilbert Islands in 2009 and 2010 were applied to analyze the length structure of bigeye tuna (n=376 individuals) and yellowfin tuna (n=348 individuals), catch at different depth layers (40-200 m, the interval is 40 m, four depth strata), and temperature ranges (25-29 ℃, the interval is 1 ℃, four temperature ranges). A one-way ANOVA was used to test if there were significant differences between the length structure of all samples and the length structure at different depth layers or temperature ranges for bigeye tuna and yellowfin tuna catch, and to test if there were significant differences among the length structures of bigeye tuna and yellowfin tuna catch at different depth layers or temperature ranges. The results showed that: (1) there was no significant difference between the length structure of all samples and the length structure at different depth layers or temperature ranges for bigeye tuna and yellowfin tuna catch ($p \geq 0.05$); (2) there was no significant difference among the length structures of bigeye tuna and yellowfin tuna catch at different depth layers ($p \geq 0.05$); (3) there was no significant difference among the length structures of yellowfin tuna catch at different temperature ranges ($p \geq 0.05$); (4) there was no significant difference among the length structures of bigeye tuna catch at different temperature ranges ($p \geq 0.05$), except the length structures of bigeye tuna catch between 25-26 ℃ and 27-28 ℃ ($p \leq 0.05$). This study suggested that the selectivity by depth or temperature does not need to be included in the assessment of these stocks.

A5. Title of Presentation: Use of multiple selectivity patterns as a proxy for spatial structure

Presenter: Felipe Hurtado-Ferro

Authors: F. Hurtado-Ferro, A. Punt, and K. Hill

Topic: A

Abstract: There is widespread recognition that spatial structure is important for fisheries stock assessments, and several efforts have been made to incorporate spatial structure into assessment models. However, most studies exploring the impact of ignoring spatial structure in stock assessments have developed population models with multiple subpopulations rather than exploring the impact spatial dynamics may have on estimation performance of non-spatially structured assessment methods. Furthermore, the data available to stock assessments usually do not include tagging or other data to estimate movement rates. One approach around this problem is to use several fleets with different selectivity patterns to represent availability within a spatially-structured assessment method. In this study, the impacts of ignoring spatial structure and the effectiveness of using multiple selectivity patterns as a proxy for spatial structure are evaluated for the northern subpopulation of

Pacific sardine (or California sardine; *Sardinops sagax*). A spatially-explicit operating model (OM) is used to explore three spatial factors: the existence of size-dependent seasonal migrations across large geographical areas, the influx of another stock into the area of the assessed stock, and the occurrence of recruitment outside the area where it is assumed to occur. The assessment model is based on the 2010 stock assessment for Pacific sardine, implemented in Stock Synthesis (SS), and includes two seasons per year and six fleets each with a different selectivity pattern. Ignoring spatial structure is found to impact the performance of SS, with seasonal movement having the largest impact on estimation ability. SS compensates for ignoring movement and spatial structure by adjusting the selectivity patterns, but selectivity alone is not able to account for all bias caused by spatial structure.

A6. Title of Presentation: The value of empirical estimates of selectivity in integrated assessments.

Presenter: John Walter

Authors: J. Walter, B. Linton, C. Porch, and W. Patterson

Topic: A

Abstract: What is generally termed selectivity within an integrated assessment model is often a product of two processes; the fraction of the animals in the population available to the gear (availability) and the fraction of animals that encounter the gear that are retained (contact selectivity, sensu Millar). While availability is often difficult to empirically determine, contact selectivity can often be empirically determined from experiments and observational studies. Depending upon how well they reflect the modeled fishery or fleet, empirical estimates of contact selectivity can be used as either direct inputs, Bayesian priors or simply to guide the choice of appropriate functional form for length-based selectivity estimation. Either of the three uses can be exceptionally valuable and influential. We demonstrate the value of empirically derived estimates of hook selectivity for Gulf of Mexico red snapper with a length and age-based SS3 assessment model. By separating selectivity into two component processes, of which contact selectivity is an eminently tractable ground for empirical study, we can greatly reduce one of the key sources of uncertainty within the stock assessment.

B. Specification and estimation

B1. Title of Presentation: Evaluating selectivity trade-offs in groundfish assessments

Presenter: Jim Ianelli (Keynote speaker)

Authors: J. Ianelli

Topic: B

Abstract: “Selectivity is not a well-defined concept” – Dave Fournier ca. 1991. A fundamental aspect of age structured models is the notion that selectivity (or availability) of living marine resources can vary by age or size. It is easy to imagine that this process might vary over time as well. For example, fishing practices which result in targeting abundant spatially-aggregated year-classes over time would cause the relative age-component of fishing mortality to vary over time. For survey data, the age-specific catchability may vary over time if the species characteristics and/or the environment affects the organism’s distribution relative to the survey gear. In this paper we evaluate the pros and cons of the myriad of alternative approaches to specifying selectivity. An example application in which trade-offs to the non-parametric smoothing approach applied to the Aleutian Islands Atka mackerel stock shows the dimensionality and how the interaction of seemingly different processes can occur. Aspects on estimating time-varying dimensions and age-specific smoothing parameters are presented relative to retrospective patterns and key management parameters.

B2. Title of Presentation: Estimation of selectivity in Stock Synthesis: lessons learned from the tuna stock assessment

Presenter: Shigehide Iwata

Authors: S. Iwata, T. Kitakado, and Y. Takeuchi

Topic: B

Abstract: The estimation of selectivity is one of the key issues of a stock assessment since it potentially has a large influence on the estimates of management quantities. Here we present some lessons learned from the tuna stock assessment about the estimation of selectivity, focusing on the assumption of its functional form and the estimation procedure. Firstly, it has been recognized that some non-parametric functional forms (e.g. cubic splines) are quite attractive in terms of their flexibility. However, we observed in the stock assessment of Pacific bluefin (PBF) tuna that they do not necessarily work too well due to the unexpected effects to the model behavior such as non-continuous dynamics of likelihood change by increasing knot numbers, although it was expected to have a better fit than parametric functional form. Secondly, sometimes the balance between likelihood contributions from CPUEs and size compositions is controlled by weighting when they show some incompatibility. To overcome this difficulty, an iterative method for estimating selectivity curves was developed for the 2012 PBF tuna stock assessment. We conducted a small experiment by applying the method to the dataset for the PBF and it showed that the iteration procedure could have a potential to reach convergence and produce somewhat reasonable results. The performance of this method warrants further investigation and should be evaluated further through simulation experiments.

B3. Title of Presentation: Monte Carlo simulation of selectivity and maturity at age in a length-based-age-structured model

Presenter: Dean Courtney

Authors: D. Courtney

Topic: B

Abstract: A length-based age-structured simulation model was developed to investigate the sustainability of Pacific sleeper shark incidental catch in Alaskan commercial groundfish fisheries. The simulation model is governed by a standard set of age-structured population dynamics equations. The relationship between proportions of sharks at age (age frequency) and proportions of sharks at length (length frequency) is modeled using a von Bertalanffy growth equation and an age-length transition matrix. The Monte Carlo simulation was used to verify the expected outcome of including uncertainty in a simulated length at age relationship on the resulting selectivity and maturity at age curves. The expected outcome was less informative (i.e. less steep) selectivity and maturity at age curves than would have been obtained from a length at age relationship simulated without uncertainty. The mean and median values for selectivity and maturity at age from Monte Carlo simulation (n = 10,000) with normally distributed error in the length at age were graphically compared to selectivity and maturity at age obtained from an age-length transition matrix with normally distributed error in length at age. Mean selectivity and maturity at age from the Monte Carlo simulation were approximately equal to selectivity and maturity at age obtained from the age-length matrix with normally distributed error in length at age. In contrast, Median selectivity and maturity at age from the Monte Carlo simulation were approximately equal to selectivity and maturity at age obtained from an age-length matrix without uncertainty in the length at age relationship. These results were consistent with the literature and provided an intuitive example of the effects of including uncertainty in the simulated length at age relationship on the resulting selectivity and maturity at age curves.

