SCIENTIFIC COMMITTEE

NINETH REGULAR SESSION

Pohnpei, Federated States of Micronesia

6-14 August 2013

Approaches to describe uncertainty in current and future stock status

WCPFC-SC9-2013/MI-WP-04

A M Berger, G M Pilling, S J Harley, and C Kirchner[1]

[1] Oceanic Fisheries Programme, SPC, B.P. D5, 98848 Noumea Cedex, New Caledonia

Approaches to describe uncertainty in current and future stock status

Executive Summary

This paper responds to the SC8 request for further development of a common approach to describe uncertainty in current and future stock status. In particular, consideration was given to defining a consistent approach that could be used to evaluate management risks (e.g., the risk of falling below limit reference points). We acknowledge that there are unavoidable tradeoffs between the degree of analytical rigor, computational demand, and time constraints that must be considered when choosing an approach that is practical, yet informative.

In this paper we highlight the relative advantages and challenges associated with common approaches to incorporate uncertainty into fisheries management decision-making. We provide recommendations for describing uncertainty, options for selecting a representative subset of assessment model runs, and discussions on the use of model plausibility weights. Key considerations for this work are also discussed.

Ultimately, the approach used to describe uncertainty will influence perceptions of management risks and the selection of a management strategy (including reference points and harvest control rules) to meet management objectives. In particular, this is critical for measuring the risk of exceeding limit reference points. The validity of using a single model to characterize uncertainty in stock status is questionable. An approach to incorporate uncertainty that takes into account model error (integrating across multiple assessment models) in addition to process and estimation error would be the most sensible while balancing the impracticalities associated with computing/resource time.

With this in mind, the following are recommended guidelines for describing uncertainty in current and future stock status for the provision of management advice to the Commission.

- Use of a hierarchical approach towards uncertainty estimation which involves:
  - selecting a representative subset of models from the structural uncertainty grid to capture the extent of model uncertainty and
  - using stochastic projections on the chosen subset of models from the grid to provide management advice that captures key sources of process uncertainty.
- Stochastic projections should incorporate uncertainty in recruitment as a minimum, but preferably with respect to catchability (effort deviates) and the age structure in the 1st year of the projection period as well.
- Expert opinion be used by the SC to determine a representative subset of models from the grid (perhaps done by individual stock assessment working groups or at pre-assessment workshops) for describing model uncertainty, guided by the assessment document provided by stock assessment scientists.
- Plausibility weights for each model in the subset be determined by the SC on a case-by-case basis (perhaps by the stock assessment working group or at pre-assessment workshops), guided by the

assessment document provided by the stock assessment scientists (the default being that all assessment model runs have equal weight).

The hierarchical approach described above is consistent with that discussed at the pre-assessment workshop held in Noumea from 8-12 April 2013. We note that these guidelines for describing uncertainty were developed as a common approach for evaluating risks associated with candidate management strategies (e.g., risk of falling below the limit reference point). However, we acknowledge that other approaches for handling uncertainty may be more suited for particular analyses.

Introduction

At SC8 it was recommended that further development towards describing a common approach to uncertainty in current and future stock status be conducted for review at SC9. In particular, SC8 was concerned about the estimation of risk in relation to limit reference points and ensuring consistency in the provision of management advice to the Commission.

A central issue within the management of fisheries is acknowledging the many types of uncertainty within the process (Table 1), and to ensure that the management framework is robust and management objectives are achieved despite that uncertainty. Here we focus on ways to incorporate uncertainty associated with process, estimation, and model structural errors into WCPFC fisheries management (defining current stock status and projecting impacts from prospective management options). Other sources of uncertainty, such as implementation error and measurement/observation error, are also very important and arise more prominently when conducting management strategy evaluations.

This paper:

- briefly reviews previously applied approaches in the WCPFC and identifies some alternative approaches;
- highlights the relative advantages and challenges associated with common approaches to incorporate uncertainty into fisheries management decision-making;
- recommends a general framework for describing uncertainty with special consideration given to quantifying risks associated with falling below limit reference points;
- discusses options for identifying a representative set of assessment model runs and defining model plausibility weights; and
- identifies key considerations for discussion.

