SCIENTIFIC COMMITTEE
FOURTEENTH REGULAR SESSION

Busan, Republic of Korea
8-16 August 2018

Selecting and Conditioning the Operating Models for WCPO Skipjack

WCPFC-SC14-2018/ MI-WP-03

Scott R.[1], F. Scott F[1], G. M. Pilling[1], J. Hampton[1] and N. Davies[2]

[1] Oceanic Fisheries Programme, The Pacific Community

[2] Te Takina Ltd

Contents

Executive Summary

The MSE evaluation framework is constructed from two main components, an operating model (OM) and a management procedure (MP). In this paper we are specifically concerned with the process of developing and parameterising the OMs that represent the behaviour and dynamics of the fish populations and the fishing fleets that exploit them, a process termed 'conditioning'.

For the MSE framework, a range of OMs should be identified, each one representing a specific plausible hypothesis on stock biology (e.g. natural mortality, movement) or fishery dynamics (e.g. effort creep). The aim is to ensure that the OMs cover all important sources of uncertainty, against which the performance of a MP should be evaluated

We outline the important sources of uncertainty that should be considered when conditioning OMs. With a focus on skipjack tuna and the tropical purse seine fishery, a candidate group of OMs is proposed which is split into reference sets (most plausible and consequential scenarios to be tested) and robustness sets (still plausible but considered as 'what if' scenarios). The following uncertainty grid with a reference set comprising 72 scenarios (see table below and section 3.2 for full details) reflect key uncertainties that may influence future management performance.

This paper presents the first round of conditioning the OMs and should be periodically reviewed and updated to ensure that the range of OMs used in the analysis remains appropriate.

Axis	**Levels**		**Options**		
	Reference	Robustness	**0**	**1**	**2**
Process Error					
Recruitment Variability	**2**		**1982-2014**	**2005-2014**	
Observation Error					
Catch and effort	**2**		**20%**	**30%**	
Size composition	1		all models (see section 2.2.4)		
Tag recaptures	**1**	2	**status quo**	low	none
Model Error					
Steepness ‡	**3**		**0.8**	**0.65**	**0.95**
Mixing period (qtr) ‡	**2**		**1**	**2**	
Tag overdispersion ‡	**3**		**high**	**medium**	**low**
Movement	**1**	1	**estimated**	El Nino/La Nina	
DD catchability (k) ‡	**1**	1	**0**	-0.5	
Implementation Error					
Effort creep	**1**	1	**0%**	2% cont.	

Table: Skipjack OM uncertainty grid. Scenarios shown in bold are proposed for the reference set. ‡ denotes those scenarios for which a dedicated fit of MULTIFAN-CL is required.

Our considerations of OM scenarios that should be carried forward to the MSE evaluations can be broadly categorised into the following five groups:

1. **Changes to the model that apply to all scenarios:** Natural mortality is modelled as a spline function (with 4 nodes) and the weighting of length composition data is fixed at 20 (the value used for the 2016 reference case assessment).

2. **Uncertainties from the stock assessment grid that have been retained for the MSE grid:** Values for steepness, tag mixing rate and effort creep are carried over from previous studies without change. In addition the currently defined year ranges for recruitment variability are retained.

3. **New settings for the MSE grid:** Additional sources of uncertainty include observation error in catch, effort and size composition data; density dependent catchability and revised values for tag overdispersion.

4. **Lower priority elements:** Variability in the age at maturity, the regional distribution of recruitment and autocorrelation in recruitment have little impact on model quantities and are not included in either the reference or robustness sets.

5. **Areas for future work:** Uncertainty in tag reporting rates and regional variation in growth and maturity have been highlighted for further investigation. Similarly procedures for applying alternative movement hypotheses and for including additional process error in projections through the effort deviations should be investigated.

We invite WCPFC-SC to consider the following questions:

- have all important sources of uncertainty been considered?
- do the ranges of parameter values adequately reflect uncertainty in the dynamics of the resource?
- are the scenarios properly allocated between reference and robustness sets?

1 Introduction

The harvest strategy approach provides a framework for taking the best available information about a stock or fishery and applying an evidence and risk-based approach to setting harvest levels. A key element of the harvest strategy approach are harvest control rules (HCRs) that determine how much fishing can take place given the status of the target stock. It is recommended that candidate HCRs are tested, prior to implementation in the fishery, to determine the extent to which they achieve defined management objectives. The most widely adopted process for testing HCRs (and management procedures) prior to implementation is based on simulation analysis and is termed management strategy evaluation (MSE).

1.1 MSE Framework

A more comprehensive overview of the MSE framework is provided in Scott et al. (2018). In brief, the MSE evaluation framework (Figure 1) is constructed from two main components, an operating model (OM) and a management procedure (MP). The OM is a mathematical representation of the "true" system, rather like a stock assessment. It simulates the real world by attempting to capture all existing knowledge and data processes for the exploited populations and associated fisheries. Our knowledge of the dynamics of populations and fisheries is often incomplete. The OM should therefore allow for the evaluation of the consequences for management by testing different hypotheses about those dynamics. In this respect a suite of different OMs may be identified, each one representing an alternative hypothesis. Very often the OMs will include a greater level of complexity than that used for the stock assessment so that all important sources of uncertainty might be appropriately included in the evaluation process.

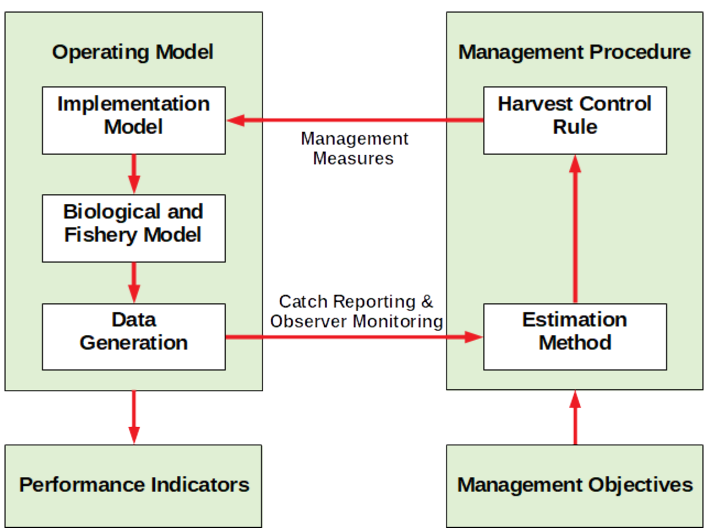

Figure 1: Conceptual diagram of the MSE framework.

The process for testing and agreeing a management procedure can be summarised in the following

7 steps (Punt and Donovan, 2007).

1. Identify and prioritise the management objectives in qualitative terms.
2. Translate the qualitative objectives into quantitative performance indicators.
3. Develop and parameterise operating models representing plausible dynamics of the resource or system (a process termed conditioning).
4. Identify candidate management procedures.
5. Simulate the future performance of each management procedure by applying them to the resources or systems represented by the operating models.
6. Summarise the performance of each management procedure in terms of the performance indicators (from step 2).
7. Identify the management procedure most likely to meet the management objectives.

In this paper we are specifically concerned with step 3, the process of developing and parameterising the operating models that represent the behaviour and dynamics of the fish populations and the fishing fleets that exploit them. This is a particularly important process in the development of the MSE framework and will require ongoing work to periodically re-evaluate the selection of OMs to use in the evaluations. Periodic review of the OMs will ensure that any new data or updated information can be incorporated into the analyses and provides an opportunity to review the bounds and limits used to define exceptional circumstances (see Section 4). This paper therefore presents the first round of conditioning the OM and may be considered a 'living document' that can be periodically reviewed to ensure that the range of OMs used in the analysis remains appropriate.

1.2 Conditioning the Operating Model

Conditioning an operating model (Rademeyer et al., 2007) involves fitting the model to data in much the same way that a stock assessment model is fit to the available catch, size composition and tag recapture data. Because we do not have perfect knowledge of the dynamics of the resource, uncertainties will always be present in any assessment of the status or productivity of the fishery. Our objective is to identify a suite of models that adequately characterises the range of uncertainty so that we can find the MP that performs best and is robust to that uncertainty.

1.2.1 Accounting for Uncertainty

Stock assessments conducted by the Pacific Community (SPC) have typically presented a range of model configurations, termed the uncertainty grid, that explore the sensitivity of the assessment results to alternative assumptions about model settings for which the data are often uninformative

(Table 1). Examples include the steepness of the stock recruitment relationship; the period required for tagged fish to become fully mixed with the general population or the relative importance (weighting) that different data sources (size composition, tag recaptures, CPUE) should be given in the assessment model. In some cases a preferred set of assumptions (diagnostic case) has been identified that might be considered to give the most likely representation of stock status. More recently, however, all of the assessment models from the uncertainty grid have been presented and management advice has been based on the resulting range of potential stock status values across the whole uncertainty grid. A weighting may be applied to each of the models in cases where some might be considered more or less likely to represent the true population.

Axis	**Code**	**Levels**	**Options**			**Importance**
			0	**1**	**2**	
Steepness	A	3	0.8	0.65	0.95	Moderate/High
Length comp. wtg	B	3	20	10	50	Moderate
Mixing period (qtr)	C	2	1	2		Moderate
Tag overdispersion	D	3	Default	Estimated	Fixed	High

Table 1: Skipjack 2016 stock assessment uncertainty grid (McKechnie et al., 2016b).

The stock assessment uncertainty grid is a useful starting point for considering the range of uncertainty that should be included in the suite of OMs for the MSE analyses. However, the assessment uncertainty grid is concerned primarily with those factors that impact the historical trajectory of the stock, as estimated by the stock assessment. When projecting assessment results forwards in time, as performed by the MSE simulations, it may be necessary to consider a different set of sensitivities in order to adequately capture the most important sources of uncertainty.

A number of previous studies have categorised the types of uncertainty based on their different sources (Rosenberg and Restrepo, 1994; Francis and Shotton, 1997; Kell et al., 2007). Punt et al. (2014) recommend that, at minimum, an MSE should consider (i) process uncertainty, in particular through variation about the stock-recruitment relationship; (ii) parameter uncertainty, with reference to stock productivity and the overall size of the resource; (iii) observation error, in the data used when applying the management strategy. They note, however, that the choice of uncertainty to include in any MSE will be case specific.

1.2.2 Reference and Robustness Sets

It is considered best practice to divide the suite of OMs into a reference set and a robustness set (Rademeyer et al., 2007). The reference set is considered to reflect the most plausible hypotheses and forms the primary basis for identifying the 'best' management strategy.

The robustness set comprises hypotheses that are considered less likely but still plausible. The scenarios considered in the robustness set will differ among fisheries but may include, for example, alternative hypotheses on the quality of historical data (eg. errors in catch statistics); the availability of future data (eg. future tagging programs); environmental effects that may impact on the spatial distribution or carrying capacity of the resource, or changes in the structure of the fishery such as gear changes that may affect selectivity or catchability.