B4. Title of Presentation: Age- vs. length-based selectivity for small pelagic fisheries: outside/inside model considerations and management conclusions

Presenter: Paul Crone

Authors: P. Crone, J. Valero, and K. Hill

Topic: B

Abstract: Pacific mackerel are a productive small pelagic species inhabiting the Northeast Pacific Ocean, characterized by highly variable and infrequent recruitment success and associated stock abundance in any given year based primarily on oceanographic conditions and less so, on direct fishing pressure. In this context, determination of appropriate selectivity assumptions and estimators to use in formal fish stock assessments is not straightforward and demands further scrutiny, given both outside and inside the model, plausible scenarios exist for using age or length data in concert with age- or length-based selectivity. The current stock assessment model was simplified by omitting/pooling particular data sources and fixing parameters to produce two baseline models that included either age-composition or length-composition data. Each baseline model was evaluated in terms of age and length selectivity parameterization. A parametric bootstrap procedure within the *Stock Synthesis* modeling platform was used to produce four simulated data sets for examining the quality (precision and bias) of derived management statistics of interest (current spawning biomass, MSY, stock depletion, etc.). The benefits of this approach for conducting future sensitivity analysis and diagnostic examinations surrounding the ongoing stock assessment are discussed in this presentation.

B5. Title of Presentation: Characterizing shape and interannual variability in selection curves of west coast groundfish

Presenter: Brandon Owashi

Authors: B. Owashi and D. Sampson

Topic: B

Abstract: Stock assessments for US west coast stocks of groundfish are generally conducted using the Stock Synthesis program. In applications of this program one generally configures the model and data set to include a small number of fleets that account for the differences in the age-compositions of the catches from the different segments of the fishery. Often the selection curve for each fleet is assumed to be either constant for the entire modeled period or constant for extensive periods with abrupt changes between periods. However, changes in the relative catches among fleets induce changes in the population-level selection curve, which is a catch-weighted average of the fleet-level selection curves. The population selection curve has a direct relation to management reference points such as MSY and B(MSY). Incorrect assumptions about population selectivity could lead to poor estimates of these reference points. This project develops yearly composite selection curves from existing stock assessments in order to characterize the shape and interannual variability in selection curves of west coast groundfish.

B6. Title of Presentation: An exploration of alternative methods to deal with time-varying selectivity in the stock assessment of yellowfin tuna in the eastern Pacific Ocean
Presenter: Alexandre Aires-da-Silva
Authors: A. Aires-da-Silva and M. Maunder
Topic: B
Abstract: Selectivity curves in the yellowfin tuna (YFT) assessment are assumed to be constant over time. However, there may be a strong time-varying selectivity process at play. This is the case of the floating-object (OBJ) fisheries which show high variability in the YFT length–compositions, which result from appearance, disappearance, and reappearance of strong cohorts over time. Misspecified selectivity is not desirable in any stock assessment model since it may cause retrospective patterns and biases in recent recruitments and fishing mortalities, which drive management actions. This paper investigates alternative approaches that could be used to model time-varying selectivity in the YFT assessment. The methods vary from ignoring time-varying selectivity to a full time-varying selectivity process through quarterly changes in selectivity, or wider time-blocks which mark changes in selectivity over time. We chose the floating-object fisheries to illustrate the different methods. A balance is required between the amount of selectivity process (numbers of parameters) that is needed to reduce bias in the recent recruitments, and the amount of OBJ length-frequency data to be used in the model fit (full time series of data or a few terminal years only). This work indicates that allowing for time-varying selectivity (quarterly deviates) in the 5 terminal years of the assessment only while fitting to the length-frequency data available for this period is a reasonable compromise. An "average" stationary selectivity curve is applied to the early period of the assessment with no need to fit to length-frequency data for the early period. This approach seems to greatly minimize the retrospective pattern and improve recent recruitment estimates and fishing mortality rates that are influential in population projection work. Improved estimates of other management quantities are also obtained.

B7. Title of Presentation: Evaluation of a practical method to estimate the variance parameter of random effects for time varying selectivity
Presenter: Hui-Hua Lee
Authors: H. Lee, M. Maunder, A. Aires-Da-Silva, and K. Piner
Topic: B
Abstract: Time varying selectivity may be desirable in many fisheries applications, particularly if fisheries with different characteristics are combined together. Virtual Population Analysis (VPA) inherently allows age-specific selectivity to change from year to year, but results in the loss of a lot of information and may not be practical if age composition data is not available for some years. Also, there may be some fisheries that have fairly constant selectivity from year to year and this consistency in conjunction with the age composition data will provide information on several of population and fishing processes. An alternative to VPA is to treat selectivity parameters as random effects, which is a standard approach in contemporary population dynamics models and is equivalent to state-space models. Inference using random effects models involves integrating out the

random effect (a high dimensional integral), but this can be too computationally demanding in contemporary integrated fishery stock assessment models. Penalized likelihood approaches have been used in fisheries stock assessments (e.g. for annual recruitment variation), but the maximum likelihood estimate of the variance of the random effect is inconsistent and degenerates to zero. A practical method combines the variance of weakly constrained penalized likelihood estimates with the variance estimated by iteratively estimation (i.e. use penalized likelihood to estimate the deviates, calculate the variance of the deviates, use the variance in the penalty function and re-estimate the deviates, and repeat until the estimate of the variance converges) to estimate the variance of the random effect. We test this method using simulation analysis roughly based on the stock assessment of bigeye tuna in the eastern Pacific Ocean.

B8. Title of Presentation: A proposal for penalized-likelihood estimation of semi-parametric models in age-structured stock assessment models

Presenter: James Thorson

Authors: J. Thorson

Topic: B

Abstract: Time-varying selectivity is an active and important area of research in stock assessment. One convenient approach is semi-parametric modeling, which incorporates prior information regarding the functional form of selectivity while also allowing systematic deviations away from this form when appropriate. Gaussian process (GP) estimation represents a gold standard for semi-parametric models, and uses mixed-effects to specify a 'prior' on selectivity while allowing available data to update the prior. However, mixed-effects estimation requires numeric integration, and this will be difficult for many existing stock assessment models. We therefore develop an analogous approach for penalized-likelihood estimation, which uses a penalty on deviations away from the specified form for selectivity. We endeavor to demonstrate that using cross-validation to tune the penalty allows for identifiability of both the parametric 'prior' and all deviations. We conclude by discussing prospects for incorporating this approach into the existing Stock Synthesis software, i.e., by specifying 20% of compositional data as a 'ghost fleet' (i.e. a fleet that does not enter the objective function) and maximizing a profile of the likelihood of this ghost fleet given different values for the penalty.

B9. Title of Presentation: Best practices for modeling time-varying selectivity

Presenter: Steven Martell

Authors: S. Martell and I. Stewart

Topic: B

Abstract: Changes in the observed size- or age-composition of commercial catch can occur for a variety of reasons including: market demand, availability, temporal changes in growth, time-area closures, regulations, or change in fishing practice, to name but a few. Two common approaches for dealing with time-varying selectivity in assessment models are the use of discrete time-blocks associated with an epoch in the history of the fishery, or the use of penalized random walk models for parametric or non-parametric selectivity curves. Time block periods, or penalty weights associated with time-varying selectivity parameters, are subjective and often developed on an ad hoc basis. A factorial simulation-estimation experiment, with discrete or continuous changes in selectivity, is conducted to determine the best practices for modeling time-varying selectivity in fisheries stock assessments. Both the statistical properties of the assessment model and the policy implications of choosing the wrong model are taken into consideration.

B10. Title of Presentation: Selectivity and two biomass measures in an age-based assessment of Antarctic krill

Presenter: Doug Kinzey

Authors: D. Kinzey and G. Watters

Topic: B

Antarctic krill (*Euphausia superba*) sampled over a 19 year period from four areas in the Antarctic Peninsula by the Antarctic Ecosystem Research Division at the Southwest Fisheries Science Center are believed to be part of a larger population of Antarctic krill that is moving through the sampled areas. Two time series of krill biomass, based on trawl nets and acoustic sampling, respectively, are available from each survey. An age-based assessment model coded in AD Model Builder is under development. The model framework currently allows either logistic or double-logistic forms of selectivity. The model integrates size composition data collected by the trawl surveys with two potential measures of biomass: trawl densities and acoustic densities. These two potential measures of biomass are uncorrelated through the time series. The acoustic measures of annual biomass were generally but not always higher than biomass based on nets for the same year. This study evaluates the ability of the integrated model to reconcile differences in the two sources of data for krill biomass by allowing selectivities to be estimated separately for each.

Five model configurations with separate selectivities for acoustic and trawl biomass that differ only by different weightings on these two data sources, including ignoring one or the other, are compared. All configurations used composition data from the net trawls. The configuration using both sources of biomass data with empirical standard errors fit the biomass data from nets and acoustics and the composition data from nets satisfactorily. Configurations forced to fit either the biomass data from nets or the biomass data from acoustics by assigning a CV of 0.01 to one or the other fit that data source very closely at the expense of fitting the other data sources. Configurations forced to fit the acoustic data produced higher estimates of krill spawning biomass than configurations forced to fit the biomass from nets. Models with only one or the other source of biomass data were internally consistent but differed in estimates of krill spawning biomass. Annual spawning biomass estimates in the study area varied between about 2e+06 and 8e+06 tonnes for the acoustic-only biomass model and between about 5e+04 and 3.5e+05 tonnes for the net-only model. Models integrating both sources of biomass data made intermediate estimates of spawning biomass.