Many of the approaches discussed are not mutually exclusive, but in fact could be applied in combination or as a hybrid of several approaches. Ultimately, there is a direct link between the perceived risk of a particular management action and the level of uncertainty acknowledged in the scientific analysis. Hence, a framework for incorporating uncertainty should be a prerequisite to consistently evaluating risk and, consequently, ensuring that the risk of exceeding a limit reference point remains at a low level.

Review of approaches used with WCPFC tuna stock assessments

A number of approaches have been used in the WCFPC to characterize the most prominent uncertainties associated with providing management advice. These have predominantly been applied to identify and account for estimation, process, and model sources of error associated with stock assessments. There are advantages and challenges associated with each approach (Table 2 and 3). These must be considered to ensure that the approach is appropriate given the ultimate aims of the analysis and its use to inform management decisions.

When considering current stock status, WCPFC stock assessments typically involve:

- the characterization of estimation error for statistically fitted parameters within a single model (e.g., standard errors or confidence bounds); and
- model error by comparing results across a structural uncertainty grid (Harley et al. 2009).

Likelihood profiles (or the delta method) have also been used to provide error bounds for derived parameters (those not statistically fitted in the model, such as biomass and fishing mortality) so that a statistically plausible range for key management parameters can be developed for specific assessment models. Comparing estimates across a grid of model runs has been a common procedure for evaluating sensitivity to model assumptions, and has been applied to define a range of plausible stock status indicators (e.g., F/F_{MSY} or SB/SB_{MSY}). It has rarely been applied to explicitly account for model uncertainty in quantifying management advice (i.e., through the use of multi-model inference; Burnham and Anderson 2002). The 2012 south Pacific albacore assessment (Hoyle et al. 2012) is one exception where the provision for management advice was deemed to be the median across the grid of model runs by SC8.

When considering projections of how a fish stock might be expected to react to a particular management decision or under certain conditions, uncertainty has been included in one of two ways:

- by conducting deterministic projections across assessment models in the uncertainty grid. In this case, a single set of input parameters (from the assessment estimation period) are used for the projection period for each estimation model in the grid; and
- by conducting stochastic projections, which has typically been applied to a single model, frequently the agreed reference case (Davies and Harley 2010). In this case, many sets of input parameters are generated (from the estimated variance-covariance matrix) and applied to the projection period for that single assessment model.

Deterministic projections have typically been used to capture model uncertainty, but ignore key sources of process uncertainty. In contrast, stochastic projections have typically been applied to account for key sources of process uncertainty (i.e., unknown population dynamic processes) which are integrated as the stochastic inputs, but have not accounted for model uncertainty (one model is selected for the projections).

Alternative approaches

There are alternative approaches for including key sources of process, estimation, and model uncertainty into fisheries management advice, each with its own set of advantages and challenges (Table 2 and 3). We briefly highlight below some of the more widely used techniques for characterizing uncertainty that could be applicable to WCPFC stock assessments for quantifying management risks associated with limit reference points. In some cases, assessment model complexity (e.g., number of parameters and data sources) and computation demand can be limiting factors.

- Weight alternative assessment models for a stock (e.g., the 'grid' of runs) according to how plausible each is considered to be. Plausibility could be defined using likelihood (information-theoretic) approaches or by expert judgment (e.g., stock assessment scientists or an SC working group). Inferences would incorporate model uncertainty and model inequity by integrating resulting management quantities across models according to the specified weights (i.e., multi-model inference).
- Define a small subset of assessment models from the uncertainty grid that adequately characterize overall model uncertainty as a means to accommodate both model (representative subset) and process (stochastic inputs) uncertainties in combination when conducting

management scenario projections. Options for selecting a representative subset of assessment models are discussed later.