The performance measures will be calculated from the reference set whilst the robustness set will be used to give a secondary indication of the performance of the management strategy.

The full list of uncertainties to be considered can be extensive. It is therefore important to identify the most important and influential sources of uncertainty in order to limit the final number of scenarios to a practical number that can be realistically considered, given the usual restrictions on processing power and evaluation time, whilst ensuring the key sources of uncertainty are sufficiently accounted for. A full factorial combination of model scenarios, in which each scenario is matched with every other scenario, rapidly becomes overly large and unmanageable. It will therefore be necessary to trim the grid of scenario combinations to exclude combinations considered less plausible, for example combinations of fast growth rates (implying a very productive stock) with low natural mortality (often associated with slow growing less productive stocks). In addition scenarios from the robustness set might be considered as one off sensitivity runs rather than a defined grid of scenario combinations.

In Section 2 we present an overview of the important issues to consider when conditioning an OM along with a description of previous work undertaken to investigate these issues and how they relate to skipjack in the WCPO. In Section 3 we draw on this general overview, and a number of additional analyses, to provide specific recommendations on scenarios to be included in the suite of OMs and how those scenarios should be divided between the reference and robustness set. It may not be possible to examine all these scenarios due to technical limitations of the modelling framework or data deficiencies, and in Section 4 we consider what additional work will be required in order to address these issues.

2 Designing the MSE

Following a review of candidate tools for OM development, MULTIFAN-CL has been identified as the most appropriate tool to use as the OM within the MSE simulation framework for WCPO stocks and fisheries (Scott et al., 2016, 2017a) and a number of technical developments have recently been implemented to provide the software with the necessary functionality (Davies et al., 2017, 2018). The most recent stock assessment of skipjack in the WCPO was conducted using MULTIFAN-CL and this assessment and the data upon which it relies will form the basis of the OM. There are however a number of additional sources of data and alternative modelling approaches that might also be considered when choosing the suite of OMs and plausible states of nature that comprise the reference and robustness sets. These alternative approaches are further discussed in Sections

2.2 and 3.1.6.

2.1 Data Sources

The data used for the stock assessment include fishery specific catch and effort, size composition data and a substantial quantity of tag release and recapture information. The details of these data have been documented in previous papers to the WCPFC-SC (McKechnie et al., 2016b,c; McKechnie, 2016; Peatman et al., 2016) and are not described further here.

An important consideration is the likely availability and quality of future data. The assessment of skipjack in the WCPO relies heavily on standardised indices of abundance derived from pole and line fisheries and on tag release and recapture information. Pole and line fisheries in many regions of the skipjack assessment have declined substantially in recent years and now represent only a small fraction of the total skipjack catch. It is unclear whether these fisheries will be able to continue to provide reliable indices of abundance for skipjack. Tagging cruises continue on a periodic basis but are subject both to future funding and to the competing data demands of other assessments. The design of tagging programmes in recent years has been modified to try to tag greater numbers of bigeye and yellowfin tuna.

2.2 Plausible States of Nature

Ideally, the range of uncertainties considered in an MSE should be sufficiently broad that new information collected after the management strategy is implemented should reduce rather than increase the range (Punt and Donovan, 2007). In practice, however, it is often not practical to consider all sources of uncertainty and decisions will need to be made on which alternative scenarios are the most important and consequential and therefore need to be included. In the following sections we address a number of potential sources of uncertainty. We consider their likely importance and review the available information on them with specific regard to skipjack in the WCPO.

2.2.1 Stock Structure

Skipjack distribution is influenced by their environment. Skipjack inhabit the oceanic surface layers of the warm waters of the tropics and warm water current systems of sub-tropical regions. They have limited tolerance to low levels of dissolved oxygen and low temperatures due to their high metabolic rate and large red muscle mass. They have limited ability to thermoregulate and, apart from occasional brief dives to deeper water, are restricted to the surface waters having an optimal temperature range between 19°C to 26°C (Boyce et al., 2008).

Skipjack in the WCPO are, for the purposes of stock assessment, assumed to comprise a single panmictic stock having a distribution that extends from about 40°N to 40°S that roughly corresponds to the 20°C surface isotherm. However, the sub-structure of the skipjack population throughout the Pacific Ocean is not fully understood. It is considered unlikely that all adults have an equal opportunity of breeding with each other and that the population is therefore not truly panmictic. Whilst it is considered that skipjack follow some form of population structuring across the Pacific it remains unclear whether this structuring is based on stable geographic boundaries or on a continuous cline where the chance of any two fish breeding is inversely proportional to the distance between them (Wild and Hampton, 1994).

The current stock assessment of skipjack in the WCPO assumes a spatial structure comprised of 5 regions and movement between these regions is assumed to occur instantaneously at the beginning of each time period. An alternative spatial structure (Ochi et al., 2016) was investigated and compared with the reference case assessment model. Model estimates from the revised spatial structure were broadly similar with the reference case assessment although overall levels of depletion showed moderate differences (McKechnie et al., 2016a).

SEAPODYM (Lehodey and Senina, 2009) is a model developed for investigating the spatio-temporal dynamics of fish populations under the influence of both fishing and the environment in which movement is constrained by physical and biological environmental observations. Similarly, Ikamoana (Scutt Phillips et al., 2018) is an individual based model that replicates the movement processes of Eulerian models, such as SEAPODYM and MULTIFAN-CL, using similar habitat-dependent movement processes. An updated version of SEAPODYM 3.0 has been applied to skipjack tuna in the Pacific Ocean (Senina et al., 2016) to estimate stock structure and potential movement rates based on a 2° and 1 month spatial and temporal resolution. The potential application of these model estimates is further discussed in section 3.1.6.

MULTIFAN-CL estimates recruitment from a single stock and recruitment relationship (SRR) and distributes these recruits amongst the assessment regions according to an estimated relative proportion of recruitment which does not change systematically through time (i.e. some random deviation may apply). In this way the overall recruitment level is determined from the overall abundance of skipjack. This assumption may be reasonable when considering the historical trajectory of the stock but may not be appropriate when conducting projections under varying rates of exploitation. For cases where one or more regions may be subject to localised depletion, skipjack abundance in those regions could be estimated at artificially high levels because the model assumes that the proportional distribution of recruitment does not change systematically. A potential solution, that would be more consistent with the assumption that skipjack are not truly panmictic, would be to estimate region specific SRRs. A common steepness parameter might apply to each SRR, to reduce the number of estimated parameters, but R0 (the level of recruitment at virgin biomass) would be estimated separately for each region. This feature is not currently available in MULTIFAN-CL but could be implemented for use in future evaluations.

2.2.2 Parameter Non-stationarity

Parameter non-stationarity refers to the variation in parameter estimates over time, space, or some other determining factor. It has implications for many of the assumptions that are necessarily made when assessing the status of a resource including the biological characteristics of the population, rates of movement and factors that may vary with stock abundance.

Growth and Maturity

Biological characteristics of fish populations can vary over time and space. Previous studies on skipjack suggest that growth rates may vary between areas of the Pacific (Wild and Hampton, 1994; Maunder, 2001) and that growth may be faster in regions close to the equator (Leroy, 2000). Growth and maturity are intrinsically linked and faster growing individuals are likely to reach sexual maturity at an earlier age (Charnov and Berrigan, 1991). Changes in growth rates may be linked to environmental conditions or to density dependent effects and in some cases temporal changes in growth and maturity have been linked to fishing, where individuals grow faster and mature at a younger age under exploitation; a process known as fishery induced evolution (Law, 2000; Heino et al., 2015) although this is often only the case for populations subject to very high rates of exploitation.

In the current skipjack assessment, growth, natural mortality and maturity are all assumed to be invariant over time and space. Skipjack mean length at age is assumed to follow the same von Bertallanfy growth curve estimated from previous skipjack assessments and maturity to follow a knife-edge pattern with all fish two quarters or less being immature and all fish older being fully mature (McKechnie et al., 2016b). The assessment estimates an age specific natural mortality that remains fixed through time and across regions.

Following SC12, analyses were conducted to address the concerns of some CCMs regarding some of the model assumptions of the 2016 skipjack stock assessment including, amongst others, the form of the natural mortality function, alternative growth functions and the assessment's sensitivity to age at maturity (McKechnie et al., 2016a). The analysis investigated a range of approaches for modelling natural mortality and concluded that a spline function (with four nodes) provided a more parsimonious model and more realistic estimates of natural mortality at the youngest age. Across the range of approaches, model quantities and resulting reference points were very similar to the reference case assessment. Based on these findings we recommend that the four knot spline function be used to model natural mortality in the OM.

Numerous studies have been conducted to determine growth rates in skipjack (see Wild and Hampton (1994)) and a selection are presented in Figure 2. These studies have typically been based on samples taken from a distinct region of the Pacific Ocean and have used a variety of capture methods and estimation techniques. Few, if any, can be considered representative of skipjack across the

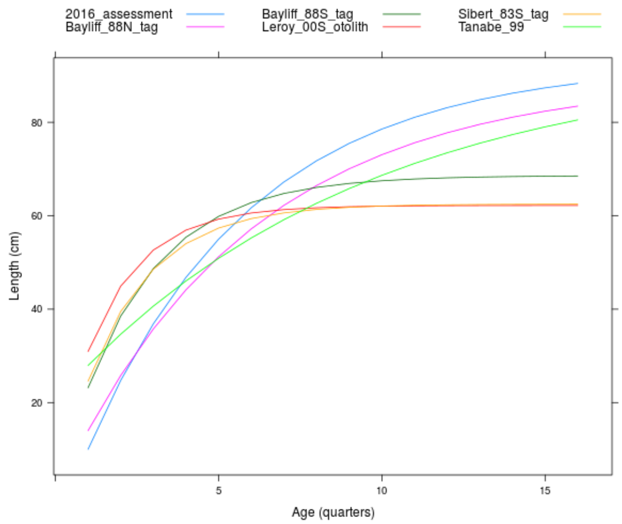

Figure 2: Sample growth models

whole assessment area. Whilst there is clearly uncertainty regarding the growth assumptions to be used in the OM it is difficult to know what an appropriate range of uncertainty might be.

The sensitivity of the 2016 stock assessment to alternative age at maturity assumptions has also been investigated (McKechnie et al., 2016a). Increasing the age at which skipjack were assumed to reach full maturity from three quarters to four quarters resulted in reduced estimates of adult biomass throughout the time series, but had relatively little impact on key reference point values (F/F_{MSY}, $SB/SB_{F=0}$) and in terms of total biomass ($B/B_{F=0}$) the two models were almost identical. From these results, uncertainty in maturity can be considered a lower priority for inclusion in the OM scenarios.