C. Model selection and evaluation

C1. Title of Presentation: Model selection for selectivity in fisheries stock assessments

Presenter: André Punt (Keynote speaker)

Authors: A. Punt

Topic: C

Abstract: The choice of how to model selectivity differs among approaches to fisheries stock assessment; VPA tends to make only weak assumptions regarding (age-specific) selectivity (flat selectivity on the oldest ages and temporal stability of selectivity for the most recent years). In contrast, selectivity is more parametric in "integrated" methods and can be age-, length- and age- and length-based. This tends to reduce estimation variation as fewer parameters have to be estimated, but incorrect choices for the functional form for selectivity can lead to bias. This paper illustrates some of the effects of poor choices for selectivity on the outcomes from stock assessments, outlines methods for evaluating whether a particular choice for selectivity is appropriate, and summarizes current ways to select among alternative functional forms for selectivity.

C2. Title of Presentation: Tradeoffs between bias, model fits, and using common sense about biology and fishing behaviors when choosing selectivity forms.

Presenter: Dana Hanselman

Authors: D. Hanselman and P. Hulson

Topic: C

Abstract: The trawl fishery for Gulf of Alaska Pacific ocean perch (POP) has changed over time from a large-vessel foreign fleet, to a large-vessel domestic fleet, to a generally small catcher-vessel fleet since 1960. Trawl survey catchability was drifting higher over time. We found that instead of fitting logistic selectivity for the fishery throughout the time series, fitting a combination of logistic and dome-shaped gamma selectivities had a far superior fit to the data and also alleviated the trawl survey catchability drift. We conducted simple simulations using a POP-like population to test when allowing more complicated selectivity functional forms is both estimable and justified. Data were generated with selectivities from the double-normal mode with error, and models were fitted with the logistic, gamma, exponential-logistic models, and double-normal models. Results were examined for differences in model fit and parameter bias. Estimability was evaluated by examining parameter correlations, uncertainty, and model convergence. The results were used to develop "rules-of-thumb" for what level of true complexity of the selectivity curve justifies applying a complex selectivity curve, or if a simpler curve can be more robust.

C3. Title of Presentation: What does each data component tell us about model misspecification in integrated stock assessment models?

Presenter: Momoko Ichinokawa

Authors: M. Ichinokawa, H. Okamura, and Y. Takeuchi

Topic: C

Abstract: The integrated model has the benefit of integrating multiple data sets such as abundance indices and size compositions. However, the relative weighting among different data sets and high correlation between stock size and selectivity parameters can be problematic. In particular, model misspecification or biased samples can easily lead to erroneous evaluation of the stock. An approach to solve this problem is to 'do not let other

data stop the model from fitting abundance data well' because 'abundance data should have primacy' (Francis 2011, Can. J. Fish. Aquat. Sci. 68:1,124-1,138.). However, size compositions are expected to have specific information on selectivity of fisheries, growth and relative abundance of year-classes, which might eventually affect total stock size estimation. We would like to determine how much model misspecification related to size composition data makes the information contained in the data unusable. In other words, can size composition data tell us anything about stock status under model misspecification? For this purpose, an operating model is established to observe how conflicts occur between abundance indices and size compositions, and between different fisheries targeting different age groups under given scenarios of model misspecification. In the operational model, age-structured population dynamics are simulated to produce observed fishery data (catch, fishery CPUE and catch at length by fishery) for estimating parameters by the integrated stock assessment model, Stock Synthesis fit to length composition data. Various scenarios of model misspecification on selectivity, somatic growth, non-proportionality between abundances and indices, and other important key parameters such as steepness are considered. Additionally, likelihood profiles of focused likelihood components of SS (e.g. size compositions vs. abundance indices) are examined. This analysis shows potential distances from maximum likelihood estimates based on different data sets to the true value, under the condition that model misspecification causes conflicts among different data sets. In addition, the ability or inability to estimate important parameters such as virgin biomass and selectivity parameters simultaneously within the length-based integrated model is discussed.

C4. **Title of Presentation**: Influence of selectivity and size composition misfit on the scaling of population estimates and possible solutions: an example with north Pacific albacore
Presenter: Steve Teo
Authors: S. Teo and K. Piner
Topic: C
Abstract: In the recent stock assessment of north Pacific albacore tuna in 2011, the scale of population estimates and assessment results were highly sensitive to the weighting of size composition data in the model. Therefore, the assessment substantially down-weighted the size composition data (lambda=0.01) in order to constrain the population estimates to a biologically reasonable scale. This is a relatively common situation for assessments of highly migratory species, where differences in the selectivity of various fleets are used as proxies for movements to and from different areas where the fleets operate in. In addition, the selectivity processes tend to be modeled as less variable in time and space than the actual movement processes. This may in turn lead to size composition misfit, which can strongly influence the population scaling. Using the north Pacific albacore assessment as an example, we perform a R_0 profile with respect to the various data components in the model to understand the influence of size composition misfit on the assessment results. Futhermore, we compare several possible solutions to the problem and discuss the pros and cons of each.

D. Impacts on management

D1. Title of Presentation: Fisheries management: does selectivity matter?

Presenter: Doug Butterworth (Keynote speaker)

Authors: D. Butterworth

Topic: D

Abstract: Assumptions about selectivity can be highly influential on estimates of management quantities from fisheries stock assessment models. Selectivity can influence the optimum yields obtainable from a fish stock as illustrated from traditional yield-per-recruit analysis and are related to the age of fish caught relative to the tradeoff between natural mortality and growth. In addition, selectivity interacts with the stock-recruitment relationship through the proportion of the catch that are spawners. We outline the impact of selectivity assumptions on management quantities and provide several case studies to illustrate the impact. We then discuss how management strategy evaluation can be used to determine what harvest rules, data, and assessment methods are most robust to selectivity misspecification.

D2. Title of Presentation: Selectivity's distortion of the production function and its influence on management advice

Presenter: Sheng-Ping Wang

Authors: S. Wang, M. Maunder, and A. Aires-Da-Silva

Topic: D

Abstract: Surplus production models (e.g. the Schaefer and Pella-Tomlinson models) aggregated the dynamics of a fish population into a simple function of abundance and do not explicitly represent biological and fishing processes. It has been clearly shown using age-structured models that the symmetrical production function of the Schaefer model is inappropriate for most fish species and the shape of the production function depends on biological parameters such natural mortality, growth, and the stock-recruitment relationship. It also depends on the age-specific selectivity of the fishery. We evaluate the influence of the selectivity curve on the shape of the production function and compare it with the influence of biological parameters. We then compare results of a stock assessment roughly based on bigeye tuna in the eastern Pacific Ocean when the production function does not match the selectivity curve and when the selectivity curve changes over time. Our results provide one more nail in the Schaefer model's coffin.

D3. Title of Presentation: Evaluating the Sensitivity of Biological Reference Points to Variation in Spatial and Temporal Selectivity

Presenter: Hiroshi Okamura

Authors: H. Okamura, M. K. McAllister, M. Ichinokawa, L. Yamanaka, K. Holt

Topic: D

Abstract: We developed a semi-age structured delay-difference model that takes spatial and temporal selectivity change into account. This model can deal with multiple fishing fleets that have different ages at recruitment and different seasonal and depth preferences. Offshore lingcod data in British Columbia were used as an example in this analysis. The commercial trawl fishery of B.C. offshore lingcod occurs in summer and winter seasons and especially since 2003 mostly in deep water (i.e., greater than 50m). Offshore lingcod tend to show seasonal vertical migration which is different by sex: most adult males aggregate in shallow water and most adult females aggregate in deep water in winter while both distribute equally in deep and shallow water in summer. Male and female juvenile lingcod distribute in only shallow water in both seasons. We examined the sensitivity of biological reference points of B.C. offshore lingcod to assumptions made about migration patterns and effort allocation across seasons and depths. Using the migration pattern which assumes males and females are equally distributed in shallow and deep water in both seasons, estimated MSY and SPR at MSY were robust against changes of effort allocation. However, using the migration pattern which assumes 95% adult males distribute in shallow water in winter and 95% adult females distribute in deep water in winter, estimated MSY and SPR at MSY were markedly sensitive to changes in effort allocation. In particular, SPR at MSY could vary considerably depending on the effort allocation to different seasons and depths. Considering that %SPR is widely used as a proxy for MSY reference points, this result suggests that incorporating spatial and temporal selectivity appropriately into stock assessment models could improve evaluations of management options for B.C. lingcod and other highly migratory species.