- Use statistical computing such as Monte Carlo Markov Chain (MCMC) methods to characterize distributions of parameters. Distributions can then be used to describe estimation error (estimated and derived parameters) or can be readily sampled/resampled to implement process errors for developing stochastic inputs for projection scenarios or for developing operating models for management strategy evaluations.
- Apply a hybrid of the above techniques that attempts to maximize the advantages of each approach while overcoming some of the impeding challenges (e.g., computational demand).

There are unavoidable tradeoffs between the degree of analytical rigor, computational demand, and time constraints that must be considered when choosing an approach that is practical, yet informative. For example, computing constraints make it difficult to run stochastic projections across the full uncertainty grid (e.g., running 200 (or more) stochastic simulations for each of 540 models is impractical, and using this approach for multiple species compounds this), or to apply MCMC methods using the types of assessment models currently used to assess WCPFC tuna stocks. However, it may be manageable to run stochastic projections across a subset of models from the grid (e.g., see Figures A1 and A2, SC9-MI-WP-03). The difficulty arises when choosing the subset of models and deciding if each is equally plausible (i.e., weighted equally) or not (weights specified for each individual model).

Guidelines for describing uncertainty

Uncertainty and risk are closely related. Therefore, a standard approach for describing uncertainty is needed to provide a basis for expressing management risks in a consistent manner for the provision of advice, such as the risk that a limit reference point has, or is likely to be, exceeded. We note that a risk-based approach for management in the WCPFC (e.g., reference points and harvest control rules) rules out the single model approach (running deterministic or stochastic projections from a single model run), because it does not capture model uncertainty.

Stochastic projections do provide the best format currently available for introducing process errors in key population dynamic parameters (e.g., recruitment and catchability) which is necessary given there are processes we know little about (e.g., steepness and time-varying catchability) and that the variability among structurally different model runs is typically much greater than parameter uncertainty within models. Given that stochastic projections are computationally restrictive when applied over many assessment models, a representative set of assessment models (e.g., a subset of the full grid) that adequately captures the uncertainty in the stock assessment would need to be chosen.

Ideally, we would like to characterize the distribution of each management indicator of interest across the set of assessment models to assign probabilities to particular values (or range of values) when evaluating risks (Figure 1). However, determining the underlying distribution is often impossible, particularly given that many model runs occur as discrete categories (e.g., steepness), leaving interpolation to formulate a continuous distribution. While an underlying distribution could be assumed, if it is incorrect, biased estimates can result. A pragmatic solution to uncertainty, as recommended below, is likely to perform just as well.

Because consistency (and plausibility) in the approach are vital, objective criteria must be considered, which includes the development of guidelines for those models to include from the structural uncertainty grid. We recommend:

- Use of a hierarchical approach towards uncertainty estimation which involves:
  - Selecting a representative subset of models from the structural uncertainty grid to capture the extent of structural uncertainty.
  - Using stochastic projections on the chosen subset of models from the grid to provide management advice that captures key sources of process uncertainty.
- Stochastic projections should incorporate uncertainty in recruitment as a minimum, but preferably with respect to catchability (effort deviates) and the age structure in the 1st year of the projection period as well.

These recommendations are consistent with those discussed at the pre-assessment workshop held in Noumea from 8-12 April 2013 (OFP 2013). We note that these recommendations were developed with specific reference to quantifying management risks in a consistent way. Other approaches for handling uncertainty may be more suited for particular analyses.

Options for selecting a subset of assessment model runs

Several options are available for selecting a representative subset of models from the structural uncertainty grid in WCPFC stock assessments. There is a need to specify formal guidelines for selecting the subset of model runs to ensure that the set is not chosen *ad hoc* (or in a subjective manner), which could lead to inconsistencies in management advice, or to encourage a particular result.

We highlight options below that have potential for application within the existing framework used by the WCPFC for establishing the provision of management advice.

