SCIENTIFIC COMMITTEE
THIRTEENTH REGULAR SESSION
Rarotonga, Cook Islands
9–17 August 2017

Generating Pseudo-Data in MULTIFAN-CL

WCPFC-SC13-2017/MI-IP-02

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

[1]Oceanic Fisheries Programme, The Pacific Community
[2]Te Takina Ltd.

Contents


```
                Version                                Built
MFCL            2.0.2.1                           March 2017
R                 3.4.0 R 3.4.0; ; 2017-04-21 19:54:24 UTC; unix
FLCore   2.6.0.20170228 R 3.4.0; ; 2017-05-28 03:25:46 UTC; unix
FLR4MFCL          0.0.2 R 3.4.0; ; 2017-07-09 06:29:55 UTC; unix
```

Abstract

The MSE simulation framework comprises two main components: an operating model and a management procedure. The operating model is the mathematical representation of the system to be managed. It represents the biological components of the stock as well as the fishery that operates on the modelled population. It also includes models for the generation of data and the procedures for implementation of management regulations.

In this exercise we have used MULTIFAN-CL to generate pseudo catch, effort and length frequency data from the 2016 WCPO skipjack reference case assessment model and compare the historical observed data to the generated pseudo data. As a further analysis we ran a series of assessment runs over a range of function evaluations to see how well and how quickly MULTIFAN-CL was able to fit the pseudo generated data. We used the converged model solution from the 2016 reference case assessment as the starting point for these runs under the assumption that if the pseudo data were an accurate representation of the historic data series the model would find a converged solution to the extended data set in a relatively small number of function evaluations.

The initial attempts at generating pseudo data, detailed here, have highlighted a number of issues, both for the generation of catch and effort data and for the generation of length composition data. In many cases, resolving these issues will require further development of MULTIFAN-CL. We note that the generation of pseudo tag data is also prioritised on the list of developments required for MULTIFAN-CL along with functionality to generate pseudo data for the historical time series.

We note that the results presented here are preliminary and have been derived from initial investigations of on-going developments to the MULTIFAN-CL software.

1 Introduction

The MSE simulation framework comprises two main components: an operating model, that represents the true underlying dynamics of the stock and the fishery; and a management procedure, that comprises the methods used to estimate the status of the stock as well as the harvest control rules (HCRs) that determine what management action should be taken for a given estimate of stock status.

The operating model is the mathematical representation of the system to be managed. It represents the biological components of the stock as well as the fishery that operates on the modelled population. It also includes models for the generation of data and the procedures for implementation of management regulations. In this paper we focus specifically on the generation of pseudo data by the operating model.

Developing and conditioning the operating model is a complex and lengthy process. A primary consideration when designing and building an operating model is to identify the most important sources of uncertainty regarding the dynamics of the stock, the operations of the fishery and the quality of the data that will be used as inputs to the management procedure (Rademeyer et al., 2007).

The data upon which stock assessments, and therefore management strategies, are based are subject to uncertainty. Following Rosenberg and Restrepo (1994), Francis and Shotton (1997) and Kell et al. (2007), uncertainties in fish stock assessment and management can be categorised as follows:

- process error - caused by disregarding variability, temporal and spatial, in dynamic population and fishery processes.
- observation error - sampling error and measurement error.
- estimation error - arising when estimating parameters of the models used in the assessment procedure.
- model error - related to the ability of the model structure to capture the core of the system dynamics.
- implementation error - where the effects of management actions may differ from those intended.

The operating model must capture the overall dynamics of the underlying system and be capable of generating data, to be passed to the management procedure, that sufficiently represent that system, and have appropriate variability to reflect important sources of uncertainty.

In this exercise we have used MULTIFAN-CL to generate pseudo catch, effort and length frequency data based on the 2016 WCPO skipjack reference case assessment model and compare the historical observed data to the generated pseudo data. These comparisons are largely based on visual inspection, along the lines of the so-called Turing Test[3].