D4. Title of Presentation: Indian Ocean bigeye tuna stock assessment: weighting length composition versus CPUE data; Issues on longline and purse seine selectivity

Presenter: Rishi Sharma

Authors: R. Sharma, A. Langley, M. Herrera, and J. Geehan

Topic: D

Abstract: A one area assessment for bigeye Tuna (*Thunnus obesus*) in the Indian Ocean is examined in relation to shape and time varying parameters for selectivity in the major longline and purse seine fisheries in the Indian Ocean. The logistic, double normal, and cubic spline selectivity functions are examined for the different fleets, namely the longline, and purse seine fleets. While marginal improvements are made in the fits to the length composition data when time varying selectivity is implemented in the SS-III framework, by far the larger effect is the weighting of the length composition data in relation to the index of abundance data. Contradicting signals between these two sources have a large effect on the spawning biomass dynamics, and inference based on these weightings can produce different management conclusions. There is a high degree of uncertainty associated with the process of sampling the length composition of the catch from the main fisheries and it is likely that the sampling procedures have changed over the history of the fishery. Further, subtle changes in the selectivity of the composite fisheries may be influencing the length composition of the catch. Alternate hypothesis are derived from the index of abundance and length-composition data. However, based on inconsistent sampling and uncertainty in the length-frequency data, we chose to discount the length-composition data, and changes observed and go with a time invariant selectivity in our assessment. Capturing the signal of the abundance trend was more important than capturing the length frequency as it provided management targets that seemed highly unrealistic given the current levels of catch observed in the fishery.

D5. Title of Presentation: A retrospective investigation of selectivity for Pacific halibut

Presenter: Ian Stewart

Authors: I. Stewart and S. Martell

Topic: D

Abstract: The Pacific halibut stock assessment represents a particularly challenging selectivity application. Contributing factors include: extremely pronounced temporal changes in size-at age, relatively late (age 6+) appearance of fish in survey and fishery data, a minimum fishery size limit, spatial heterogeneity in demographic parameters, and pronounced spatial trends in population abundance over time. Historical stock assessments have variously modeled selectivity as a function of size or age, and also employed nonparametric forms in attempting to account for these various factors. Despite these efforts, a strong retrospective bias in model results was detected during three separate time periods that ultimately required modification of the selectivity parameterization to ameliorate. A summary of historical and current approaches, along with some of the methods employed to explore the most recent retrospective pattern will be presented.

APPENDIX C: Focus questions

Focus questions and answers are presented under the four major sub-topics of selectivity addressed during the workshop. It is important to note that it was beyond the scope of this workshop to produce an exhaustive list of questions/answers that have been thoroughly vetted and can serve as general consensus. Rather, researchers should consider this information as an initial step to gain insight into appropriate practices for modeling selectivity in a variety of fishery settings, particularly those based on contemporary, fully-integrated stock assessment models. Collectively, the following list represents important questions with tentative answers based on workshop discussion. See Background and Summary above for related information.

A. Underlying processes

- The definition and scope of selectivity was generally discussed at the onset of the workshop. It was agreed that selectivity was best addressed during the workshop as parameterized in age/length-structured, multiple data source, and integrated fishery assessment models. In such models, selectivity parameterization typically describes both contact selectivity (i.e., the relative probability that fish of length *l* or age *a* are caught following contact with a fishing gear), as well as fish availability (i.e., the relative probability that a fish of length *l* or age *a* was available to the gear.
- Although underlying processes can provide insight on the *a priori* expected shape of selectivity curves, the combination of spatial processes, both in the biology of the fish (e.g., migration, spatial structure, etc.) and fishing process can result in unique selectivity forms. For example, it may be expected that asymptotic selectivity is warranted based on characteristics of the gear itself. However, in the assessment model, some fits may indicate dome-shape selectivity due to differences in the availability of fish spatially and/or temporally by age/size.
- It is common practice to use fleets as a proxy for spatial processes exhibited by the fish or fishery, but more research is needed to identify cases where this is a robust practice and not contributing to further misspecification.
- Gear experiments can provide insight on some components of selectivity, but other processes cannot be directly addressed through gear experimentation. Although tag-recapture data could provide information on selectivity in general, workshop discussion was limited on this topic.
- On first principles, the actions of fishermen, managers, and scientists should be expected to produce, to some degree, real changes in selectivity over time. Interpretation of past and expected impacts of stakeholders' and analysts influence on the fishing process will require a clear understanding of their goals and incentives.

A.1) Can underlying processes be used to determine a selectivity form (e.g., dome-shape, asymptotic, non-parametric, etc.)?
Answer: Physical gear processes estimated from direct gear experiments can be used to define one component of the selectivity process. However, it is unlikely that the effect of availability due to spatial distribution of different ages/sizes can be estimated externally of the assessment model. If multiple gears are employed in the same area and selectivity due to gear effects can be estimated outside the model, then it may be possible to better evaluate the extent availability plays in defining selectivity appropriately in a stock assessment. In such cases, the spatial component of selectivity could be shared for all gears operating in similar areas, with the gear component of selectivity shared across all similarly operating fisheries. In contemporary stock assessment models, it is assumed that it is possible to harvest all fish in a specified area and time period with infinite fishing intensity and low selectivity. In contrast, availability determines if a fish is available to the gear and the fishing process, but this phenomenon is not modeled explicitly in stock assessments.

A.2) Can gear selectivity experiments be used to estimate selectivity used in stock assessment models?
Answer: Gear selectivity experiments cannot be used to represent all the components represented by selectivity in a stock assessment model. That is, in the vast majority of cases, gear experiments address specifically contact selectivity (and not fish availability). However, such experiments may be useful and informative to estimate a component of selectivity when developing stock assessments.

A.3) Does targeting of strong cohorts influence selectivity, how common is it, and how can one determine confidently that it is occurring?
Answer: In most cases, it should be expected that targeting strong cohorts will lead to changes over time in selectivity, with increased selection for older individuals as a cohort ages. Targeting strong cohorts is common, particularly, in pelagic species that school by size (e.g. sardine, hake, tuna, Pollock, etc). Finally, James Ianelli presented a ‘triple separability’ approach to parameterize selectivity in a way that takes cohort effects into account (see Appendix B, Presentation B1 above).

A.4) Is spatial distribution by fish age/size or gear characteristics most influential on selectivity?
Answer: Both fish biology and fishery gear factors can have a substantial impact on selectivity. For example, the mesh size of trawls and gillnets and escape gaps in pots and traps obviously exclude small fish from the catch. However, availability is driven largely by intrinsic (biological) factors, which are more difficult to examine objectively and likely play a more important role than extrinsic (gear) factors in many fisheries. It is important to note that spatial distribution and in effect, availability, is likely an outcome of behavioral characteristics of the fish as well. For example, small bigeye and yellowfin tuna are attracted to floating objects where they are

vulnerable to purse seines, while large bigeye are generally not available to purse seines, but caught with longline operations and large yellowfin generally associate with dolphins where they are vulnerable to harvest by purse seines.

A.5) What actions can fishermen, managers, and scientists, take to deliberately influence population selection?
Answer: There are a variety of tools managers can use to influence population selection, but the effectiveness of each method is species- and fishery-specific. Some examples are: minimum mesh size and escape gaps to exclude smaller fish; spatial or seasonal closures will allow certain size/age fish to be avoided (e.g., marine reserves for species that segregate by size/age spatially and/or temporally); and banning or modifying gear can limit the catch of certain size/age fish. Ultimately, the actions that can be taken will depend on the overriding goals of the fishery of interest (e.g., maximize overall MSY, maximize MSY by gear/fleet/areas, or minimize bycatch)

A.6) What will be the effect of a no-take marine reserve be on population selectivity?
Answer: The effect of no-take marine reserves on population selectivity will depend on the size of fish in the reserve compared to outside the reserve, movement, and exploitation rates. Some degree of dome-shaped selectivity may be expected in such cases, given availability differences, inside and outside of the marine reserve.

A.7) Does modeling selectivity adequately account for spatial differences in age structure and movement?
Answer: Using fleets as a proxy for spatial processes is a common practice in stock assessments. However, preliminary simulation experiments for Pacific sardine (see Appendix B, Presentation A5 above) indicate that capturing spatial structure using fleet selectivity was not able to adequately account for spatial differences in age structure, migration, sampling, and related fishery processes. The ability for selectivity to explain potential spatial differences in age or size structure is likely fishery-specific and will require further research before forming a broadly applicable answer.