- A comprehensive meta-analysis to assign weights to different assessment models (based on information from other fisheries and science papers on the general influence of particular parameters) and then select those models with the highest weight. However, it may be difficult to be objective and applicability to WCPFC tuna stocks is unclear.
- The use experimental design approaches such as 'fractional factorial designs' (FFD; Montgomery 1991; Hoyle et al. 2008) to reduce the number of assessment model runs by a factor of 2^n; where *n* is the number of axes in the uncertainty grid and 2 refers to the number of factors associated with each axis. However, most assessment grids have more than two factors associated with each grid axis (e.g., steepness of 0.65, 0.8, and 0.95) so the application of the FFD approach for reducing WCPFC assessment model runs is not straightforward.
- The use of expert opinion to define a subset of grid runs that are deemed to be representative of the full uncertainty grid or the range believed to be plausible given the stock assessment. Expert opinion could come from the scientists conducting the stock assessment or, preferably, the Science Committee (e.g., selecting the 'key' model runs to characterize uncertainty could be an additional task assigned to assessment working groups).

- The assigning of plausibility weights to assessment models from either the full grid or a subset of the grid. This implicitly suggests that not all assessment models are equally plausible, and that the characterization of uncertainty should take into account assessment model inequity.
- A combination of the above techniques.

Model-based approaches such as the use of likelihood and information theory (i.e., model selection techniques; Burnham and Anderson 2002) to select the 'best' models are well defined and used widely, but are limited to situations where comparisons are made across assessment models using the exact same data. Most WCPFC stock assessment model runs include sensitivities to alternative data sources or formats so this is not a viable option.

Considerations

There are several considerations that warrant discussion at SC9.

1. The recommended hierarchical approach to describe uncertainty:
   - select a representative subset of assessment model runs (capture model uncertainty); and
   - apply stochastic simulations across the subset of models (capture process uncertainty) that integrate over recruitment (also catchability (effort deviates) and the age structure in the 1st year of the projection period when feasible).

2. The option for selecting a representative subset of assessment model runs (if deemed applicable by SC9):
   - several options exist, which is the most appropriate for the WCPFC?
   - we recommend at this time that expert opinion be used by the SC to determine a representative subset (perhaps done by individual stock assessment working groups or at pre-assessment workshops) for describing model uncertainty, guided by the assessment document provided by stock assessment scientists (see 'expert opinion poll' example in Appendix 1).

3. Weighting assessment model runs to establish inequities in model plausibility (if deemed applicable by SC9)
   - should weights be assigned to assessment model runs (across the grid or a subset of grid) given that weights are difficult to determine even with the aid of computer intensive statistical applications such as MCMC, which are currently impractical for WCPFC tuna stock assessments
   - we recommend that SC determine whether to use plausibility weights on a case-by-case basis (perhaps by the stock assessment working group or at pre-assessment workshops), guided by the assessment document provided by the stock assessment scientists (the default being that all assessment model runs have equal weight)

Intuitively, weighting alternative models from the grid seems like a logical way forward. However, it is not a trivial matter to determine these weights, often times selected in a subjective manner (Maunder 2012).

The use of weights to determine model plausibility is an important consideration and should be discussed further at SC9.

Ultimately, the approach used to describe uncertainty will influence perceptions of management risks and the selection of a management strategy (including reference points and harvest control rules) to meet management objectives. In particular, this is critical for measuring the risk of exceeding limit reference points. At a recent international workshop on tuna RFMO reference points and harvest control rules (Anonymous 2013), it was noted that when calculating probabilities relative to reference points (such as the risk of exceeding the limit reference point) model uncertainty should be included to the extent practical (Anonymous 2013). Consequently, the validity of using a single model to characterize uncertainty in stock status is questionable (Harley et al. 2009). An approach to incorporate uncertainty that takes into account model error (integrating across multiple assessment models) in addition to process and estimation error would be the most sensible while balancing the impracticalities associated with computing/resource time.

For WCPFC tuna stocks, computational time is often the limiting factor due to the complex nature of the stock assessments. The computational burden is exacerbated when considering dynamic projections (e.g., those used to conduct harvest control rule evaluations) because of the iterative nature of projections that include management interventions through time.