As a further analysis we ran a series of assessment runs over a range of function evaluations to see how well and how quickly MULTIFAN-CL was able to fit the pseudo generated data. The time taken to achieve an acceptable fit is an important consideration when selecting a model to use within a simulation framework that will be run over a large number of iterations and model

[3] Alan Turing proposed the "Imitation game", now more widely referred to as the "Turing Test" which considers the issue of whether people are able to distinguish between the responses of a person and a machine.

scenarios. In order to reduce model fitting time we used the converged model solution from the 2016 reference case assessment as the starting point for these runs under the assumption that, if the pseudo data were an accurate representation of the historic data series, the model would find a converged solution to the extended data set in a relatively small number of function evaluations.

1.1 Basis of pseudo-observations

Methods for generating pseudo observations in MULTIFAN-CL are detailed in WCPFC-SC-2017/SA-IP-05. A summary of the process is provided here for convenience. Observation error for catch and effort data are derived assuming a log-normal error distribution

$$X_{t,f} = x_{t,f}.exp^{\epsilon} \tag{1}$$

and

$$\epsilon \sim N(0,\sigma) - \frac{\sigma^2}{2} \tag{2}$$

Where $X_{t,f}$ is the model predicted catch or effort in fishery f in time period t.

Pseudo data incorporating observation error for the length/weight compositions was generated from a multinomial distribution.

$$q_{t,f} \sim Multinom(n_{t,f}, p_{t,f}) \tag{3}$$

where $p_{t,f}$ are the model predictions and $n_{t,f}$ is the assumed effective sample size for a fishing incident in fishery f in time period t.

2 Pseudo Data Generation for skipjack

The 2016 skipjack reference case stock assessment inputs have been used as the basis for the generation of pseudo data. Pseudo catch, effort and length frequency data have been generated from MULTIFAN-CL for a specified future period and the existing reference case MULTIFAN-CL solution was refitted to the extended time series.

To investigate the properties of the pseudo data we generated future data for a 30 year time period. Fishery specific catch and effort for the future period were based on quarterly mean values calculated for the recent three year period (2013 to 2015). The standard deviation for observation error in catch or effort was set to 0.1 and the effective sample size (ESS) for the multinomial re-sampling of length frequencies was set to 50. Pseudo catch and effort data were generated for all fleets except for the long-line fisheries for which catch (by number) remained fixed at 500 for each year-quarter and effort was set to -1 (to signify that effort data are missing).

Comparisons of pseudo and observed catch and effort data were conducted prior to any refitting of the assessment model. Subsequently, the model was refitted to the extended data set and time series of model estimates (effort deviations, catchability, selectivity, etc.) compared with those of the reference case assessment. Rather than undertake a full assessment fit we used the converged solution of the 2016 reference case assessment as a starting point to initialise the refitted model