B. Specification and estimation

- There has not been adequate research to date to recommend a selectivity specification that performs well under most circumstances. There are several parametric, non-parametric, and semi-parametric forms available for modeling selectivity in stock assessments, with the ‘double-normal’ distribution (e.g., *Stock Synthesis*) providing one of the most flexible and commonly used forms.
- Researchers should be cognizant of inherent tradeoffs between selectivity flexibility and ease of interpretation and communication to stakeholders.

- Size (length or weight), age of fish, and a combination of both have been used to parameterize selectivity in stock assessments. More research is needed to evaluate the interactions between selectivity type (e.g., age- and/or size-based), use of available data (age and/or size), and population processes (e.g. growth, natural mortality, migration, and recruitment) in stock assessments. Such interactions are likely to be assessment specific. In this context, caveats will be needed in developing good practices, e.g., information regarding the specific conditions for which a recommended practice was arrived at.

B.1) When should one use length- or age-based selectivity?
Answer: On first principles, the choice of length- or age-based selectivity should involve knowledge concerning the mechanisms and underlying processes that ultimately govern the fishing process. Preliminary research indicates that if only age-composition data are used, either choice for selectivity performs similarly, given the analytical mapping of length into age selectivity involved in age-structured models. However, if length-composition time series are used, the biological attribute of choice will often lead to different conclusions regarding the status of a stock. With age-based selectivity, all lengths that map into a similar age will have the same selectivity. However, if length-based selectivity is assumed, selectivity estimates will be different, i.e., the fit to the length composition data will be different even if the length distribution reverts to a normal curve in a subsequent time period. Estimation of growth, particularly if it is time varying, may lead to greater differences between age- and length-based selectivity. It is important to note that both length- and age-based selectivity may be operating simultaneously. It is possible to model length and age selectivity in a stock assessment, e.g., using length selectivity to address gear-related phenomena and age selectivity to account for availability. This general approach may be suitable in some cases. However, there may be interactions between fish size and age that affect availability. For example, directed ontogenetic movement within a population, whereby larger fish move faster and farther than smaller fish of the same age, suggesting that fish farther away have faster growth, when it is plausible that this is an outcome from differential movement of larger fish of the same age. In effect, such a supposition may introduce additional bias in terms of attempts to accurately characterize spatial structure for selectivity, as well as growth attributes (e.g., size-at-age). If selectivity is size based, there will be a difference in the distribution of length-at-age in the catch vs. the population at large. That is, the estimated growth from a length-at-age relationship derived via a size-selective gear will be different than the growth of individuals in the population.

Age-based selectivity may not be appropriate in situations where there are large differences in estimated selectivity in consecutive ages or over ages where size is generally similar (e.g., older fish in a population). The ASCALA model included a length-based smoothing penalty in addition to age-based ones to ensure ages with similar sizes have roughly equivalent selectivities. In many age-structured stock assessment models, it is assumed that for each time step, the distribution of length-at-age reverts back to a normal distribution that does not change over time.

However, this assumption may be violated substantially in particular settings if length-based selectivity is used (e.g., knife-edge selection and high fishing intensity), given size-specific fishing mortality will change the length-at-age distribution. For such cases, model specification for selectivity should explicitly account for both age and length or potentially, a fully length-based model should be used. *Stock Synthesis* includes a growth platoon or morph option that allows for approximation of an explicit description of the distribution of size-at-age in the model.

B.2) Is it inappropriate to rely on simplifying assumptions when modeling selectivity?
Answer: Simplifications can have a substantial impact on results and necessarily depend on the characteristics of the assessment and the degree of simplification. Some research indicates that data should be disaggregated into multiple fisheries and selectivities shared among fisheries rather than using pooled data. In this way, fits to biological-composition data of individual fisheries can be used for diagnostic purposes and to validate/refute assumptions. Alternatively, it would be useful to first fit a simple model with a single fishery, or perhaps with multiple fisheries but making assumptions about selectivity and ignoring biological-composition data for some or all of the fisheries. In this latter approach, complexity and further processes can be added to the model, rather than starting with a complex model specification that could make diagnostic examinations more difficult.

B.3) What are the relative merits of parametric vs. non-parametric selectivity forms?
Answer: The following points reflect preliminary research concerning the most appropriate form for characterizing selectivity in stock assessment models.
Parametric selectivity

- It may not be flexible enough to adequately model selectivity and in effect, may bias results.
- Often easier to compute and interpret.

Non-parametric selectivity

- Often require estimation of smoothing-related parameters that are relatively difficult to determine objectively.
- Can produce unique patterns that have multiple modes, adjacent bins (lengths or ages) that have considerably different values, and are generally unrealistic and not supported by auxiliary information.
- May lead to model instability.
- Difficult to interpret how uncertainty is characterized and propagated, and may generate unrealistic levels of precision associated with estimates.

Semi-parametric selectivity

- Appears promising at this time, however, research is limited to date.
- Some of these methods default to parametric forms when data are missing and/or uninformative.
- In some settings, these methods will be computationally intensive.

B.4) Can selectivity be directly estimated by comparing the spatial distribution of fishing effort to the distribution of fish at large by size and/or age?
Answer: At the onset, this requires knowledge of the spatial distribution of fish abundance by size or age. For example, data from trawl surveys might be used to represent fish abundance, but this requires information about the selectivity of the survey gear. Spatial distribution of the fishing effort will strictly provide information on the availability of fish and not selectivity impacts associated with the fishing gear. The comparison of biological-composition data from a fishery and a survey operating in the same vicinity, assuming full selectivity for all ages encountered in the survey, may provide information regarding contact selectivity of the gear and perhaps some insight into interactions between fish availability and gear selectivity.

B.5) How can estimation best be enhanced based on directed experiments?
Answer: Gear experiments can be used to model the contact selectivity component of stock assessments, but not fish availability. Some experiments can be conducted for purposes of calibrating/standardizing different fishing gears that may have been employed historically for either surveys or fisheries. However, even if a calibration experiment is performed outside the model, diagnostics should still be conducted internally in the stock assessment model to ensure generated results are meaningful and not contradictory. Estimating changes inside the model addresses the uncertainty in the calibration that would be lost by using a single conversion factor from an experiment. Finally, auxiliary tagging and mark-recapture data could be beneficial for informing selectivity parameterization.

B.6) Should a stock assessment model always have one fishery that is characterized by asymptotic selectivity to stabilize estimation within the overall model?
Answer: This is a common practice. A dome-shape selectivity for all gears will result in estimated 'cryptic' biomass and highly uncertain estimates of stock biomass. Further simulation studies could be used to evaluate the implications of incorrectly assuming a given fishery has asymptotic selectivity.

B.7) Is there a selectivity form that performs well in most fishery settings?
Answer: Data quality and quantity, as well as characteristics and dynamics of the stock and fishery directly impact overall selectivity estimation, which collectively, hamper straightforward determination of a single, flexible, and robust selectivity pattern to employ. Alternative approaches can aid in determining a broadly applicable selectivity form. For example, when developing a stock assessment, start with a simple parametric approach and increasingly add in more complex parameterizations, or start with a more flexible non-parametric approach and systematically move towards a more parsimonious and interpretable model. Again, simulation studies will be useful for examining the general performance of a candidate selectivity curve that can be employed in a variety of applications. See B.3 above.

B.8) How should prior distributions for the parameters of the double-normal curve be specified (e.g., in the *Stock Synthesis* model)?
Answer: Use of prior distributions for parameters of the double-normal curve, as available in *Stock Synthesis*, should be used with caution. Some of the double-normal parameters cannot be estimated with reasonable precision. Construction of priors is necessarily complicated due to the parameter transformations used in the double-normal curve. Some assessment authors have used priors to avoid convergence problems, but comprehensive research has not been conducted to determine the effect of influential priors on models results of interest.

B.9) Is aggregating all fisheries together and modeling time-varying selectivity an appropriate approach or should fisheries be disaggregated into units were selectivity is assumed to be reasonably constant?
Answer: To date, there are several views on this question. Some researchers suggest separating fisheries into as many groups as possible and sharing ('mirroring') selectivity across particular fleets in a systematic manner to evaluate individual fits to fishery data and overall model performance. Others consider that there typically exists so much temporal variation in selectivity that there is no reasonable method to separate the fisheries objectively and thus, one should at the onset, combine and model time-varying selectivity. Most methods to model time-varying selectivity assume that the variability in selectivity patterns is similar over the entire modeled time frame, which is often not the case. Traditionally, VPA-like methods typically aggregate catch and implicitly model time-varying selectivity. A hybrid parameterization whereby multiple fisheries are used with a combination of time-varying and constant selectivity would be another reasonable approach to consider. Alternatively, the age- and/or size-composition data for some fisheries that have variable selectivity could be ignored and a representative selectivity pattern relied on to ensure fish are removed from the population at about the correct size (this approach should consider the quality and temporal coverage of the catch-composition data obtained from both fisheries and surveys).