Recruitment, the age structure in the final year of the assessment, and catchability (or effort deviates) have been previously identified as key population dynamic processes to consider (Davies and Harley 2010).

References

Anonymous. 2013. Report of the 2013 ISSF stock assessment workshop: harvest control rules and reference points for tuna RFMOs. ISSF Technical Report 2013-03. International Seafood Sustainability Foundation, Washington, D.C., USA.

Burnham, K.P., and Anderson, D.R. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer, New York, NY. 2nd ed.

Davies, N., Hoyle, S., Harley, S., Langley, A., Kleiber, P., and Hampton, J. 2011. Stock assessment of bigeye tuna in the western and central Pacific Ocean. WCPFC-SC7-2011/SA-WP-02.

Davies, N., and Harley, S.J. 2010. Stochastic and deterministic projections: a framework to evaluate the potential impacts of limit reference points, including multi-species considerations. WCPFC-SC6-2010/MI-WP-01.

Harley, S.J., Hoyle, S.D., Hampton, J., and Kleiber, P. 2009. Characteristics of potential reference points for use in WCPFC tuna stock assessments. WCPFC-SC5-2009/ME-WP-02.

Hoyle, S. D., Bouyé, F., Langley, A., and Hampton, J. 2008. General structural sensitivity analysis for the bigeye tuna stock assessment. WCPFC-SC4-2008/SA-WP-03, Port Moresby, Papua New Guinea, 11-22 August 2008.

Hoyle, S., J. Hampton and N. Davies. 2012. Stock assessment of albacore tuna in the south Pacific Ocean. WCPFC-SC8-2012/SA-WP-04-REV1.

Maunder, M. N., Aires-da-silva, A., & Deriso, R. 2012. A critical evaluation of the construction of the Kobe strategy matrix: lessons learned from bigeye tuna in the eastern Pacific Ocean. IATTC Scientific Advisory Committee Draft Report. SAC–03–06c.

Montgomery, D. C. (1991). 'Design and analysis of experiments.' Third.Edn. (John Wiley and Sons: New York.)

OFP (Oceanic Fisheries Program, SPC). 2013. Report from the SPC pre-assessment workshop, Noumea, April 2013. WCPFC-SC9-2013/SA-IP-01.

Rosenberg, A. A., and Restrepo, V. R. 1994. Uncertainty and risk evaluation in stock assessment advice for U.S. marine fisheries. Canadian Journal of Fisheries and Aquatic Sciences, 51: 2715-2720.

Table 1. Key uncertainties for fisheries management that may manifest errors and misguide management advice (adapted from Rosenberg and Restrepo 1994).

Uncertainty	Description	WCPFC example
Process error	Natural variation	Year-on-year variation in number of young fish produced
Measurement error	When collecting information	Species composition in purse seine catches
Estimation error	When modelling natural processes	Fitting movement models based upon tagging information
Model error	When assuming that an assessment model mimics real life	The Multifan-CL model and assumptions on spatial structure
Implementation error	Management decisions are never implemented perfectly	Challenges in implementing all CMM requirements

Table 2. General characteristics of some common approaches to incorporate uncertainty (process, estimation, and model errors) into fisheries management decisions about current stock status. The advantages and challenges under each approach are not comprehensive, but rather identify broad issues.

Approach	Computation Demand	Advantages	Challenges
Delta method from a single reference model	Low	Normal approximation approach makes it easy to calculate statistical uncertainty in quantities of interest	Implicitly assumes a single model describes uncertainty well
Likelihood profiles	Low-Med.	Characterizes within model statistical uncertainty without making normal approximation assumption	Time consumptive when performed over many parameters and models; non-normal distributions can make interpretation difficult
Select number of one-change sensitivity analysis	Low	Quick processing time; ability to explore one or two sources of uncertainty on point estimates	Selecting the set of sensitivity runs; ignores interactions among different factors
Full uncertainty grid	Med.	Moderate processing time but manageable; ability to explore a wide range of model uncertainty	Selecting the set of models to populate the grid; often times ignores interactions among different factors
Uncertainty grid with models weighted by plausibility	High	Moderate processing time but manageable; theoretically more realistic given that some models are likely to be more plausible than others	Selecting appropriate weights in addition to selecting the set of models to populate the grid; often times ignores interactions among different factors
Monte Carlo Markov Chain (MCMC) on a single reference case model	High	Incorporates interactions among uncertain parameters; produces distributions of estimated and derived parameters	Computationally intensive and time consumptive; mostly impractical for WCPFC MFCL-driven assessment models at this time