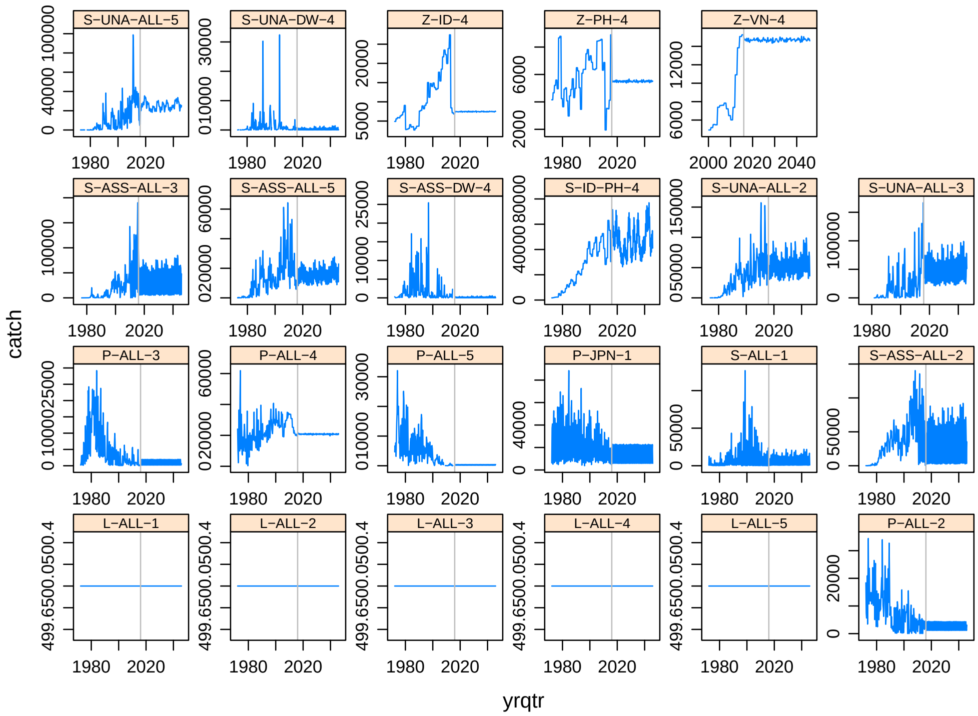

Figure 1: Historical and future catch for all fleets for a single iteration. 30 year projection

and to reduce the time taken to refit the model to the extended data series. A range of function evaluations were trialled to investigate convergence of the refitted model.

2.1 Catch and Effort Data

Future catch (Figure 1) and effort (Figure 2) are based on the quarterly arithmetic mean of the most recent three years (2013 to 2015). In some cases this involves setting future catches at either the maximum or minimum of the historically observed catch and effort values. For the most part, however, the future generated catch and effort values are broadly consistent with the historical time series, at least for the period over which they have been determined.

For several fleets, however, time series plots for CPUE (Figure 3) show abrupt changes in their overall level and seasonal patterns between the historical and future periods. This is particularly evident for the pole and line fleets in regions 1,2,3 and 5 (P-JPN-1; P-ALL-2; P-ALL-5 and P-ALL-3) but less so for pole and line in region 4 (P-ALL-4). CPUE indices of abundance are standardised for pole and line fleets in regions 1,2 and 3, whilst nominal catch and effort are used for those in regions 4 and 5.

To a lesser extent, many of the purse seine fleets are subject to a reduction in CPUE between the last year of observed data (2015) and the first year of pseudo data (2016).