Maturity and growth are intrinsically linked and it might be expected that any temporal and spatial variation in growth rates will be accompanied by corresponding changes in maturity. There is, however, limited opportunity to incorporate temporal or spatial variability in biological characteristics in the current implementation of MULTIFAN-CL. The biological characteristics of a stock are important and potentially influential inputs to the OM and there remains a requirement for studies that can determine these values and their likely variability across the geographic range of the stock assessment.

Density Dependent Processes

Other model parameters, such as fishery catchability, may also vary over time, space, or in response to some density dependent process. Typically a linear relationship is assumed between CPUE and abundance and in the current skipjack assessment, catchability is estimated individually for each of the fisheries represented in the assessment. For those fisheries for which CPUE was not standardised, catchability was allowed to vary progressively over time, but was assumed to be independent of abundance. The potential for non-linear dynamics in the CPUE abundance relationship has been proposed by several studies (Harley et al., 2001; Maunder et al., 2006) and investigated with specific reference to WCPO skipjack (Scott et al., 2015b). The true relationship is difficult to determine, however, the erroneous assumption of a linear relationship with stock abundance, when a non-linear relationship actually exists, can lead to an underestimation of the extent to which a stock is being depleted through fishing and from a fisheries management perspective it can lead to bias in the estimated magnitude of effort change required to manage a stock to biomass targets. Given the uncertainty in the relationship between skipjack abundance and purse seine CPUE in particular, this will be considered further in Section 3.

Movement

The stock assessment of skipjack in the WCPO is based on a spatial structure comprising 5 regions. MULTIFAN-CL distributes the population amongst these regions by estimating the relative proportion of recruitment in each region and age-specific, seasonal movement, both of which are assumed to remain constant across all years. However, movement among the assessment regions is likely to be strongly influenced by climatic conditions that are known to vary over time, most notably through the El Niño Southern Oscillation (ENSO). The occurrence of La Niña and El Niño ENSO events affects the east-west displacement of the warm pool which, in turn, affects the spatial distribution of skipjack in the WCPO and the fishing fleets that target them (Lehodey et al., 1997).

Parameter non-stationarity is a broad subject area with implications for numerous sources of uncertainty. Aspects of parameter non-stationarity are considered further in section 3 with specific reference to recruitment (Section 3.1.1); growth (Section 3.1.2); movement (Section 3.1.6) and density dependent catchability (Section 3.1.7).

2.2.3 Process Error

Process error arises through natural variability in the biotic and abiotic processes that impact on population dynamics. For example the extent of natural mortality experienced by larval fish can vary substantially depending on several factors such as the availability of suitable prey items or the abundance of predators. Small changes in the timing of key events, such as plankton blooms or predator migration rates, can lead to substantial changes in conditions for survival and consequently in the numbers of fish that recruit to the fishery (Cushing, 1990). These conditions are difficult, if not impossible, to predict but can have a significant impact on the state of the resource. The simplifying assumption is typically made that future variability will be consistent with the variability observed over some historical period. A common approach when conducting projections is to run multiple iterations with future variability imposed through re-sampling of the historical deviates from a fitted stock and recruitment relationship. The period over which this historical sample is taken and the extent of variability that it covers is therefore an important consideration when running stochastic projections (Pilling et al., 2016). Uncertainty in future recruitment levels and the nature of variability (e.g. autocorrelation) is a key consideration when testing the robustness of HCRs and specific recruitment scenarios are further considered in Section 3.

Additional mechanisms for incorporating process error into the simulations include the effort deviations, the extent to which model estimates of effort deviate from observed effort. These represent 'noise' in the relationship between fishing effort and fishing mortality and are further discussed with respect to both observation error and process error in Section 2.2.4 and 3.1.3.

2.2.4 Observation Error

Observation error is the difference between a measured value of a quantity and its true value. It includes natural errors that occur in any data collection procedure as well as systematic errors (affecting all measurements) that can arise from, for example, the miscalibration of instruments. Observation error is a key source of uncertainty and a particularly important consideration with respect to the input data to the MP. The key input data include fishery specific estimates of catch, effort, size composition and tag recaptures, all of which will be subject to observation error to a greater or lesser extent.

When generating tag recapture data for the future the number and spatial distribution of tag releases must be specified by the user. Observation error is introduced into the tag recapture data based upon the OM estimation of the multinomial probability of recapture given the release samples. In this sense the probability of recapture of tagged fish is determined from the internal calculations of MULTIFAN-CL and cannot be specified by the user. The user must, however, specify the quantity of tags to be released, the regions from which those releases will be made and the fishery selectivity from which the length distribution of the releases will be generated. A summary of historical tag releases is shown in Table 2.

Region	**Historical**							
	SSAP		**RTTP**		**PTTP**		**JTTP**	
	1977-1980		1989-1992		2006-2014		1998-2014	
1	1	(83)	0		0		57	(305)
2	7	(1125)	9	(866)	11	(1335)	34	(103)
3	9	(4554)	5	(2049)	8	(825)	4	(49)
4	2	(2699)	7	(1933)	3	(6880)	18	(66)
5	3	(3924)	10	(2541)	16	(5641)	0	

Table 2: Tag release summary for WCPO skipjack: Number of tag release events by region and tagging program. The average number of fish tagged and released per release event is shown in brackets.

2.2.5 Model Error

Model error arises from the possibility that the model is deficient in some way. For example, a simple linear model might be considered deficient if it were used to describe a clearly non-linear process. Model error is associated with the structural design of the model such as the functional form of the stock and recruitment relationship or the particular growth model employed (Richards, Von Bertalanffy, etc.).

Parameter error is the possibility that the parameters used to define the model are incorrect, given that the model form is correct. Parameter error occurs because there is only a limited amount of data from which to estimate the parameters and because the parameters themselves may evolve through time (non-stationarity). For the purpose of this analysis we consider model error and parameter error together as a single category and several important elements of model error have already been discussed in section 2.2.2.

With specific reference to the skipjack assessment, potential sources of model error include the assumed steepness of the stock and recruitment relationship (Figure 3) and the period of time to allow for adequate mixing of tagged fish with the wider population. Both of these error sources are already included in the stock assessment uncertainty grid of models. The existing range of values (0.65, 0.8 and 0.95 for steepness; 1 and 2 quarters for mixing) are considered to sufficiently span the range of uncertainty and it is proposed that both of these axes of uncertainty be maintained in the reference set of OMs.

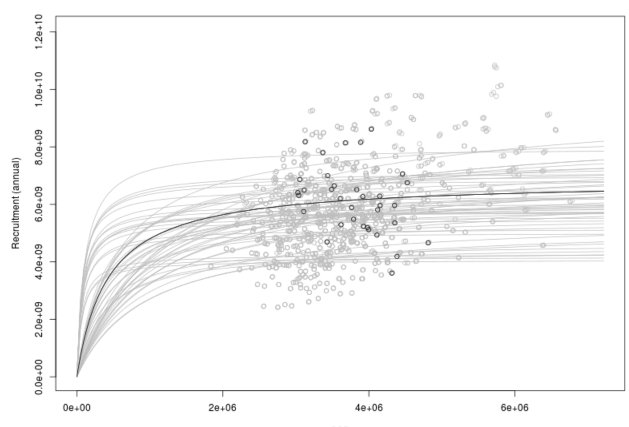

Figure 3: Fitted SRR and estimated annual recruitment (1982:2015) for all 54 models in the uncertainty grid. Reference case estimates shown in black.

The current skipjack assessment applies a single, annual stock and recruitment relationship for all regions of the assessment. Recruits are then proportionally allocated across quarters and regions according to a time invariant distribution. A comparison with the results of a model in which a quarterly SRR was fitted (Figure 4) shows the distribution of recruits across regions and quarters is very similar. The seasonal and regional distribution of recruits appears to be well estimated by the assessment and there is no requirement for additional uncertainty in this regard.

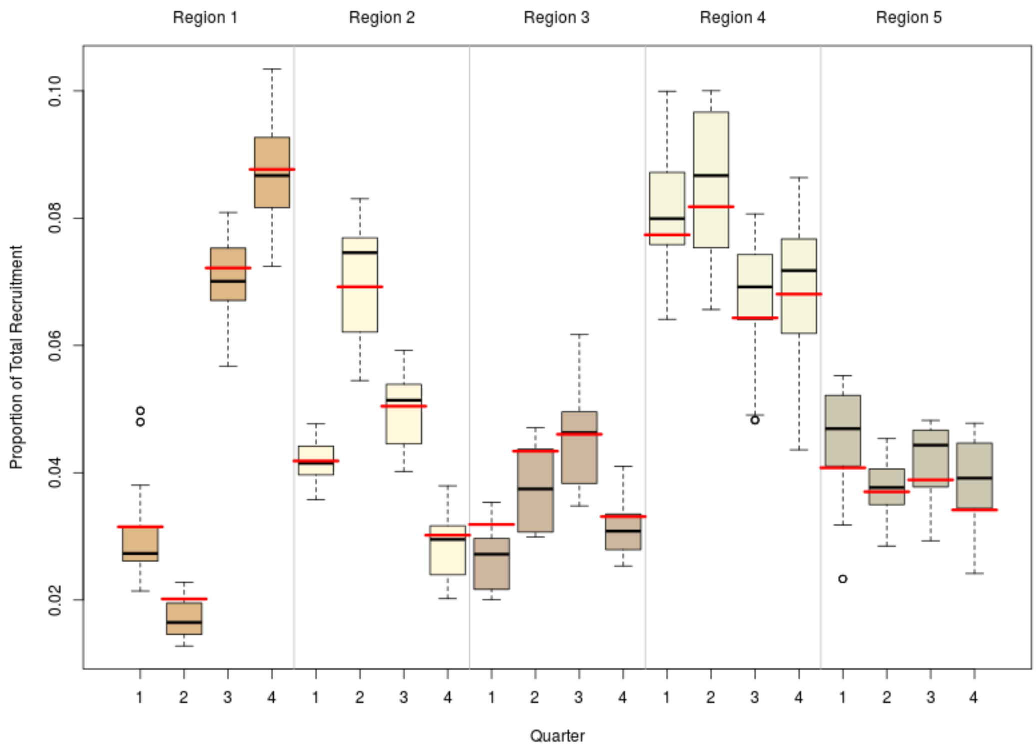

Figure 4: Quarterly regional recruitment distribution across the grid, averaged over years (1982:2015). Red lines show the quarterly regional recruitment distribution from the quarterly SRR sensitivity model.

Other important sources of parameter error relate specifically to the use of tag release and recapture data within the assessment model. They include factors such as the tag overdispersion settings and the reporting rate of recaptured tags. Overdispersion occurs when the observed variance exceeds the variance assumed by the model and is an important consideration with respect to tag data that often show high variance in recapture rates. A range of overdispersion values were included in the uncertainty grid of the 2016 stock assessment and in Section 3 we further consider the extent of overdispersion that may apply to WCPO skipjack tag recapture data.