Table 3. General characteristics of some common approaches to incorporating uncertainty (process, estimation, and model errors) into projections that evaluate prospective fisheries management decisions. The advantages and challenges under each approach are not comprehensive, but rather identify broad issues.

Approach	Computation Demand	Advantages	Challenges
Deterministic from a single reference model	Low	Quick processing time and straight forward interpretation; inferences from a single model	Implicitly assumes that a single model describes uncertainty well; uncertainty in point estimates only
Deterministic on a subset of models from the uncertainty grid	Low-Med.	Incorporates some model (structural) error; analyses not computationally prohibited	selecting a representative subset of grid models
Deterministic across the uncertainty grid	Med.	Incorporates a more full range of model (structural) error; analyses usually not computationally prohibited	selecting the full set of grid models
Stochastic from a single reference model using key sources of uncertainty	Med.	Quick processing time; inferences from a single model; incorporates key process errors	Assuming that a single model describes uncertainty well; selecting a manageable set of key sources of uncertainty
Stochastic using a small subset of grid models chosen by expert opinion	Med.	Incorporates process errors for a subset of grid models; moderately time consumptive but manageable	Selecting a representative subset of grid model for a particular analysis; selecting the key sources of uncertainty;
Stochastic using a small subset of grid models using methods for experimental design (such as the FFD[1] method)	Med.	Significantly reduces the number of runs needed to capture the distribution of model uncertainty; easy to implement once developed/tested	Needs further development and testing for WCPFC applications; feasibility in terms of computational time may be analysis dependent
Stochastic across the uncertainty grid	High	Incorporates combined model and process errors; allows for a full range of interaction between these two sources of error	Computationally intensive and time consumptive; may be impractical depending on the size of the grid and for particular analyses
Stochastic from a single reference model using Monte Carlo Markov Chain (MCMC)	High	Produces distributions of estimated and derived parameters; samples easily drawn for use as stochastic inputs	Computationally intensive and time consumptive; impractical for WCPFC MFCL-driven assessment models at this time

[1] Fractional factorial design (Montgomery 1991; Hoyle et al. 2008)