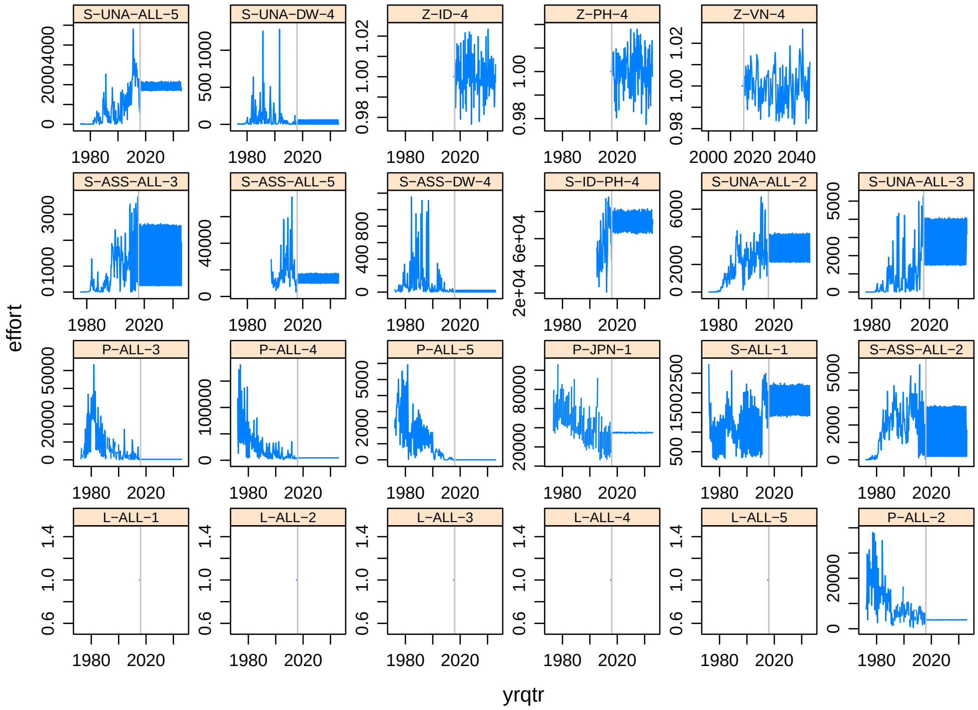

Figure 2: Historical and future effort for all fleets for a single iteration. 30 year projection

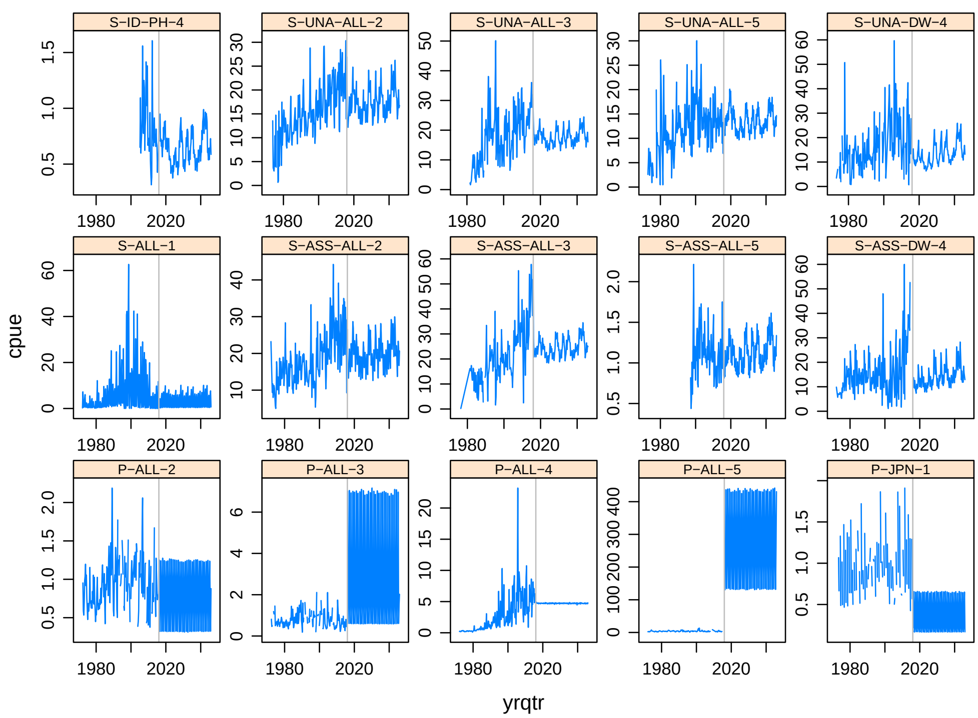

Figure 3: Historical and future CPUE for all fleets for a single iteration. 30 year projection. (Long line fleets and miscellaneous fleets in region 4 omitted)