The reliability with which recaptured tags are reported as having been recaptured is a key consideration when analysing tag recapture data and is an area of active research in support of WCPO stock assessments (Berger et al., 2014; Peatman et al., 2016). Information from tag seeding experiments, undertaken by on-board observers, is used to determine prior estimates of the likelihood that a recaptured tag will be detected by the crew and reported as having been recaptured. These prior estimates of reporting rate are further modified by MULTIFAN-CL during the assessment fitting process and in many cases the resulting rate is greater than the rate estimated from the tag seeding experiments. The tag reporting rates as estimated from MULTIFAN-CL across the assessment uncertainty grid (Figure 5) show a range of values for most reporting groups (although a small number are consistently estimated at the upper bound) indicating that some variability in the tag reporting rate parameter is imposed on the model through the range of settings that are already considered in the uncertainty grid. A single trial run in which the reporting rates were fixed at the prior values gave estimates of very high biomass in some regions that could not be considered plausible. We do not, therefore, propose any additional uncertainty be included in the uncertainty grid, with specific regard to tag reporting rates, beyond that which is already imposed through the variation of other parameters. We do, however, recommend that this component of uncertainty be re-considered in future iterations of the OM conditioning process.

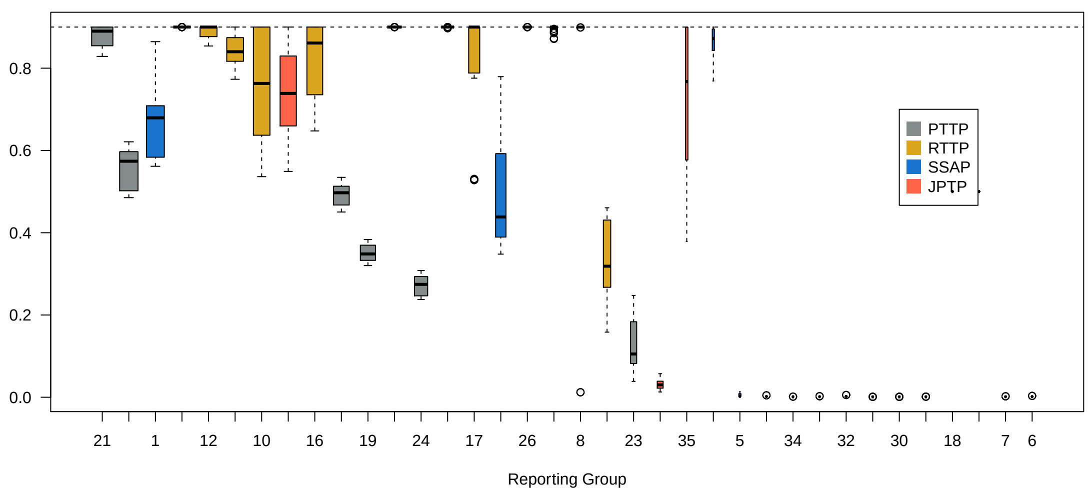

Figure 5: Estimated tag reporting rates across the grid of model runs.

2.2.6 Implementation Error

Implementation error results when the management actions specified by the HCR are not followed precisely by the fishery. Some level of implementation error is almost always inevitable especially in cases where a single species HCR is applied to fisheries that may opportunistically target a range of species. At a finer scale, the HCR may specify the overall catch or effort to be applied but not the distribution of that catch or effort amongst the various fisheries. Purse seiners increasingly account for the majority of skipjack catches and previous analyses have shown that the biomass target reference point for skipjack can be achieved over a range of different FAD and free-school effort distributions (OFP, 2014). However, the results may be impacted in cases where the true effort distribution deviates significantly from that assumed in the evaluations. The simplest assumption is that the outcome of the HCR will be precisely implemented and that each fishery will be subject to a proportional change in catch or effort as a result. Deviations from this assumption will be subject to the results of ongoing negotiations through WCPFC.

3 Conditioning the OM

In section 2 we have outlined various sources of uncertainty and discussed their importance with regard to the assessment and management of skipjack in the WCPO. In this section we describe additional runs and analyses that have been conducted to further investigate the key sources of uncertainty and our attempts to define appropriate bounds for them. In section 3.2 we propose the components of the MSE uncertainty grid and how the different scenarios might be grouped into reference and robustness sets.

3.1 Model Runs and Analyses

Several elements of the assessment uncertainty grid have been retained for inclusion in the MSE uncertainty grid (steepness, tag overdispersion, tag mixing), and a number of additional elements are proposed. In some cases additional analyses have been conducted to further investigate the most appropriate parameter values, whilst in others the same parameter values have been carried over. We describe specific analyses to investigate appropriate parameter values for growth, tag overdispersion and the extent of observation error in catch and effort data and the reasons for selection or rejection of models from the MSE uncertainty grid. In cases where it has not been possible to conduct the necessary analyses we identify potential avenues for future investigations.

3.1.1 Recruitment Variability

The latest skipjack assessment indicates a progressive trend for increasing recruitment over time (Figure 6). It is not known if the recently observed higher levels of recruitment will persist into the future or if future recruitment will revert to the longer term, lower levels. Previous analyses for skipjack have typically considered two levels of recruitment variability. The first based on long term recruitment over the period 1982 to 2014 and the second on recent recruitment for the period 2005 to 2014 (see Figure 6). We propose two recruitment scenarios be represented in the reference set of operating models, one based on longer term recruitment over the period 1982 to 2014 and the second based on the higher recent recruitments for the period 2005 to 2014.

Due to the low number of observations for the most recent cohorts, recruitment estimates in the terminal time periods of the model can be poorly estimated. Retrospective analyses conducted as part of the 2016 skipjack stock assessment (McKechnie et al., 2016b) indicated that recruitment in the last two quarters of the assessment were poorly estimated and were therefore fixed at geometric mean of the time series. We therefore omit the recruitment estimates for the terminal assessment year (2015) when re-sampling recruitment for stochastic projections.

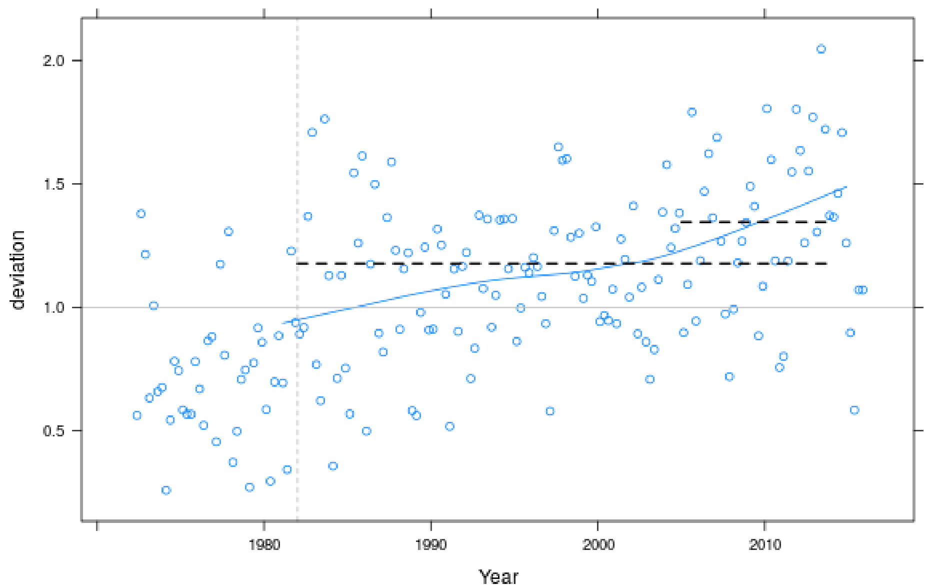

Figure 6: Deviations from the fitted Beverton Holt stock and recruitment model (reference case) with loess smoother fitted over the period 1982:2014. Future projections draw random deviates, at the annual level, from either the long-term (1982:2014) or short-term (2005:2014) periods. Dotted lines show mean recruitment over the long and short-term periods.

Autocorrelation measures the similarity of observations as a function of the time lag between them. The extent of autocorrelation in recruitment is an important consideration both for the historical fit of an assessment and when generating future recruitment series for projections. Estimates from MULTIFAN-CL fits for WCPO skipjack indicate that first order autocorrelation in annual recruitment (ρ) is low (median $\rho = 0.04$ across all grid scenarios) and recruitment and biomass estimates across model combinations with and without autocorrelation in recruitment gave almost identical results. We therefore consider autocorrelation in recruitment to have a low priority for inclusion in the OM scenarios.

3.1.2 Growth

A small number of sensitivity trials were conducted to investigate the consequences of alternative growth assumptions on model estimates. Relatively simple trials in which either length at the youngest and oldest ages was fixed and the rate of growth between these two points varied or growth rate was fixed but length at the oldest age varied (Figure 7) showed that increasing the growth rate or the size at the oldest age resulted in lower estimates of adult biomass and an $SB/SB_{F=0}$ around 4% to 6% lower than the reference case assessment, whilst reducing the growth rate or the size at the oldest age gave estimates of adult biomass that were slightly higher and an $SB/SB_{F=0}$ around 4% to 6% higher than those of the reference case assessment. It should be noted

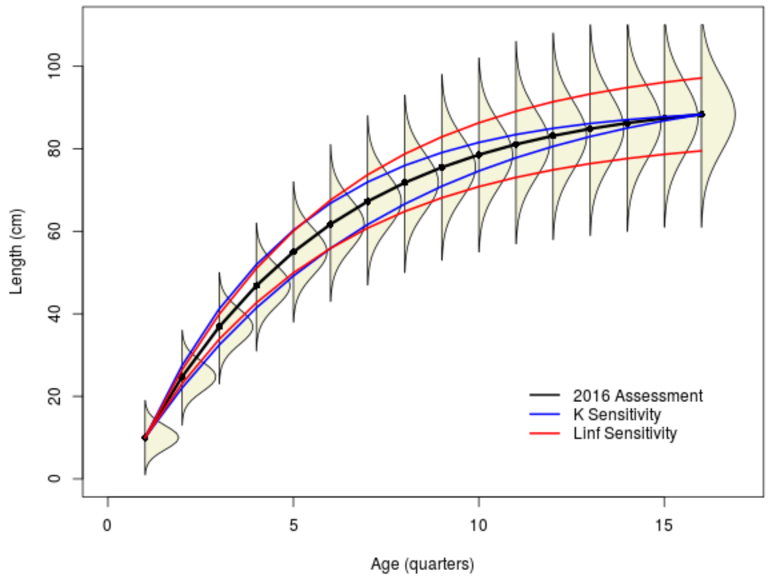

Figure 7: Growth model sensitivity

that no corresponding changes were made to maturity at age and that the range of growth models investigated was relatively small in comparison to the range of models determined from previous studies.

The alternative growth parameters assumed for these trials was based on relatively arbitrary choices. The results provide an indication of the sensitivity of the model estimates to alternative growth assumptions but do not provide any guidance on appropriate distributions for those parameters. In Appendix A.2 we describe an approach using the variance covariance matrix as calculated from the hessian of the fitted model to derive appropriate parameter distributions for generating variability in growth. However, we consider this work to be preliminary and does not yet provide a sufficiently reliable basis for the specification of variability in growth in the OM.