B.10) When there are no sex-specific biological data, but growth and/or selectivity is thought to differ among sexes, should one: 1) have a single combined-sex selectivity; 2) estimate sex-specific selectivity; or 3) share sex-specific differences from another fleet?
Answer: There has been little research on this question, but there exists potential for substantial bias in results if differences in selectivity between females and males are assumed, without evidence or justification, to be negligible (noting that the less common assumption of estimating sex-specific selectivity when there exists little to no life history differences also has the potential for producing biased results). Assuming the same size-specific selectivity for males and females when selectivity is not simply based on a size-at-age relationship can also potentially produce significant biases in overall model results. Since sex-specific differences are exhibited by many species in terms of size-at-age, spatial distribution, seasonal movement, etc., simulation analysis could provide useful information to better assess the appropriateness of using a split-sex

modeling effort based on differences in selectivity between the sexes. Ideally, sex-specific biological data should be collected in the field for purposes of explicitly modeling selectivity by sex in a stock assessment.

B.11) What methods are available to derive age selectivity from length selectivity?
Answer: There are a variety of approaches that can be used to convert length-based selectivity into age-based selectivity and remove fish correctly for each age in an age-structured model. The most thorough approach is to apply the length-based selectivity curve to the length-at-age distribution for purposes of producing the predicted length-composition data. Subsequently, appropriate conversion to age-specific selectivity is achieved by weighting the length-based selectivity by the proportion of length-at-age for each age class. This overall approach assumes that growth, including changes over time, has been correctly modeled and estimated. In strict terms, incorrectly specifying growth will necessarily compromise this general method for linking length and age selectivity and in effect, produce potentially biased estimates of removals.

B.12) The size/age composition of the fishery catch should be catch-weighted in order to provide unbiased estimates that are representative of the biological characteristics of the catch. However, the fishery CPUE should be area weighted to best approximate results from a fishery independent survey. If fish distribution by age varies over space, will the selectivity curve estimated from the size/age composition samples be appropriate for interpreting the CPUE data?
Answer: It is unclear at this time if these methods and assumptions are true in all applications. However, if the data are analyzed in this manner, it is likely that estimated selectivity for the fishery catch and the CPUE index will be different and in effect, should not be shared, i.e., strictly unbiased methods would involve different age-composition data for the two data components, fishery catch and CPUE, respectively.

B.13) In fitting to CPUE time series, if selectivity is changing over time, how should catchability be standardized appropriately?
Answer: If selectivity is changing over time, it is likely that catchability is varying temporally as well. This may indicate that the CPUE index is not proportional to abundance and thus, one should be circumspect including such auxiliary information in a stock assessment model as tuning indices. Alternatively, time-varying catchability could be modeled in concert with selectivity. Caution should be taken and distinction made between fisheries and surveys when considering allowances for changes in selectivity and catchability over time.

B.14) Should selectivity be standardized using the same factors (explanatory variables) used to standardize CPUE?
Answer: This is difficult to do in practice. One approach would be to create fisheries according to the factors used to standardize the CPUE. Alternatively, the selectivity could be standardized

outside the model to develop a single fishery catch-at-age time series and then used in a VPA-type model.

B.15) How does the influence of the fishing mortality level on the relative fishing mortality-at-age generally impact selectivity specification and when should an exploitation rate be used instead?
Answer: The proportion of the population removed (exploitation rate) is not proportional to the fishing mortality. As the fishing mortality gets very high, additional fishing mortality (effort) has less effect on the proportion of fish removed from the population. Thus, when relying on the separability assumption for fishing mortality, fish associated with a low selectivity will be more likely to exhibit a proportional relationship between fishing effort and exploitation rate than fish that are fully selected. In effect, the relative exploitation rate-at-age will differ from the fishing mortality. It is unclear if this is a reasonable representation of the real processes and further, more research is needed to better understand how this disproportional mortality influences results generated from a stock assessment. This distinction is also important to consider when modeling availability. For example, if one-half the population of adults resides in a marine reserve, no matter how high the fishing effort, the selectivity for adults will never exceed 0.5. That is, this is true for models based on exploitation rates (e.g., Pope's approximation), but not for those based on fishing mortality (e.g., solving the catch equation). Thus, a different approach to modeling fishing mortality (e.g., Pope's approximation) may be warranted when selectivity represents availability, i.e., when the fish are outside the fishing area and unavailable for capture.

B.16) What computational issues should be monitored when developing stock assessment models?
Answer: Care should be taken in establishing proper benchmarks to determine model convergence criteria. For example, the double-normal distribution available in *Stock Synthesis* is very unstable in some applications, which can lead to convergence problems. Scrutiny is needed when deciding on bin sizes for biological-composition data to ensure that abrupt changes in selectivity do not occur within a bin such that there is little to no information regarding the parameter representing this change. Finally, the derivative information generated from ADMB-based models should be closely reviewed to evaluate appropriate model convergence.

B.17) Do selectivity forms perform as expected in posterior distributions?
Answer: Several stock assessment researchers have experienced Markov chain Monte Carlo (MCMC) convergence issues likely attributable to selectivity parameterizations. At this time, it is uncertain if such convergence problems have any effect on the estimation of management quantities. Re-parameterization of selectivities to reduce parameter confounding may help with MCMC convergence.

C. Model selection and evaluation

- Model selection and evaluation are not straightforward for any selectivity parameterization. This is in part due to the underpinnings of fully-integrated stock assessments with multiple data sets, where typical statistical criteria for model selection may not be appropriate, given the ad hoc specifications of likelihood functions, sample sizes, and variances by assessment analysts.
- Simulation analysis can provide insight on the performance of alternative approaches to model selectivity. Operating models should be more realistic to avoid conclusions based on oversimplifications that can mask or ignore critical interactions affecting model performance.
- There were insightful discussions on alternative approaches to deal with problematic age- or size-composition data. There was a general agreement in the workshop to emphasize survey indices of abundance over fishery-composition data. However, no agreement was realized regarding how best to proceed with the weighting exercise.
  - Limited discussion on possible alternatives ranged from down-weighting the fishery-composition data, allowing for time-varying selectivity, and entirely omitting the biological time series and fixing selectivity. See General selectivity specification and estimation above.
- The role that selectivity specification (and misspecification) has in diagnosing retrospective bias in stock assessments was briefly discussed. Participants noted much more research is needed on this topic.
- The underlying mechanisms, merits, and risks (misspecification) of constant vs. time-varying selectivity were also identified as high priority areas in need of further research. Finally, future research should also investigate how to implement time-varying selectivity, e.g., annually, extended time blocks, random walks, etc. See Constant or time-varying selectivity above.

C.1) When selectivity estimation or catch-composition data are problematic should one: a) down-weight the fishery data (relative to indices of abundance), which will lead to model estimation tradeoffs and selectivity parameterization being influenced more heavily by other sources of data in the assessment model; b) model more process (e.g., time-varying selectivity); or c) fix the selectivity and eliminate the composition data?
Answer: One potential approach for addressing this question would be to allow for time-varying selectivity and evaluate likelihood profiles of a population scaling parameter (e.g., virgin recruitment level or catchability) for the biological-composition data to check for consistency with the index of interest. A more systematic approach would be to: add more processes to the model, such as time-varying selectivity (alternative b), and produce likelihood profiles of interest; if unsuccessful, down-weight the composition data (alternative a) and evaluate residual patterns in concert with increases to the weight of the biological-composition data; and finally, if still unresolved, investigate fixing selectivity assuming a few plausible forms and omit fishery-

composition data (alternative c). Problematic age- or size-composition data from fisheries may also indicate a misspecified fleet structure. Care should be taken when defining fleets and should include a good understanding of the actual fishing processes of the vessels, spatial-temporal characterization of the fleets, as well as the data collection schemes and related sampling processes that ultimately influence the resultant biological-composition time series used in the stock assessment model.

C.2) What methods can be used to diagnose selectivity misspecification and assist appropriate specification?
Answer If estimated selectivity indicates dramatic changes from one age or size bin to the next, one should further investigate potential misspecifications in the model. Residual analysis and comparison of estimated effective sample sizes with analogous input sample sizes are useful exercises to identify problems associated with overall selectivity parameterization. However, residuals are seldom independent statistically and the sample size comparisons may prove unsuccessful and could be misleading. More research is needed for developing a list of standard diagnostics for selectivity misspecification.