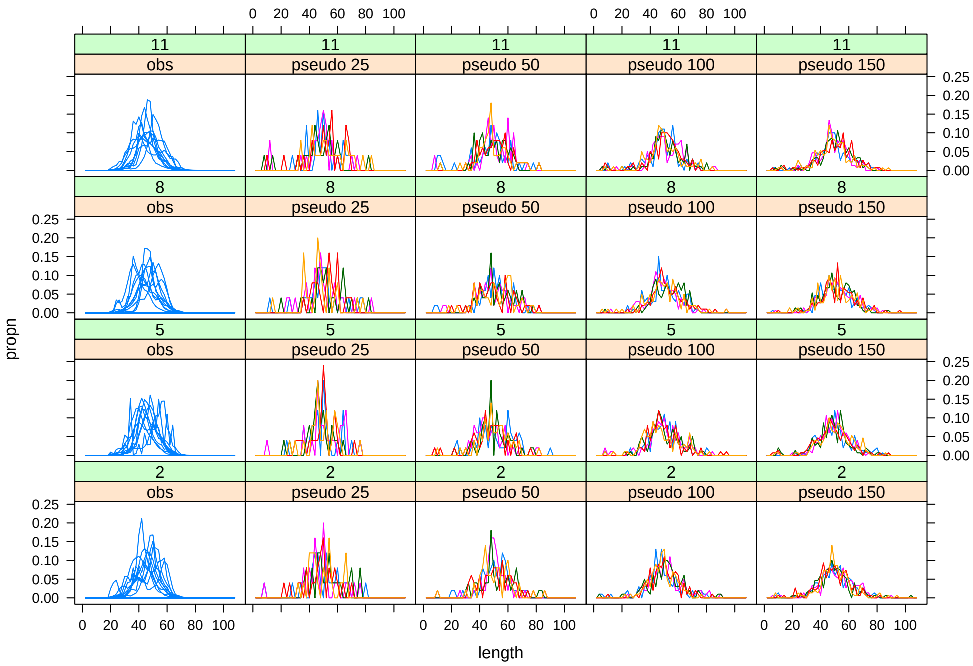

Figure 4: Comparative quarterly (months 2, 5, 8, 11) length compositions for the skipjack associated purse seine fishery in assessment region 2 (S-ASS-ALL-2). Historical observations for the period 2005:2015 are shown in the left column, and five iterations of pseudo data generated from MFCL for varying effective sample sizes (25, 50, 100, 150) in subsequent columns.

2.2 Length Composition Data

Length frequency data were generated for a range of ESS for comparison with historical observed length frequencies (Figure 4). Across the range of ESS values considered (25,50,100,150) the pseudo data exhibit more observations in the tails of the distribution, less auto-correlation and less modal variation than the historical observations. With increasing values for the ESS the proportion of pseudo observations in the tails of the length frequency declined and the distributions become increasingly smooth but the modal distribution remains invariant.

3 Re-fitting MFCL

At present the generation of pseudo data from MULTIFAN-CL is possible only for the projection period and it is not possible to conduct a direct comparison of pseudo generated and observed data. Instead we have generated pseudo data for a projection period and attempted to re-fit the MULTIFAN-CL to the extended data series.

MULTIFAN-CL was refitted to the extended time series of input data using the flag settings and par file estimates derived from the terminal phase of the 2016 reference case assessment model. A small number of changes were made to ensure that model would fit appropriately.

- The year range settings for the stock and recruitment curve were adjusted to accomodate the increased time series of the new data set. The year range over which the stock and recruitment curve was fitted was maintained at that of the refcase assessment (1982 to 2014).
- The number of function evaluations was set to the required value and the convergence criteria reduced to 10^{-3}.

A range of function evaluations were used when refitting the extended model with pseudo data to investigate the rate at which the model converges and the sensitivity of the model estimates to the number of function evaluations. Table 1 below shows the maximum gradient value for each of the runs with increasing function evaluations. It shows relatively poor levels of convergence, for a model data set that has been increased by 30 years, for model runs of less than 5000 function evaluations. Estimates of adult biomass from each of the nine runs with increasing function evaluations are shown in Figure 5. Adult biomass is initially estimated to be relatively low, but the estimates increase progressively as the number of function evaluations increases and eventually appears to stabilise at relatively high biomass.