3.1.3 Observation Error in Catch, Effort and Size Composition Data

Catch and effort data

Methods for generating future data with observation error in MULTIFAN-CL are described in Davies et al. (2017). Observation error for catch and effort data are derived assuming a log-normal error distribution. A single, user defined, coefficient of variation is specified for catch, or effort, which applies across all fisheries and all periods in the model. Previous investigations of the inclusion of observation error in catch and effort data (Scott et al., 2017b) have been based on an arbitrary CV of 0.1. To try to determine a more appropriate value we applied an approach based on a linear regression of catch against effort. We fit a simple linear model to the observed catch and effort data for purse seine fisheries for each fishery and quarter (Figure 8) from which we calculated the standard error of the residuals about the fitted model and divided this by the mean of the predicted values to obtain a coefficient of variation for each fishery and quarter (Table 3). The calculated CVs vary markedly between fisheries and to a lesser extent across quarters.

$$\sigma_{est} = \sqrt{\frac{\sum(Y - Y')^2}{N-1}} \tag{1}$$

$$CV = \frac{\sigma_{est}}{\sum Y'/N} \tag{2}$$

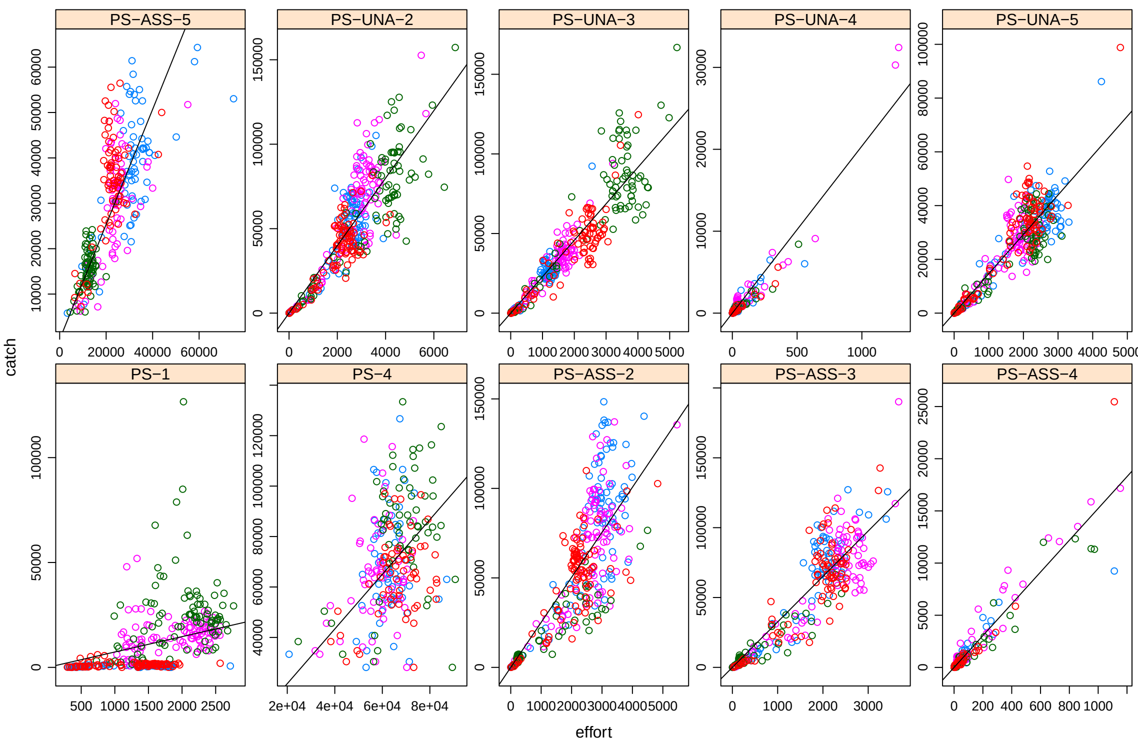

Figure 8: Linear regression of catch on effort for purse seine fisheries. (Note that in this figure a single regression is fit through all 4 quarters whereas in the analysis above separate regressions were fit to each fishery and quarter).

Using this approach, we estimated the CVs for the historical period (1972 to 2015) and for the results of future stochastic projections with increasing values for the input CV (0.1, 0.2 and 0.3). We found the estimated CV for the future period was always greater than the input CV and that the range of estimated values was always smaller than that of the historical period (Figure 9). An input CV of 20% resulted in a median output CV of 0.36 which is close to that calculated for the same fisheries for the historical period (0.38). Based on these results CVs of 0.2 and 0.3 may be considered appropriate values for the generation of catch and effort observation error to include in the reference set of operating models for skipjack.

We note however, that there are well established procedures for monitoring and reporting catch and effort in most of the fisheries targeting skipjack and that it is unlikely that error in the observation of these quantities will be as high as 30% to 40%. The CV estimated through this approach represents a combination of observation error in the reported values of catch or effort and also process error in the underlying relationship between catch and effort. This process error is captured by the MULTIFAN-CL fitting process in the estimation of effort deviations. Stochastic projections that incorporate re-sampling from the effort deviations (in the same way that recruitment deviations are currently re-sampled) would provide a more comprehensive simulation approach and would separate the sources of error more appropriately into their constituent components. This feature for MULTIFAN-CL projections is currently under development but once implemented would be the preferred approach.

Fishery	**Quarter**			
	1	**2**	**3**	**4**
PS-1	0.601	0.556	0.755	1.001
PS-ASS-2	0.334	0.335	0.452	0.345
PS-4	0.308	0.306	0.279	0.238
PS-ASS-4	1.323	0.418	0.646	0.747
PS-UNA-2	0.283	0.274	0.297	0.288
PS-ASS-5	0.285	0.310	0.254	0.288
PS-UNA-5	0.239	0.311	0.304	0.263
PS-UNA-4	0.997	0.503	0.706	0.699
PS-ASS-3	0.252	0.308	0.498	0.306
PS-UNA-3	0.281	0.219	0.274	0.285

Table 3: Fishery and quarter specific CVs for catch and effort for purse seine fisheries.

Size composition data

Observation error in size composition data is generated from a multinomial distribution (Davies et al., 2017). The variability in the generated data can be controlled through the effective sample size (ESS) which is specified by the user for each individual size composition. As such, there is considerable flexibility for the specification of size composition observation error between fisheries and across time. As a first choice of appropriate values to use we have taken the estimated ESS as determined from a MULTIFAN-CL model fit using the self-scaling multinomial fitting option (Davies et al., 2018)which estimates fishery specific ESS (Table 4). No alternative ESS values are proposed at present.

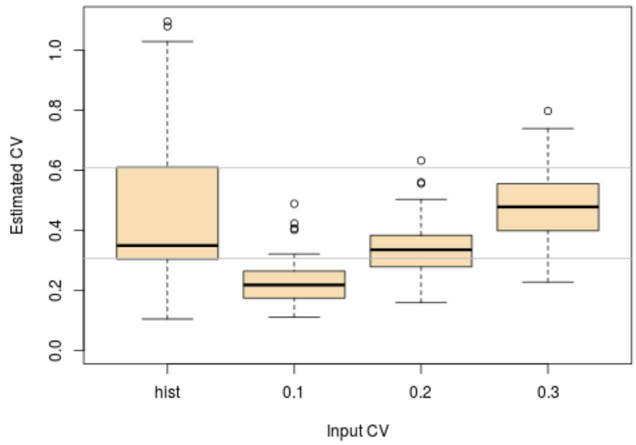

Figure 9: Boxplot of the estimated CV in the catch effort relationship (ie. the CV as determined from eqns 1 and 2) for purse seine fisheries for varying input CVs. Horizontal grey lines indicate the approximate inter-quartile range of the historically observed CVs.

3.1.4 Overdispersion

The 2016 assessment investigated the sensitivity of the assessment model to the value of the overdispersion parameter in the negative binomial likelihood for the tagging data. Alternative assumptions about overdispersion produced the most significant deviations from the estimates of the reference case model of all of the sensitivity analyses considered (McKechnie et al., 2016b).

Options for both fixing the overdispersion parameter at some pre-defined value and for estimating it within the model have been previously explored. Estimated values for overdispersion varied amongst model runs between values of around 2.5 and 4.5. However, the option to estimate the extent of overdispersion in the distribution of tag recaptures is a relatively new feature in MULTIFAN-CL and the extent to which the data support the estimation of this parameter requires further investigation. Alternative modelling approaches based on the Tagest model (Sibert et al., 2012) using the same skipjack tagging data suggest that direct estimation of the extent of overdispersion is difficult.

The default setting of MULTIFAN-CL approximates the Poisson distribution which is generally considered to significantly under-estimate the variance in tag recaptures. It is therefore desirable that some level of overdispersion be specified, however, the degree of overdispersion that should be applied is less clear. Based on the findings of the sensitivity analyses conducted for the last skipjack assessment we propose that 3 levels of overdispersion (τ = 2.5, 4, 8) be carried forward to the reference set of OMs (see Figure 10), but recommend that further work is conducted in this area to determine the most appropriate values of overdispersion.

Fishery	Region	Catch/Effort	Scaler	ESS
PL	1	1	1	55
PS	1	2	1	29
LL	1	1	1	7
PL	2	1	1	14
PS-ASS	2	2	1	44
PS-UNA	2	2	1	39
LL	2	1	1	12
PL	5	1	1	22
PS-ASS	5	2	1	38
PS-UNA	5	2	1	35
LL	5	1	1	10
PL	3	1	1	14
PS-ASS	3	2	1	47
PS-UNA	3	2	1	40
LL	3	1	1	49
Dom-PH	4	1	1	26
Dom-ID	4	1	1	21
PS	4	2	1	42
PL	4	1	1	16
PS-ASS	4	2	1	12
PS-UNA	4	2	1	10
Dom-VN	4	1	1	43
LL	4	1	1	6

Table 4: Fishery specific effective sample sizes as estimated from MULTIFAN-CL using the self-scaling multinomial fitting option.

3.1.5 Length Composition Weighting

One of the difficulties in fitting integrated models is the need to assign relative weightings to the different data components. It is increasingly recognised that some data components should be actively down-weighted in some situations to allow other more reliable, or less numerous, data to inform the model (Francis, 2011). The uncertainty grid of the 2016 skipjack assessment included three levels of weighting for the length composition data, although in practice these weightings had relatively little influence on model quantities, in particular $SB/SB_{F=0}$ (McKechnie et al., 2016b).

For skipjack tuna the tagging data provide most of the information on scaling overall biomass and on parameters such as natural mortality and movement. The assumed level of overdispersion for the tagging data therefore has greater influence on model quantities as it impacts the most influential data components in the assessment. We therefore propose that alternative weighting of the length composition data be dropped from the MSE uncertainty grid and that a single weighting value of 20 be applied across all scenarios.

