Theme session S

Stock assessment methods, model complexity and uncertainty

Conveners: Arni Magnusson (ICES), Patrick Lynch (USA), Erik Olsen (Norway)

Synopsis

A three part session comprehensively examined fishery stock assessments to foster awareness and understanding of recent advances and new assessment techniques. Participants discussed the ability of modellers to decide which model, or suite of models, is sufficient for a particular stock and situation. They also highlighted methods and approaches to quantify and communicate uncertainty at all steps of the fishery advisory process.

Fishery stock assessments are an applied science that is fundamental to sustainable fishery management. Recent advances in stock assessment methods include integrated models with time-varying processes, spatially explicit approaches, state-space algorithms, inclusion of multispecies and environmental processes as well as economic factors, techniques for data-limited situations, and rigorous evaluations of management strategies.

Choosing a model structure and assumptions has a direct effect on resultant management advice, where the appropriate model complexity is based on the data collected, findings from simulation analyses, and available expertise. Conversely, determining an appropriate model structure may help to guide investments in data collection and analytical capacity. Thus, determining the level of assessment complexity is a key decision in the fishery management process. Model ensembles offer a way to base management advice on a spectrum of model structures and assumptions.

The final step in a stock assessment is effective communication of the main results, along with the corresponding uncertainties. Quantifying the uncertainty about stock status and catch forecasts is a challenge because of the accumulating sources of uncertainty in the observed data and analytical assumptions, as well as the intrinsic variability in the biological and human components.

Scientific advice should help stakeholders and managers to understand uncertainty, trade-offs, and risks, so management policies for each stock can incorporate the desired precaution.

The session was organized around the following three themes:

1. **State of the art in stock assessment methods**
   a. Improvements in stock assessment methods, implementation of random effects in time and space, analytical methods for data-limited situations.
   b. Performance evaluation (e.g. MSE) of assessment methods and harvest policies.
   c. Multispecies and environmental processes in stock assessment, as well as ecosystem and economic factors in harvest policy evaluation.

2. **Stock assessment model complexity and model ensemble techniques**
   a. Appropriate complexity of an assessment model for a given stock, the basis of choosing a model, implications for data collection.
   b. Application of model ensemble techniques in stock assessment.
   c. Performance of model ensembles, when applied to historical datasets and simulated data.

3. **Quantifying and communicating uncertainties in stock assessment**
   a. Quantifying uncertainties in fisheries catch statistics and survey data, assessment model choices and assumptions.
   b. Developing harvest control rules and evaluating management strategies.
   c. Quantifying overall uncertainty at the advice level, communication of uncertainties, and interpretation of advice.

Presentations

There were 47 oral presentations on the final program (three of which were withdrawn), and 10 posters (one of which was withdrawn). There were three thematic elements of the session, however, the presentations cannot be categorized accordingly, because many touched on more than one, if not all, of the session themes. Over half of the presentations discussed attributes, advancements, and applications of the more complete stock assessment approaches, such as statistical catch-at-age assessments. Several presentations focused on methods and issues that are specific to the assessment and management of data-limited stocks. Several presentations related to the collection and processing of data inputs for stock assessments, and at least five presentations offered higher-level reviews and perspectives that fell under the session themes.

Highlights

The number and quality of the presentations in the session showed the relevance importance for developing stock assessment approaches within the ICES community. Four key topics focused the discussion at the session: I) characterization and communication of uncertainty, II) limits and possibilities of Virtual Population Analysis (VPA) based assessments, III) trade-offs of when to expand complexity, and IV) ensemble modelling approaches.

I - Characterization and communication of uncertainty

It is necessary to accurately and effectively characterize and communicate the uncertainty around stock assessment results. However, while assessments must be transparent regarding their uncertainty, some managers prefer a single number to support decision-making. Thus, the challenge in communicating stock assessment uncertainty is finding the balance between providing a complete characterization of uncertainty while minimizing confusion. Decision tables have been developed to deal with this challenge by providing single point estimates over a range of scenarios, which allow decision-makers to extract the value needed while understanding the precision around that estimate across a range of plausible assumptions. Currently, decision tables are not being used regularly by many management organizations, in part, because the importance of making management decisions that appreciate the uncertainty around the advice has not been stressed enough. This is in contrast to other disciplines, such as weather forecasting, where meteorologists routinely present the uncertainty of weather forecasts in ways that the public and managers are able to use to make decisions - hurricane forecasting is a timely example where management decisions (e.g. evacuation) are based on advice with clearly identified uncertainties (cone of uncertainty around a hurricane path). Weather and hurricane forecasts are however easier to verify than fisheries stock assessment forecasts as historic weather is known with certainty while stock assessments cannot be verified. Although it is possible to communicate uncertainty, we should be mindful of these issues when deciding how much uncertainty is conveyed and when in the process it is communicated, so as not to overwhelm decision-makers and stakeholder or create an impression that stock assessments are fundamentally flawed. A complete understanding of uncertainty in stock assessments is important for implementing a precautionary approach to

management. This connection to the precautionary approach is perhaps the most important link to showing the relevance of uncertainty in a management setting. Using ensemble modelling (similar to techniques employed by meteorologists) is a viable and publicly accepted approach that stock assessment scientists should explore further to better characterize and communicate the uncertainty of advice.