Differences in the model diagnostics between the 2016 reference case assessment and the extended model with pseudo data are more difficult to interpret. Figures 6 and 7 show the effort deviation coefficients for extended models run for either 25 or 10,000 function evaluations in comparison with the reference case assessment model. Although similar trends are apparent for both the reference case and the extended model at 25 function evaluations, it is clear that, for many fleets, the trends are much closer after 10,000 function evaluations. It is notable that the effort deviations for the future period for the pole and line fishery in region 5 differ markedly between runs for 25 and

Table 1: Maximum gradients of the fitted model (additional 30 yrs pseudo data) after n function evaluations.

run	function evaluations	max gradient
1	25	4.77e+03
2	50	1.86e+03
3	100	1.46e+02
4	250	1.22e+02
5	500	2.33e+01
6	1000	5.03e+00
7	2000	3.56e+00
8	5000	2.92e-01
9	10000	3.72e-02

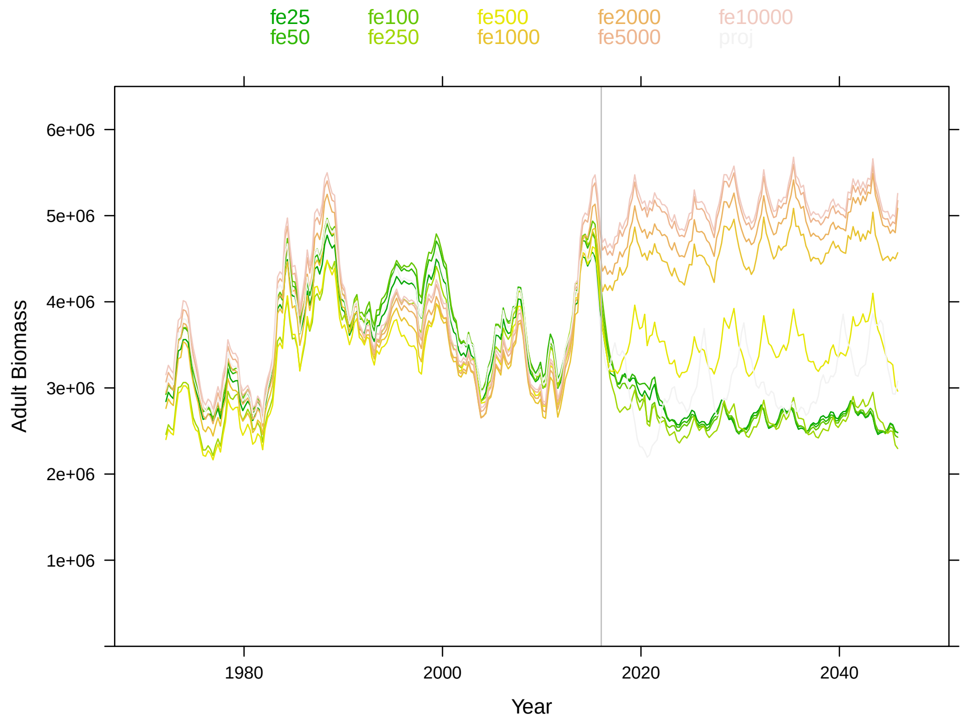

Figure 5: Estimates of adult biomass derived from the refitted MULTIFAN-CL model with increasing (green to brown) numbers of function evaluations. Vertical grey line denotes the transition between historically observed data and "future" pseudo generated data

10,000 iterations. However, neither of these fisheries account for a substantial volume of catch and it is unlikely that this difference alone is the cause of the change in future biomass with increasing function evaluations.

Trends in catchability estimates between the reference case and the extended model after 10,000 function evaluations (Fig 8) show differences for most fleets although in many cases, particularly those where catchability is constant in time, the differences are small, however, varying trends in catchability in time are apparent for many of the seine fisheries. These differences in catchability are not always in the same direction, estimates for some fisheries have increased whilst others have decreased. The shifts in estimates of catchability may be a significant factor affecting the estimates of future biomass, although further work will be required to determine the true cause. Estimates for other model quantities such as selection at age (Figure 9) and recruitment (Figure 10) remained very consistent between the reference case assessment run and the extended model.