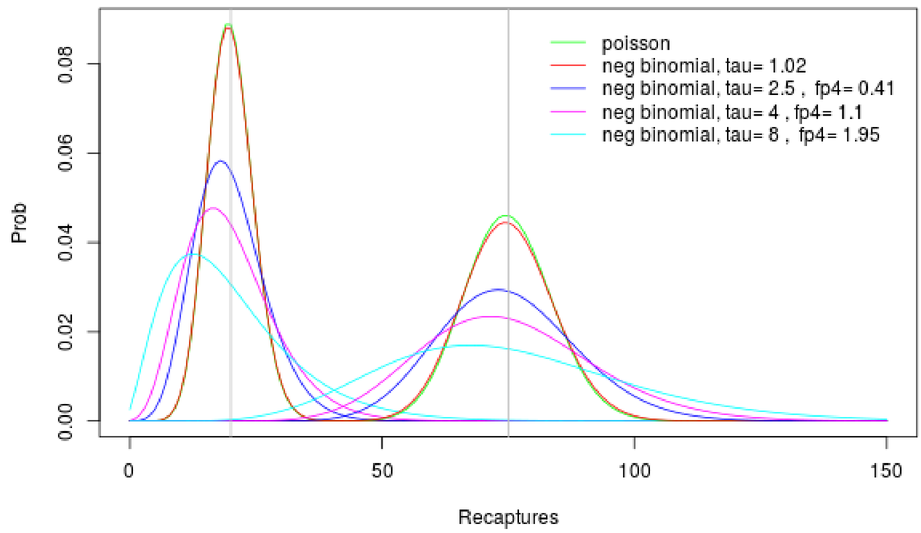

Figure 10: Negative binomial probability densities with varying levels of overdispersion (τ) with mean values of 20 and 75 corresponding approximately to the range of the number of recaptures for a given release group and fishery combination. At low overdispersion ($\tau \to 1$) the negative binomial approximates the Poisson distribution. Probability densities corresponding to overdispersion values of 2.5, 4 and 8 are also shown.

3.1.6 Movement

A potential approach for including temporal variability in movement rates would be to estimate separate movement matrices for specific ENSO conditions (El Niño, La Niña, neutral) and to randomly replace the movement matrix used in the MULTIFAN-CL projections to simulate the appropriate oscillatory dynamics in climate. ENSO specific estimates of movement for the skipjack regions can be determined from SEAPODYM and have already been estimated using a similar model that employs an individual-based modelling approach, *Ikamoana*, (Scutt Phillips et al., 2018).

The estimation of ENSO specific movement rates for skipjack using SEAPODYM has yet to be completed and no trials of this approach have, to date, been undertaken. The approach may not be as straight forward as simply plugging in an alternative movement matrix. Both SEAPODYM and *Ikamoana* estimate greater connectivity between assessment regions than MULTIFAN-CL i.e. higher levels of movement in particular between regions 2 and 5. This may occur because movement rates in MULTIFAN-CL do not contain inter-annual variation, consequently, contrasting movements associated with El Niño and La Niña conditions may lead to averaging of the estimates over the assessment time period. An initial comparison between the movement estimates of MULTIFAN-CL and *Ikamoana* suggest that the MULTIFAN-CL movement rates are close to those estimated by *Ikamoana* for neutral ENSO conditions.

3.1.7 Density Dependent Catchability

Previous analyses to investigate the potential impact of hyperstability on skipjack (Scott et al., 2015b) concluded that, underestimating the extent of hyperstability in CPUE will likely lead to an overestimation of the necessary change in effort that is required to achieve a desired change in stock biomass. However, the degree of hyperstability that might be operating in the WCPO purse seine fisheries is difficult to estimate, in particular due to potential confounding with changes in purse seine efficiency that occur through time.

Hyperstability in CPUE can affect the performance of a HCR. We propose that some consideration of the potential for hyperstability in CPUE be retained in the MSE evaluations and the functionality currently exists in MULTIFAN-CL to incorporate this. However, given the high degree of uncertainty of the extent of hyperstability that may be operating in WCPO purse seine fisheries we recommend that it be considered as part of the robustness set rather than the core reference set of model scenarios.

3.2 Operating Model Selection

The key sources of uncertainty to be considered in the OMs and their proposed parameter values are summarised in Table 5. Where appropriate the scenarios have been provisionally allocated to reference and robustness sets. The scenarios listed in table 5 represent a combination of model fits and projection settings. As such, some of the scenarios, specifically those scenarios listed under model error, will require refitting of MULTIFAN-CL assessments, whereas scenarios relating to the projection settings, such as recruitment variability or the level of observation error, will not.

3.2.1 Reference and Robustness Sets

The selection of OMs identified in Table 5 results in a reference set of 18 models and 72 scenarios, assuming a full factorial design. Additional models identified for the robustness set may be run as one off scenarios thereby reducing the overall number of evaluations that need to be conducted.

We note that scenarios based on changes in the rate of movement between regions are currently included in the uncertainty grid but no trials have yet been conducted to investigate these scenarios. These analyses will be possible once the relevant movement rates have been estimated from SEAPODYM. In the meantime this axis of uncertainty has been provisionally tabled for the reference set.

Axis	**Levels**		**Options**		
	Reference	Robustness	**0**	**1**	**2**
Process Error					
Recruitment variability	**2**		**1982-2014**	**2005-2014**	
Observation Error					
Catch and effort	**2**		**20%**	**30%**	
Size composition	1		all models (see section 2.2.4)		
Tag recaptures	**1**	2	**status quo**	low	none
Model Error					
Steepness ‡	**3**		**0.8**	**0.65**	**0.95**
Mixing period (qtr) ‡	**2**		**1**	**2**	
Tag overdispersion ‡	**3**		**4**	**2.5**	**8**
Movement	**1**	1	**estimated**	El Nino/La Nina	
DD catchability (k) ‡	**1**	1	**0**	-0.5	
Implementation Error					
Effort creep	**1**	1	**0%**	2% cont.	

Table 5: Skipjack OM uncertainty grid. Scenarios shown in bold are proposed for the reference set. ‡denotes those scenarios for which a dedicated fit of MULTIFAN-CL is required.

3.2.2 Model Validation

In this section we consider the validity of the models we have identified for the OM uncertainty grid. Throughout the process of conditioning the OMs, consideration must be taken of how the suite of models corresponds to the real world system, or at least our perception of it. At the simplest level, the models should appear to be realistic representations of the stock and fisheries. There is no simple test to establish the validity of a model and instead we rely on a collection of indicators based on diagnostics of the fit of the model to data and consideration of whether the quantities estimated from it are reasonable.

Axis	parameter	Label	**Settings**		
			0	**1**	**2**
Steepness	h	A	0.8	0.95	0.65
Tag mixing (qtrs)		B	1	2	
Tag overdispersion	τ	C	4	2.5	8
Hyperstability	k	D	0	-0.5	
Recruitment autocorrelation	ρ	E	0	est	

Table 6: Settings and labels for the models considered for inclusion in the reference and robustness sets. Note that recruitment autocorrelation has been included in these investigations but may not be included in the final selection of models.

The maximum gradient is a measure of how well the model has converged to a solution. The lower the maximum gradient the better the model fit is considered to be. Maximum gradients for the initial fits of the OM model grid (Figure 11) are consistently below 0.01, indicating that all of the models have achieved a satisfactory level of convergence. Other model diagnostics include likelihood profiles of key model parameters and retrospective analyses that indicate how sensitive the model is to the addition of new data. Due to time constraints we have not been able to include these diagnostic checks in this paper but will complete these as part of the ongoing work.

Temporal trends in the estimates of adult biomass are relatively consistent across the grid of models. The absolute levels of adult biomass vary between runs but are comparable to that of the 2016 reference case assessment (Figure 12). Model settings for the tag data account for the greatest variability. Higher estimates of biomass are associated with longer tag mixing periods (2 qts) and low values of overdispersion (τ 2.5), whilst the lowest estimates of adult biomass are associated with short tag mixing periods (1 qtr) and high values of overdispersion (τ 8). Different values of steepness for the SRR introduced very little variation into the historical estimates of adult biomass but are retained within the reference set of models as steepness will be an important source of uncertainty when conducting projections.

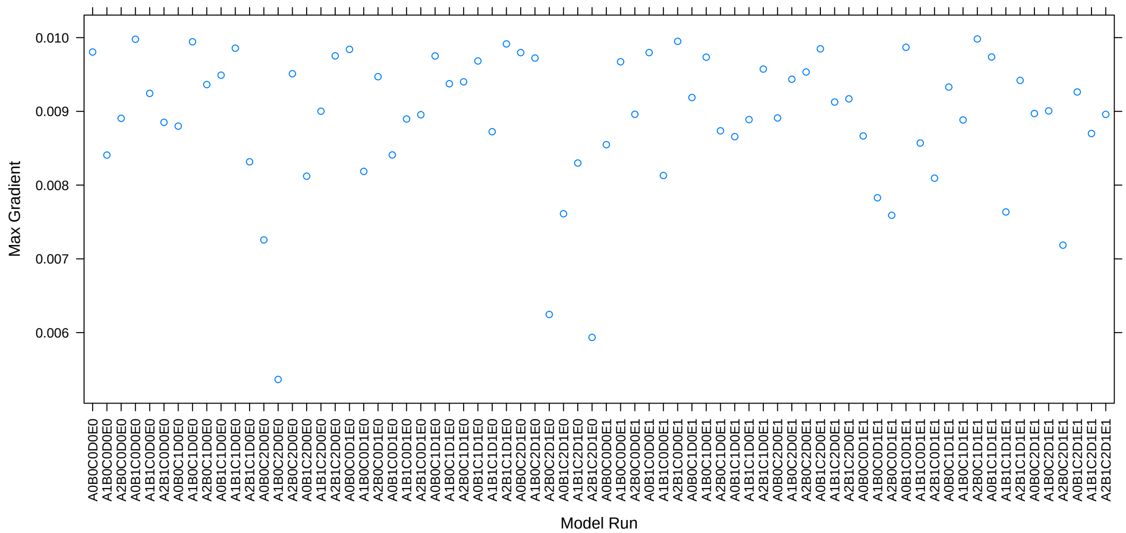

Figure 11: Maximum gradient across the grid of models. (All models were run for an additional 18000 function evaluations after the final fitting phase)

4 Future Work Areas

We note again that this document represents the first round of conditioning the OMs for the stock of skipjack tuna in the WCPO and its associated fisheries. The selection of OMs that constitute the reference and robustness sets will need to be periodically reviewed to allow for new data or updated information to be incorporated and to ensure that all aspects of our uncertainty are adequately accounted for. The next assessment of WCPO skipjack tuna is anticipated to be in 2019 and will provide a good opportunity to reconsider some of the outstanding issues.