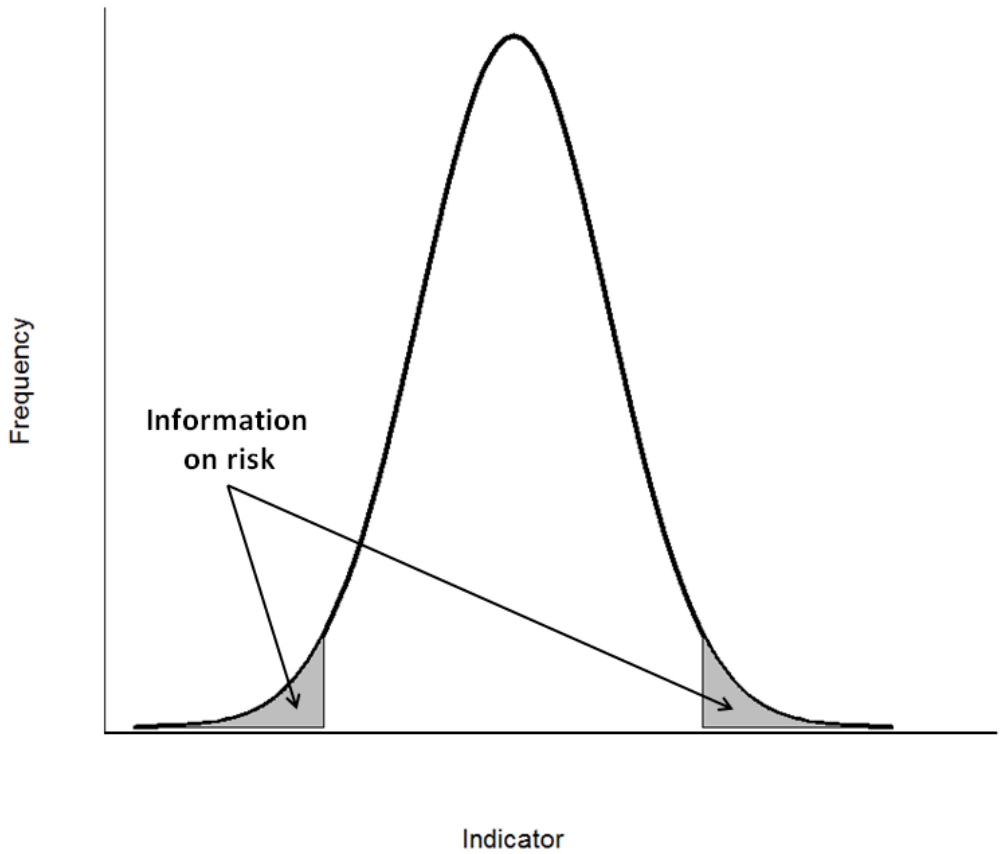

Figure 1. A standard normal distribution that could hypothetically represent the range of uncertainty for a particular management indicator across a set assessment models. In most cases, the underlying distribution is unknown and must be assumed. The distribution tails (shaded) are the most informative when considering risk.

Appendix 1 – Expert opinion poll

A representative from each delegation will be asked to fill out the two tables below during SC9 (to the best of their ability) to come up with assessment model weights for bigeye tuna as an example of how an 'expert opinion' process for incorporating model uncertainty into management decision-making might function. These plausibility weights will be averaged across participating representatives and results for stock status will be collated across models according to their weight and presented back to SC9.

First, assign weights for each of the six grid axes developed for the 2011 bigeye tuna assessment model (WCPFC-SC7-2011/SA-WP-02) according to how important you feel each axis is to the interpretation of assessment results for management (higher weight = more importance). Weights should be assigned between the values of 0 and 1, with the sum across all six axes equal to 1.

Grid axis (sensitivity)	Axis weight
1. Purse seine catches and size composition	
2. Longline CPUE	
3. Steepness	
4. Tagging data	
5. Longline size data	
6. Natural mortality	
Sum total	**1.0**

Second, assign weights to the model runs associated with each grid axes. In this section, you will be weighting the different runs for each axis independently of other axes, such that your weights will indicate how important you feel the assumption made in a particular model run is to the interpretation of assessment results for management. For example, you may give a relatively high weight to models that include PTTP tagging data (4A) if you thought it was a very important aspect to the stock assessment (thus 4B would have a low weight). Weights should be assigned between the values of 0 and 1, with the sum across model runs within each grid axis equal to 1.

Model weights must sum to 1!

Axis	Model run	Model Weight
1. Purse seine catches and size composition		
	A. Spill sample corrected grab samples	
	B. Grab sample only (SBEST)	
	Sum total	**1.0**
2. Longline CPUE		
	A. Operational indices, temporal weighting of standardized effort	
	B. Exclude all CPUE prior to 1975	
	C. Aggregate indices, regions 1-6	
	Sum total	**1.0**
3. Steepness		
	A. Fixed = 0.8	
	B. Fixed = 0.65	
	C. Fixed = 0.95	
	Sum total	**1.0**
4. Tagging data		
	A. Included PTTP	
	B. Exclude PTTP	
	Sum total	**1.0**
5. Longline size data		
	A. Full weight	
	B. Down-weighted	
	Sum total	**1.0**
6. Natural mortality		
	A. Base	
	B. Increased for juveniles	
	Sum total	**1.0**