The data used for re-fitting MULTIFAN-CL comprise a mixture of historical observations and future pseudo data. Differences in the statistical properties of the two sets of data will affect the ability of MULTIFAN-CL to converge to a fitted solution that is both close to the model from which the data were generated and achieved within a reasonable number of function evaluations.

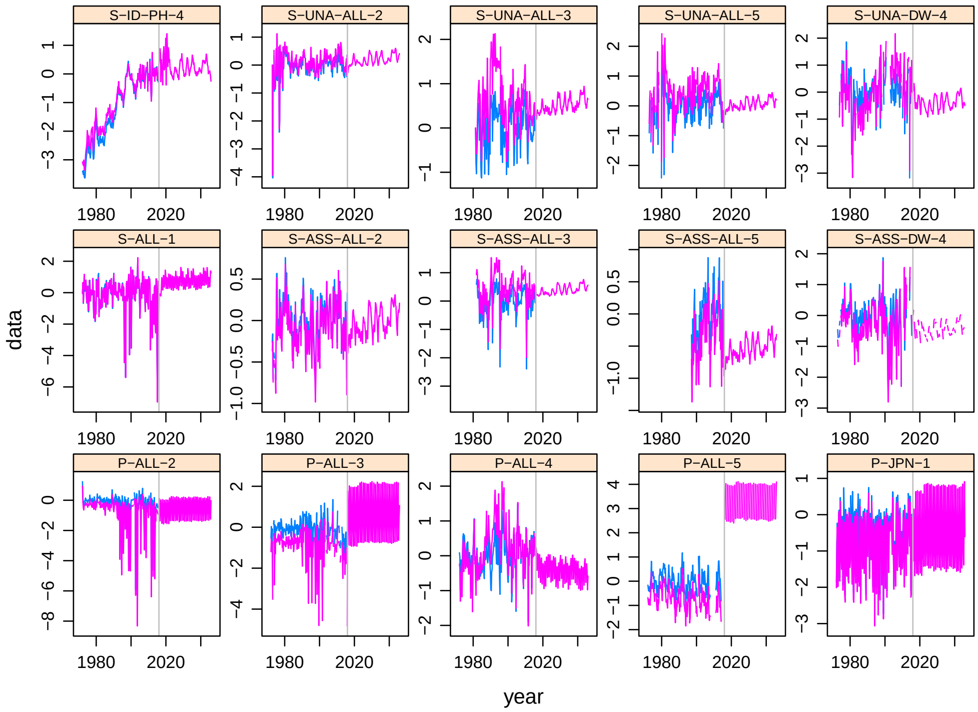

Figure 6: Effort dev coffs - Reference case (blue) and refitted MFCL model based on 25 function evaluations. Long line fleets and miscellaneous fleets in region 4 omitted