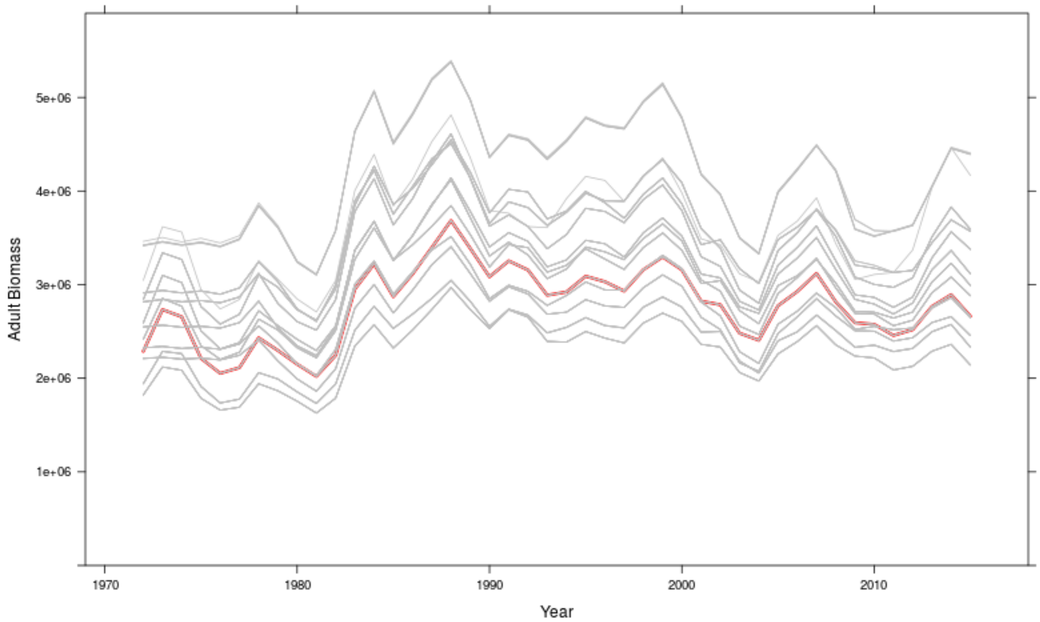

Figure 12: Estimated adult biomass across the grid of OM models. Red line shows the estimated adult biomass from the 2016 reference case skipjack assessment model

Throughout this document we have made a number of recommendations for future work to address some unresolved issues, either in cases where the necessary data are not currently available or the necessary analyses have not yet been undertaken, or because our modelling tools and software require further development.

Estimates of growth rates and maturity that are representative of the skipjack population across the whole assessment area as well as information on how growth varies between the different regions will be required before appropriate bounds on parameter estimates can be determined and, as has been noted in previous studies, additional information on growth, such as length-increment data from tagging studies, should be given high priority for future investigation.

The MULTIFAN-CL software is subject to ongoing development and a number of potential modifications have been identified both with regard to the model fitting process (region specific stock and recruit relationships) and approaches for including additional uncertainty (e.g. effort deviations) in stochastic projections. In addition it is recommended that future analyses further investigate the most appropriate model settings for overdispersion and reporting rates when modelling tag release and recapture data.

4.1 Plausibility Weightings

Numerous analyses for WCPFC fisheries have previously been conducted in which simulations have been run across a grid of models. In some cases the models have been weighted according to their perceived plausibility. Examples are analyses to determine acceptable levels of risk of exceeding the LRP (OFP, 2014) or the evaluation of management options for the tropical tuna CMM bridging measure (OFP, 2017). However, some fora recommend not to apply weightings to the scenarios on the basis that the weightings are almost always subjective and there is often little or no data to inform on the plausibility of one scenario relative to another. Instead it is recommended that each scenario receive equal weighting unless there is a compelling reason to do otherwise.

Alternative approaches to a fixed grid of models are available. The approach adopted by CCSBT for the evaluation of southern bluefin tuna was based on a suite of OMs that were randomly sampled until a specified number of evaluations had been completed. Through this approach all of the OMs were given equal weighting with their relative occurrence in the evaluation depending on the random sample. A sufficiently high number of evaluations will be required in order to ensure that a sufficiently representative sample is drawn from the suite of OMs, although depending on the number of OMs the necessary number of evaluations may be smaller than in the case of a fixed grid.

4.2 Exceptional Circumstances

Throughout the evaluation and testing process it must be borne in mind that the identification of a 'best performing' management procedure will be dependent on the range of assumptions over which it has been tested and the extent to which those assumptions reflect the true underlying dynamics of the stock and fishery. The scenarios we have selected are expected to cover the most likely eventualities, but we cannot assume that they cover all eventualities. It is therefore necessary to identify those situations, termed "exceptional circumstances", in which one might abandon either the HCR or the management procedure all together and take some other form of management action.

In general terms, exceptional circumstances include any event that falls outside the range of assumptions over which the management procedure has been tested, but may also include situations where the trajectory of the stock does not respond as expected to management action. For example if biomass falls below the limit reference point, or catches continually exceed some upper threshold.

We do not identify any exceptional circumstances here, but note that the range of OMs chosen for the evaluation will be a key consideration when the time comes to identify them.

5 Conclusions

Throughout this paper we have outlined the important sources of uncertainty that should be considered when conditioning OMs along with the outcomes of analyses to address key uncertainty issues. In Section 3.2 we make recommendations for the scenarios to include in the reference and robustness sets of OMs and propose a reference set comprising 18 models and 72 scenarios, assuming a factorial design.

Our considerations of OM scenarios that should be carried forward to the MSE evaluations can be broadly categorised into the following five groups:

1. **Changes to the model that apply to all scenarios:** Natural mortality is modelled as a spline function (with 4 nodes) and the weighting of length composition data is fixed at 20 (the value used for the 2016 reference case assessment).

2. **Uncertainties from the stock assessment grid that have been retained for the MSE grid:** Values for steepness, tag mixing rate and effort creep are carried over from previous studies without change. In addition the currently defined year ranges for recruitment variability are retained.

3. **New settings for the MSE grid:** Additional sources of uncertainty include observation error in catch, effort and size composition data; density dependent catchability and revised values for tag overdispersion.

4. **Lower priority elements:** Variability in the age at maturity, the regional distribution of recruitment and autocorrelation in recruitment have little impact on model quantities and are not included in either the reference or robustness sets.

5. **Areas for future work:** Uncertainty in tag reporting rates and regional variation in growth and maturity have been highlighted for further investigation. Similarly procedures for applying alternative movement hypotheses and for including additional process error in projections through the effort deviations should be investigated.

We note that a number of important areas of uncertainty have been identified that have yet to be fully incorporated into either the reference or robustness sets. It is important that future work be undertaken to address these unresolved issues. In this respect, this paper presents the first round of conditioning the OM and should be periodically reviewed to ensure that the range of OMs used in the analysis remains appropriate.

Acknowledgments

We gratefully acknowledge funding for this work from the New Zealand Ministry of Foreign Affairs and Trade (MFAT) funded project "Pacific Tuna Management Strategy Evaluation" and from the European Union through their funding support for WCPFC Project 75 "Technical Support for the Development of Harvest Strategies". In addition we thank the Center for High Throughput Computing (CHTC UW-Madison) for generously providing access to their computing resources.

References

Berger, A. M., McKechnie, S., Abascal, F., Kumasi, B., Usu, T., and Nichol, S. J. (2014). Analysis of tagging data for the 2014 tropical tuna assessments: data quality rules, tagger effects, and reporting rates. WCPFC-SC10-2014/SA-IP-06, Majuro, Republic of the Marshall Islands, 6–14 August 2014.

Boyce, D., Tittensor, D. P., and Worm, B. (2008). Effects of temperature on global patterns of tuna and billfish richness. *Marine Ecology Progress Series*, 355:267–276.

Charnov, E. and Berrigan, D. (1991). Evolution of life history parameters in animals with indeterminate growth. *Evolutionary Ecology*, 5(1):63–68.

Cushing, D. (1990). Primary production and year-class strength in fish populations: an update of the match/mismatch hypothesis. *Advances in Marine Biology*, 26:249–293.

Davies, N., Fournier, D. A., Takeuchi, Y., Bouye, F., and Hampton, J. (2017). Developments in the MULTIFAN-CL software 2016-2017. WCPFC-SC13-2017/SA-IP-05, Rarotonga, Cook Islands, 9–17 August 2017.

Davies, N., Fournier, D. A., Takeuchi, Y., Bouye, F., and Hampton, J. (2018). Developments in the MULTIFAN-CL software 2017-2018. WCPFC-SC14-2018/SA-IP-02, Busan, South Korea, 9–17 August 2018.

Francis, R. I. C. C. (2011). Data weighting in statistical fisheries stock assessment models. *Canadian Journal of Fisheries and Aquatic Sciences*, 68:1124–1138.

Francis, R. I. C. C. and Shotton, R. (1997). "Risk" in fisheries management: a review. *Canadian Journal of Fisheries and Aquatic Science*, 54:1699–1715.

Harley, S., Myers, R., and Dunn, A. (2001). Is catch per unit effort proportional to abundance? *Canadian Journal of Fisheries and Aquatic Sciences*, 58:1760–1772.

Heino, M., Pauli, B., and Dieckmann, U. (2015). Fisheries induced evolution. *Annual Review of Ecology, Evolution and Systematics*, 46:461–480.

Kell, L., Mosqueira, I., Grosjean, P., Fromentin, J., Garcia, D., Hillary, R., Jardim, E., Mardle, S., Pastoors, M., Poos, J., Scott, F., and Scott, R. D. (2007). FLR: an open source framework for the evaluation and development of management strategies. *ICES Journal of Marine Science*, 64:640–646.

Law, R. (2000). Fishing, selection and phenotypic evolution. *ICES Journal of Marine Science*, 57(3):659–668.

Lehodey, P., Bertignac, M., Hampton, J., Lewis, A., and Picaut, J. (1997). El Nino Southern Oscillation and tuna in the western Pacific. *Nature*, 389:715–718.

Lehodey, P. and Senina, I. (2009). A user manual for SESEAPODYM version 2.0: application with data assimilation (draft version). WCPFC-SC5-2009/EB-IP-13., Port Villa, Vanuatu. 10-21 August 2009.

Leroy, B. (2000). Preliminary results on skipjack (*Katsuwonus pelamis*) growth. SKJ-1, 13th Meeting of the Standing Committee on Tuna and Billfish. Noumea, New Caledonia.

Maunder (2001). Growth of skipjack tuna (katsuwonus pelamis) in the easter pacific ocean, as eestimate from tagging data. Technical Report Bulletin of the Inter-American Tropical Tuna Commision, vol. 22.

Maunder, M., Hinton, M., Bigelow, K., and Langley, A. (2006). Developing indices of abundance using habitat data in a statistical framework. *Bulletin of Marine Science*, 79(3):545–559.