C.3) When should a visual examination of residual patterns override classical model selection test criteria., such as AIC, BIC, and DIC?
Answer: At the onset, using 'by eye' techniques for identifying broad patterns in residual plots is a worthwhile exercise. More work is needed to better determine the appropriateness of relying on statistical criteria, particularly, in cases involving time-varying selectivity assumptions. Residual patterns do not necessarily inform on misspecification of selectivity specifically, given other processes can be misspecified within the model, such as growth, natural mortality, recruitment variability, migration, etc. The AIC, BIC, and DIC tests strictly depend on correct likelihood specifications (e.g., error structure, variances, sample sizes, and random effects), which is seldom accomplished in contemporary, fully-integrated stock assessments that include several sources of data. In summary, classical model selection test criteria, such as AIC, BIC, and DIC should not be used with the typical 'rules of thumb' denoting 'statistically significant' differences among results, but as general diagnostic guides during the model selection process.

C.4) Time-invariant (constant) fishery selectivity is a strong assertion. Is it justified and how does it influence model results?
Answer: Estimating time-varying selectivity may increase parameter uncertainty, but may reduce bias. It was suggested, based on initial simulation analysis, that the default (for fisheries and not surveys) should be time-varying selectivity to avoid bias. However, more work is needed to better understand the extent to which increased parameter uncertainty contributes to model estimation tradeoffs and impacts on final results generated from the assessment model.

C.5) How much trust can be placed in simulation studies?
Answer: Simulations are only as good as the operating model (OM) used to evaluate/test a hypothesis of interest. The OMs are typically very, if not too simple and can produce simulated data that are too precise and unrealistic. At this time, there is a critical need to develop standards for defining meaningful OMs. Most simulation studies are based inherently on the assumption that the likelihood form for estimation is correct. More work is needed to address potential over-dispersion in the data and to capture the overall sampling process more realistically. Given these limitations, one should avoid too many generalizations from simulation studies. To appropriately evaluate selectivity from simulation analysis, one should include the entire process of developing selectivity curves for a stock assessment model, including how fleets are selected and how overall model selection is achieved (objectives of simulation study). Presently, most models used in simulation studies for stock assessments are founded on age-structured models, i.e., more work is needed regarding simulations based on length-structured models.

C.6) What diagnostic tools should be considered for evaluating parameter estimability within the stock assessment model?
Answer Simulation analysis can be used to determine parameter estimability under ideal conditions (i.e., known model and error structures), but other factors should also be investigated, such as misspecified sampling distributions (e.g., due to influential outliers) and process error. Residual analysis, model stability investigations, and cross-validation studies would be useful for determining parameter estimability.

C.7) For selectivity parameters modeled in time blocks, how many time blocks should be specified, and how should their ranges be determined?
Answer: It may be preferable to employ a random walk for modeling time-varying selectivity rather than time blocks, given the somewhat subjective and ad hoc methods typically used for defining blocks. Alternatively, if using blocks, time periods should be specified based on changes in the fishery (e.g., sudden and substantial differences in mesh size regulations). Stakeholders should be consulted for defining such blocks, since many critical changes in the fishery may exist historically and are likely undocumented. Simple visual residual analysis can also be helpful for identifying appropriate breakpoints for the time block.

C.8) In non-parametric selectivity specifications, how should one do model selection (e.g., DIC), and for selectivity parameters modeled as random deviations, how should the overall error level (σ) be determined?
Answer: Ideally, random deviations should be modeled as random effects and integrated out of the likelihood function so that variances can be calculated. Unfortunately, this is not computationally practical in most contemporary stock assessment models. Grant Thompson (AKFSC) has developed a method to estimate these variances using a penalized likelihood that approximates the random effects approach (see Appendix B, Presentation B7 above). Cross-

validation, implemented by systematically omitting a portion of the biological-composition data to produce data sets for testing, can be used to estimate smoothing parameters for nonparametric selectivity forms.

C.9) What is the descending limb of dome-shape selectivity confounded with: a) the specification of natural mortality (*M*); b) increased age-specific *M* for older fish; c) the asymptotic length (in fully length-structured models)?
Answer: The descending limb of a dome-shape selectivity curve is likely to be confounded with the magnitude of *M*, age dependency on *M*, asymptotic length, as well as other critical parameters of interest, such as catchability for indices of abundance. The effect of asymptotic vs. dome-shape selectivity on profiles of critical model parameters (e.g., unfished virgin recruitment (R_0), terminal year spawning stock biomass, projected catch, etc.) needs further research attention. Also, more work is needed to assess the impact of different data types (vs. forms) on likelihood profiles of interest or if rescaling of absolute abundance is warranted. Finally, growth parameterization and potential confounding with particularly dome-shaped selectivity should be examined.

C.10) Can we develop a standard procedure for quantifying the degree of pattern in the residuals generated from fits to the biological-composition data (across size/age and time), and should it be used if bootstrapping methods are employed?
Answer: Simulation analysis and parametric bootstrap methods typically assume that the error structure is known correctly. Including the characteristics of the patterns visually observed in residual plots (e.g., correlation statistics) would likely produce more realistic estimates of uncertainty from both bootstrap and simulation procedures.

C.11) Are multimodal selectivity curves estimated by nonparametric methods reasonable or do they indicate some form of serious model misspecification? Can this information be used as part of a diagnostic approach?
Answer: This will depend to a large extent on the life history strategy of the species of interest, as well as the underlying processes influencing the specific shape of a selectivity pattern. Preliminary simulation analysis using sub-stocks with limited movement indicates that unexpected selectivity curves with unique shapes are not only possible, but potentially may be typical in practice for many fishery applications. It is important to differentiate between population selectivity forms that are intended to represent overall fishing mortality-at-age (all fisheries combined) from the way that selectivity is used in integrated stock assessment models. In the latter case, different fleets/selectivities would typically be used to describe the sub-stocks with a limited movement scenario as described above, rather than create a combined selectivity curve.

D. Impact on management

- Alternative selectivity parameterizations can have large impacts in the estimation of quantities of interest to management (MSY, F_{MSY}, depletion, etc.) and in the provision of scientific advice.
- Workshop discussion included considerations beyond the statistical characteristics of competing models (e.g., quality of fits of asymptotic vs. dome-shape selectivity) and more importance on the risks associated with modeling choices in the context of potentially misspecified parameters. *"... this is not about the best assessment model, but about the best management"* (Doug Butterworth).
- Surplus production models have been shown to produce biased results in cases where selectivity has changed over time, but is ignored.
- Given the sensitivity of biological reference points and management thresholds to misspecifications in natural mortality and possibly selectivity, a case was made to evaluate the alternative use of other targets (e.g., recovery) and even socio-economic vs. biological tradeoffs.

D.1) Is the assumption that one fishery has asymptotic selectivity always precautionary?
Answer: The assumption that one fishery has asymptotic selectivity will not always be precautionary, given parameter interactions and related bias. It may not be a good practice to use asymptotic selectivity to solely provide stability in the model, without other ancillary information available to support this assumption. Preliminary research indicates that dome-shaped selectivity for fleets is expected under some situations due to incomplete mixing of individuals and spatial heterogeneity in fishing. Assessment scientists should use the best scientific approaches available and not make assumptions expecting precautionary outcomes. Precautionary decisions should be left to managers and thus, assessment analysts should provide results and evaluate risks of potential management actions under alternative model scenarios. A thorough Management Strategy Evaluations (MSE) could be used to choose a model that may be precautionary and have asymptotic selectivity, even if the true underlying dynamics of the fishery indicate a dome-shaped selectivity pattern is appropriate. Noting that in such cases, the MSE process would directly address management objectives and robustness of the overall model to an assumption of dome-shape selectivity. See Asymptotic or dome-shape selectivity above.

D.2) Aside from the implications of non-asymptotic selectivity, are selectivity uncertainties generally of significant concern compared to typical other assessment uncertainties, such as the form of the stock-recruitment relationship, catchability estimation, choice of a F_{MSY} proxy, etc.?
Answer: Unaccounted temporal changes in selectivity can lead to substantial biases in estimates of management quantities (e.g., $B_{MSY,}$ depletion), but limited research on this topic precludes forming a clear understanding of misspecification of selectivity relative to misspecification of the stock-recruitment relationship, natural mortality, growth, etc.