II - Limits and possibilities of VPA-based stock assessments

With the development of statistical catch-at-age assessment models, state-space algorithms, and spatial assessment models, some assessments still use simpler Virtual Population Analysis (VPA)-based models. To fit VPA models substantial effort needs to be dedicated toward modelling growth and recruitment, and VPA models implicitly model fisheries with time-varying and flexible (e.g., dome-shaped) selectivities. However, it can be argued that too much time is directed toward modelling stock-recruitment relationships and too little toward modelling natural mortality, where a constant value is typically assumed when we know that natural mortality likely varies temporally and spatially. Using yield-per-recruit strategies is one way to get around issues with estimating stock-recruit parameters. Continued research is needed to minimize biases, and integrated assessment models represent a useful approach, especially in situations where we don't have a full catch-at-age matrix.

III - Trade-offs of when to expand complexity

Routine stock assessment is comparable to an industrial process where previously vetted and agreed on models are used to crank out assessments and develop catch advice. An increase in the use of management strategy evaluation (MSE) to test management procedures at the benchmark level prior to conducting routine assessments would result in better advice to managers. MSEs require a realistic operating model, or set of operating models, otherwise the results will be unreliable. Stock assessment models used in routine stock assessment offer one 'vetted' option, but these are perhaps not independent enough from the stock assessment process to be tested in the MSE loop. However, single-stock analysis may be too narrow in scope for MSE as such approaches should consider harvest rates at a system level to address multi-species and ecological interactions. Thus, there is a need to develop clear criteria for choosing an appropriate operating model for a given situation, as well as a process that provides guidance on what level of model complexity is necessary to distinguish changes in abundance and fishing intensity from model noise. Being able to understand the intricacies of the models is paramount; it is not enough to just 'push the button.' Scientists conducting MSEs should understand what actually happens within the model to understand its strengths and weaknesses and thus the limits of its results.

IV - Ensemble modelling approaches

Statistical assessment models and statistically sound sampling regimes allow for estimating the precision in our input data and model estimates. Biases, especially those due to structural uncertainty of model choice, are more difficult (nigh impossible) to evaluate with a single-model approach, which is currently the standard procedure in most stock assessments around the world. Ensemble modelling efforts have shown promising results when applied to a suite of data-limited assessment models, or multi-parameterization of a single model. Weather forecasts routinely use multiple (vetted) models in an ensemble to make weather (or hurricane) forecasts. An ensemble of complex models may be a way of developing an understanding of a system, where simple models just capture basic trends or certain aspects of the system. Perhaps the ensemble of complex models could serve as a guide for which simple approaches can be used. The ICES Advisory Committee has no preferences regarding

single-model vs ensemble-model assessments, but these should be evaluated through the regular benchmark process to be applied in an advisory context. However, benchmark reviews are currently dominated by single model approaches, so overcoming the bias to choose a single model will be challenging. Ensemble approaches carry other important challenges, most notably, that including poor models will result in a poor ensemble result, and also that running multiple models requires more work. However, developing new assessment frameworks will simplify the use of model ensembles. Still, good protocols for selecting and vetting models that are included in an ensemble need to be developed. This approach requires clear commitment from ICES and member states. Developing procedures for ensemble modelling should also consider the model domain to fully address the structural aspects; economists, oceanographers, and meteorologists should be engaged in the process to develop best practice guidelines.

Discussion & way forward

The stock assessment session gathered much interest and was the longest-running session at the 2017 ASC. Together with the open session on stock assessment, key stock assessment scientists from institutions in both Europe and North and South America came together to discuss new research and common issues. Dealing with uncertainties in stock assessments is a priority issue to properly meet the challenges of a precautionary approach to management. This requires both better statistical techniques to improve estimates of precision, but also broader approaches like ensemble modelling to deal with the structural uncertainties and more properly address bias issues. Such new developments also open the possibilities for more advanced and better MSEs.

The session chairs are drafting a mini-review paper based on the session, and are inviting session presenters to contribute to it.

The session also supports the ICES Methods Working group that has its first meeting 13-17 November 2017, as well as the tentative Global Assessment Methods Experts (GAME) initiative that ICES is developing as a follow up from the Strategic Initiative on Stock Assessment Methods (SISAM).