McKechnie, S. Hampton, J., Pilling, G., and Davies, N. (2016a). Aditional Analyses to ssupport the 2016 stock assessment of skipjack tuna in the western and central Pacific Ocean. WCPFC-SC12-2016/SA-WP-04, Bali, Indonesia, 3–11 August 2016.

McKechnie, S. (2016). Summary of fisheries structures for the 2016 stock assessment of skipjack tuna in the western and central Pacific Ocean. WCPFC-SC12-2016/SA-IP-06, Bali, Indonesia, 3–11 August 2016.

McKechnie, S., Hampton, J., Pilling, G. M., and Davies, N. (2016b). Stock assessment of skipjack tuna in the western and central Pacific Ocean. WCPFC-SC12-2016/SA-WP-04, Bali, Indonesia, 3–11 August 2016.

McKechnie, S., Ochi, D., Kiyofuji, H., Peatman, T., and Caillot, S. (2016c). Construction of tagging data input files for the 2016 skipjack tuna stock assessment in the western and central Pacific Ocean. WCPFC-SC12-2016/SA-IP-05, Bali, Indonesia, 3–11 August 2016.

Ochi, D., Ijima, H., and Kiyofuji, H. (2016). A re-consideration of growth pattern of skipjack on the western central Pacific. WCPFC-SC12-2016/SA-IP-08, Bali, Indonesia, 3–11 August 2016.

OFP (2014). Multi-species implications of reference points: what might a target reference point of 50 MOW3-WP/05, SPC-OFP, Faleata Sports Complex, Apia, Samoa, 28th Nov 2014.

OFP (2017). An evevaluation of the management options for purse seine and longline fisheries defined by the tt cmm intersessional meeting. WCPFC14-2017-10 rev.1, Manilla, Philippines, 3-7 December 2017.

Peatman, T., Caillot, S., Leroy, B., McKechnie, S., Roupsard, F., Sanchez, C., Nicol, S., and Smith, N. (2016). Analysis of tag seeding data and reporting rates. WCPFC-SC12-2016/SA-IP-13, Bali, Indonesia, 3–11 August 2016.

Pilling, G., Scott, R., Davies, N., and Hampton, J. (2016). Approaches used to undertake management projections of WCPO tuna stocks based upon MULTIFAN-CL stock assessments. WCPFC-SC12-2016/MI-IP-04, Bali, Indonesia, 3–11 August 2016.

Punt, A. E., Butterworth, D., de Moor, C., De Oliveira, J., and Haddon, M. (2014). Management strategy evaluation: best practices. *Fish and Fisheries*, (DOI:10.111/faf12104).

Punt, A. E. and Donovan, G. (2007). Developing management procedures that are robust to uncertainty: lessons from the international whaling commission. *ICES Journal of Marine Science*, 64:603–612.

Rademeyer, R., Plaganyi, E., and Butterworth, D. (2007). Tips and tricks in designing management procedures. *ICES Journal of Marine Science*, 64:618–625.

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

Scott, F., Jardim, E., Millar, C., and Cervino, S. (2015a). An applied framework for incorporating multiple sources of uncertainty in fisheries stock assessments. *PLoS ONE*, 11(5).

Scott, F., Scott, R. D., Davies, N., Pilling, G., and Hampton, S. (2018). Technical developments in the MSE modelling framework. WCPFC-SC14-2018/MI-IP-02, Busan, South Korea, 5–13 August 2018.

Scott, R., Pilling, G. M., Hampton, J., Reid, C., and Davies, N. (2016). Report of the expert consultation workshop on management strategy evaluation. WCPFC-SC12-2016/MI-WP-05, Bali, Indonesia, 3–11 August 2016.

Scott, R. D., Davies, N., Pilling, G., and J., H. (2017a). Developments in the MSE modelling framework. WCPFC-SC13-2017/MI-WP-04, Rarotonga, Cook Islands, 9–17 August 2017.

Scott, R. D., Davies, N., Pilling, G., and J., H. (2017b). Generating pseudo-data in MULTIFAN-CL. WCPFC-SC13-2017/MI-IP-02, Rarotonga, Cook Islands, 9–17 August 2017.

Scott, R. D., Tidd, A., Davies, N., Pilling, G., and Harley, S. (2015b). Implementation of alternative CPUE/abundance dynamics for purse seine fisheries within MULTIFAN-CL with application to effort-based projections for skipjack tuna. WCPFC-SC11-2015/SA-IP-02, Pohnpei, Federated States of Micronesia, 5–13 August 2015.

Scutt Phillips, J., Sen Gupta, A., Senina, I., Sebille, E., Lange, M., Lehodey, P., J., H., and Nichol, S. (2018). An individual-based model of skipjack tuna Katsuwonus pelamis movement in the tropical pacific ocean. *Progress in Oceanography*, 164:63–74.

Senina, I., Lehodey, P., Calmettes, B., Nichol, S. J., Caillot, S., J., H., and P., W. (2016). Predicting skipjack tuna dynamics and effects of climate change using seapodym with fishing and tagging data. WCPFC-SC12-2016/EB WP01.

Sibert, J., Senina, I., Lehodey, P., and Hampton, J. (2012). Shifting from marine reserves to maritime zoning for conservation of Pacific bigeye tuna (*Thunnus obesus*). *Proceedings of the National Academy of Sciences*, 109(44):18221–18225.

Wild, A. and Hampton, J. (1994). A review of the biology and fisheries for skipjack tuna, *Katsuwonus pelamis*, in the Pacific Ocean. FAO Fisheries Technical Paper (FAO) 336(2):1–51.

A Appendix

A.1 Overdispersion

Dave's notes on overdispersion

The probability mass function (pmf) for the negative binomial can be written as

$$Pr(X = k) = \left(\frac{a}{a+\mu}\right)^a \frac{\Gamma(a+k)}{k!\Gamma(a)} \left(\frac{\mu}{a+\mu}\right)^k \quad for \quad k = 0, 1, 2, \ldots \tag{3}$$

where μ is the mean and the variance is $\mu(1 + \mu/a)$. Note that for the Poisson with mean μ the variance is μ so that $(1+\mu/a)$ is the amount by which the variance of the negative binomial exceeds that for the Poisson. So $(1 + \mu/a)$ measures the overdispersion. This can be parameterised by τ, so $\tau = (1 + \mu/a)$. Since the pmf is often given in terms of a it is convenient to solve for a in terms of τ namely

$$a = \frac{\mu}{\tau - 1} \tag{4}$$

Note that τ must always be greater than 1 which follows from the fact that the negative binomial always has a variance greater than the mean.

The code in tagfit.cpp ensures that τ is greater than 1.

```
dvariable tau=1.0+exp(fish_pars(4,gp_fish32(ig)));
a=rtc/(1.e-20+(tau-1.0));
```

So if $\tau \to 1$, the distribution approaches a Poisson. For $\tau = 4$ the variance is 4 times the mean, that is 4 times greater than it would be for the Poisson.

Parameterisation of overdispersion in MFCL

If fish flag 43 = 1 and parest flag 305 = 1 then MFCL estimates overdispersion using the new parameterisation and stores the estimated values in fish pars (i,4).

If fish flag 43 = 0 and parest flag 305 = 0 then it uses the default value for overdispersion (about 1,02) which is very close to Poisson and fish pars(i,4) will be set to 0.

If fish flag 43 = 0 and parest flag 305 = 1 it will use the values given in fish pars(i,4) but does not estimate the overdispersion.

A.2 Growth

Skipjack mean length at age is fixed within the current assessment and is assumed to follow the same von Bertalanffy growth curve estimated from previous assessments. During the preparation of the 2016 assessment a one-off analysis was conducted in which growth was allowed to be estimated by the model. The estimates of growth from this model differ very slightly from the fixed values that were assumed in the final assessment runs (Figure 13).

To investigate the potential variability in growth parameters we generated 1000 sets of log transformed von Bertalanffy parameter values from a multivariate normal distribution with variance covariance matrix determined from the inverse of the hessian matrix. The multivariate normal assumption provided parameter distributions with appropriate median values but with extremely long right hand tails that contained unreasonable values. To reduce the occurrence of excessively wide distributions we determined variability in the growth parameters using a t-copula with triangle marginal distributions (Scott et al., 2015a). Copulas allow for flexibility in multivariate distribution models, allowing more robust distributions than the common multivariate gaussian. Triangle distributions make few assumptions about the parameters and simplify the definition of bounds. We used the median and 10th and 90th percentiles to specify the marginal distribution of each parameter.

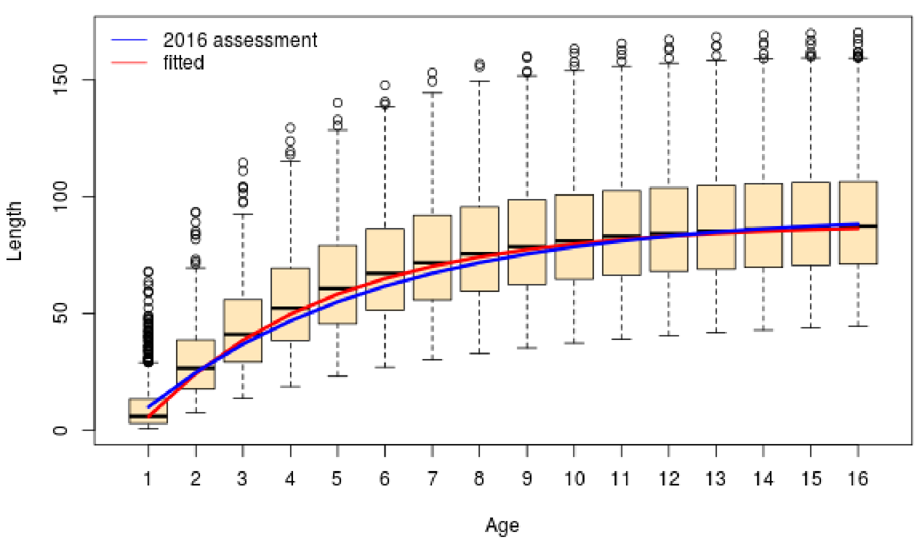

Figure 13: Von Bertalanffy growth models as used for the 2016 stock assessment (blue) and from a comparative model in which growth parameters were estimated (red). Boxes show the distribution of length at age determined from a t-copula with triangle marginal distributions

The resulting distributions of length at age (Figure 13) show wider ranges than those specified by the inputs to the 2016 assessment and a much wider range than the growth sensitivity runs described in Section 3. This would suggest that estimation error associated with the growth parameters is potentially large and may not provide useful information on the variability in growth to be assumed in the OM. However, we consider these analyses to be preliminary and recommend further work in this area. McKechnie et al. (2016b) note the difficulty in reliably estimating growth for skipjack within the assessment model and recommend that additional information on growth, such as length-increment data from tagging studies, be given high priority for future investigation.