D.3) Is it appropriate to use a surplus production model when the fishery selectivity is different from the maturity ogive or from the survey selectivity, or when multiple fisheries with different selectivities operate on the stock?
Answer: Biological processes, fishing operations and resultant catch, and indices of abundance are all functions of different measures of biomass, with the surplus production modeling approach only representing a single measure of ‘abundance.’ Preliminary analysis suggests that the measure of abundance used to represent the biomass dynamics model and its parameters can have a substantial impact on results from a stock assessment. Finally, using a surplus production model with changes in selectivity may lead to biases given the impact on the production curve of alternative selectivities in relation to the life history of the species.

D.4) Does optimizing fishing mortality rates (*F*) or age-specific selectivity provide the most potential to maximize yield?
Answer: It is unclear if *F* or selectivity plays the most influential role in terms of maximizing yield, but regardless, the answer is very likely to differ among species. Further, optimizing yield in any setting is likely to depend on single vs. multiple species contexts, whether ecosystem implications are considered, if bycatch is a concern, precautionary management and associated constraints on overall levels of fishing intenstity, etc. Recent research suggests that in some cases, a balanced harvest approach, as opposed to a selective harvest scheme, can maximize overall ecosystem yield.

D.5) How do selectivity model assumptions interact with key management parameters? How does one identify these?
Answer: Asymptotic vs. dome-shape selectivity can lead to very different implications regarding the status of a stock relative to biological reference points of interest and conclusions concerning over- or under-utilized fish resources. Time-varying selectivity affects the computation and interpretation of biological reference points and other quantities of interest to management, e.g., decisions are needed regarding which time period is relied on to characterize the selectivity used to compute the management statistics.

D.6) Should forecasts include trends in selectivity?
Answer: More research is needed on this topic. However, recent work indicates that a triple separability approach could be used to propagate cohort effects on selectivity as year classes progress into the forecast period of the assessment model.

D.7) How can impacts of different selectivity assumptions on management be best evaluated?
Answer: A well-designed simulation study can provide meaningful information to evaluate ‘risks’ from a management perspective. Formal MSEs can also assist in identifying management

strategies that are robust to uncertainties and biases associated with misspecification of selectivity.

APPENDIX D: Participants

First	Last	Affiliation	City	State	Country
Alex	Aires-Da-Silva	IATTC	La Jolla	CA	USA
Jon	Brodziak	PIFSC-NOAA	Honolulu	HI	USA
Elizabeth	Brooks	NEFSC-NOAA	Woods Hole	MA	USA
Doug	Butterworth	University of Cape Town	Rondebosch		South Africa
Yi-Jay	Chang	PIFSC-NOAA	Honolulu	HI	USA
Ray	Conser	Fish Stock Assessment Consulting	La Jolla	CA	USA
Dean	Courtney	SEFSC-NOAA	Panama City	FL	USA
Paul	Crone	SWFSC-NOAA	La Jolla	CA	USA
Rick	Deriso	IATTC	La Jolla	CA	USA
Emmanis	Dorval	SWFSC-NOAA	La Jolla	CA	USA
Brad	Erisman	SIO-UCSD	La Jolla	CA	USA
Dana	Hanselman	AKFSC-NOAA	Juneau	AK	USA
Xi	He	SWFSC-NOAA	Santa Cruz	CA	USA
Kevin	Hill	SWFSC-NOAA	La Jolla	CA	USA
Michael	Hinton	IATTC	La Jolla	CA	USA
Daniel	Hoenig	RSMAS / University of Miami	Coco. Grove	FL	USA
Peter	Hulson	AKFSC-NOAA	Juneau	AK	USA
Felipe	Hurtado-Ferro	UW	Seattle	WA	USA
Saang-Yoon	Hyun	University of Massachusetts	Fairhaven	MA	USA
Jim	Ianelli	AKFSC-NOAA	Seattle	WA	USA
Momoko	Ichinokawa	NRIFSF	Yokohama	Kanagawa	Japan
Shige	Iwata	NRIFSF	Shizuoka	Shizuoka	Japan
Darren	Johnson	SIO-UCSD	La Jolla	CA	USA
Doug	Kinzey	SWFSC-NOAA	La Jolla	CA	USA
Toshihide	Kitakado	Tokyo Univ. of Marine Science	Tokyo	Tokyo	Japan
Hui-Hua	Lee	NOAA-PIFSC-JIMAR	Honolulu	Hawaii	USA
Kirk	Lynn	CDFW	La Jolla	CA	USA
Steven	Martell	IPHC	Seattle	WA	USA
Mark	Maunder	IATTC	La Jolla	CA	USA
Jenny	McDaniel	SWFSC-NOAA	La Jolla	CA	USA
Richard	Methot	HQ-NOAA	Seattle	WA	USA
Carolina V.	Minte-Vera	IATTC-UEM	La Jolla	CA	USA
Cole	Monnahan	UW	Seattle	WA	USA
Jeffrey	Moore	SWFSC-NOAA	La Jolla	CA	USA
Hiroshi	Okamura	NRIFSF	Yokohama	Kanagawa	Japan

First	Last	Affiliation	City	State	Country
Devon	O'Meara	SIO-UCSD	La Jolla	CA	USA
Brandon	Owashi	OSU	Corvallis	OR	USA
Kevin	Piner	SWFSC-NOAA	La Jolla	CA	USA
André	Punt	UW	Seattle	WA	USA
Marlon	Roman	IATTC	La Jolla	CA	USA
Dave	Sampson	OSU	Newport	OR	USA
Amy	Schueller	SEFSC-NOAA	Beaufort	NC	USA
Brice	Semmens	SIO-UCSD	La Jolla	CA	USA
Rishi	Sharma	Indian Ocean Tuna Commission	Victoria	Mahe	Seychelles
Tim	Sippel	SWFSC-NOAA	La Jolla	CA	USA
David	Somerton	AKFSC-NOAA	Seattle	WA	USA
Liming	Song	Shanghai Ocean University	Shanghai	Shanghai	China
Dale	Squires	SWFSC-NOAA	La Jolla	CA	USA
Gene	Stebley	SWFSC-NOAA	La Jolla	CA	USA
Ian	Stewart	IPHC	Seattle	WA	USA
Brian	Stock	SIO-UCSD	La Jolla	CA	USA
Matthew	Supernaw	SEFSC-NOAA	St. Petersburg	FL	USA
Dale	Sweetnam	SWFSC-NOAA	La Jolla	CA	USA
Yukio	Takeuchi	NRIFSF	Shizuoka	Shizuoka	Japan
Ian	Taylor	NWFSC-NOAA	Seattle	WA	USA
Steven	Teo	SWFSC-NOAA	La Jolla	CA	USA
Jim	Thorson	NWFSC-NOAA	Seattle	WA	USA
Laura	Urbisci	SWFSC-NOAA	La Jolla	CA	USA
Juan	Valero	CAPAM	Seattle	WA	USA
Russ	Vetter	SWFSC-NOAA	La Jolla	CA	USA
John	Walter	SEFSC-NOAA	Seattle	WA	USA
Sheng-Ping	Wang	National Taiwan Ocean Univ.	Keelung		Taiwan
Lynn	Waterhouse	SIO-UCSD	La Jolla	CA	USA
Athol	Whitten	UW	Seattle	WA	USA
Yi	Xu	SWFSC-NOAA	La Jolla	CA	USA

Remote (WebEx) Participants

First	Last	Affiliation
Steve	Barbeaux	AKFSC-NOAA
Jie	Cao	University of Maine
Rob	Cheshire	SEFSC-NOAA
Nancie	Cummings	SEFSC-NOAA
Martin	Dorn	AKFSC-NOAA
Deborah	Hart	NEFSC-NOAA
Anne	Hollowed	AKFSC-NOAA
Jeff	Isley	SEFSC-NOAA
Mandy	Karnauskas	SEFSC-NOAA
Brian	Langseth	SEFSC-NOAA
Chris	Legault	NEFSC-NOAA
Roberto	Licandeo	University of British Columbia
Sandra	Lowe	AKFSC-NOAA
Gary	Nelson	MA Division of Marine Fisheries
John	Quinlan	SEFSC-NOAA
Adyan	Rios	SEFSC-NOAA
Carlos	Rivero	SEFSC-NOAA
Steve	Saul	SEFSC-NOAA
Paul	Spencer	AKFSC-NOAA
William	Stockhausen	AKFSC-NOAA
Cody	Szuwalski	University of Washington
Grant	Thompson	AKFSC-NOAA
Jack	Turnock	AKFSC-NOAA
Erik	Williams	SEFSC-NOAA
Xinsheng	Zhang	SEFSC-NOAA
Jing	Zhao	Shanghai Ocean University

Group photo from the CAPAM workshop on *Selectivity: theory, estimation, and application in fishery stock assessment models* (NOAA/NMFS/SWFSC, March 11-14, 2013).