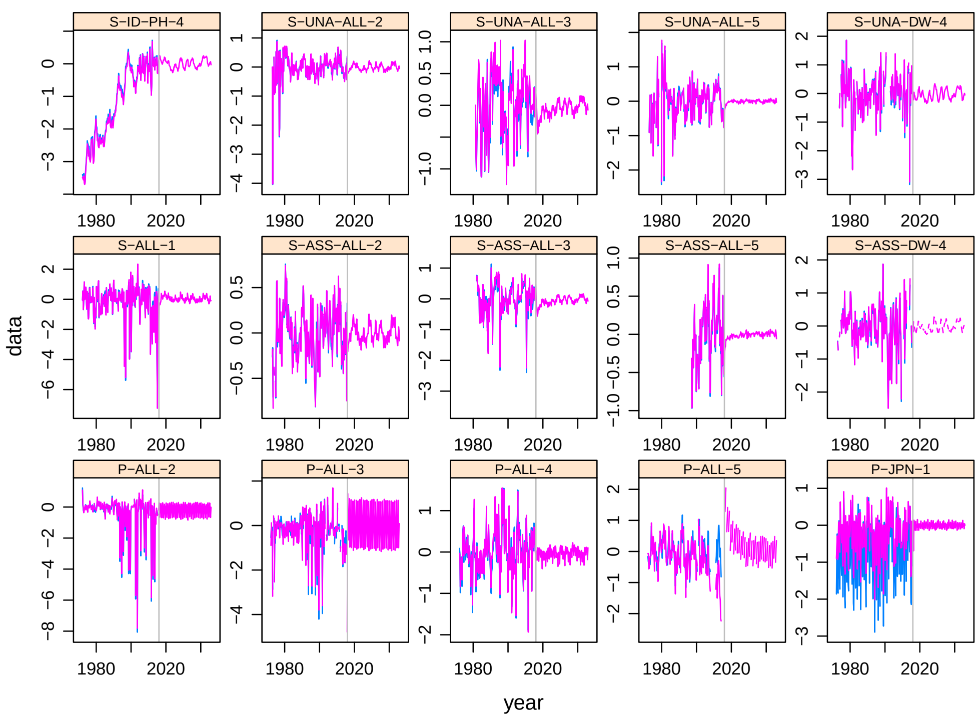

Figure 7: Effort dev coffs - Reference case (blue) and refitted MFCL model based on 10000 function evaluations. Long line fleets and miscellaneous fleets in region 4 omitted

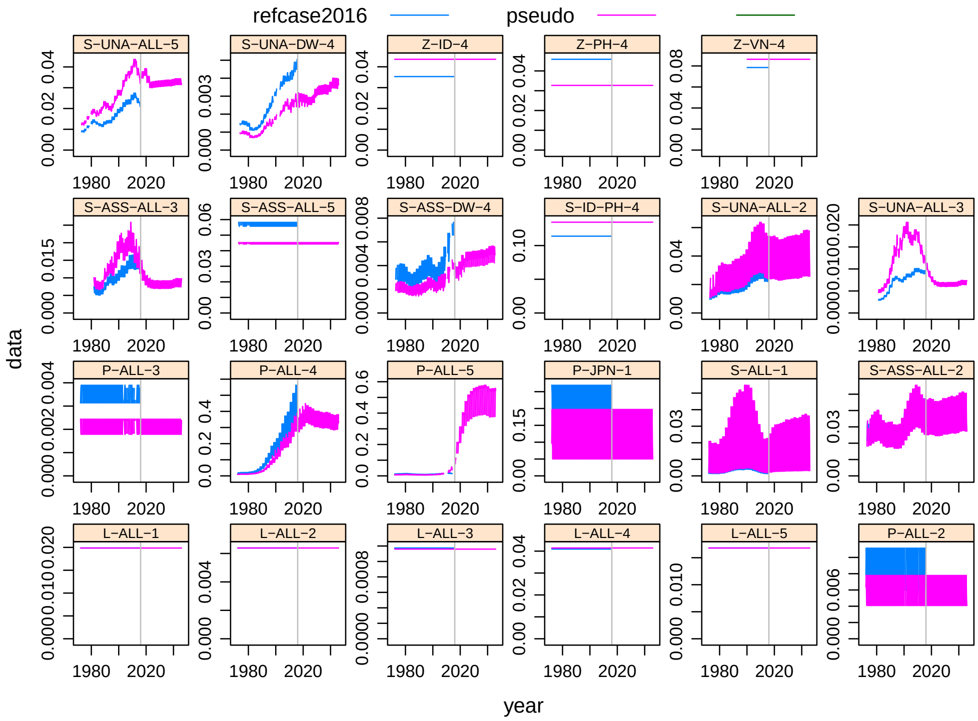

Figure 8: Catchability - Reference case (blue) and refitted MFCL model based on 10000 function evaluations.

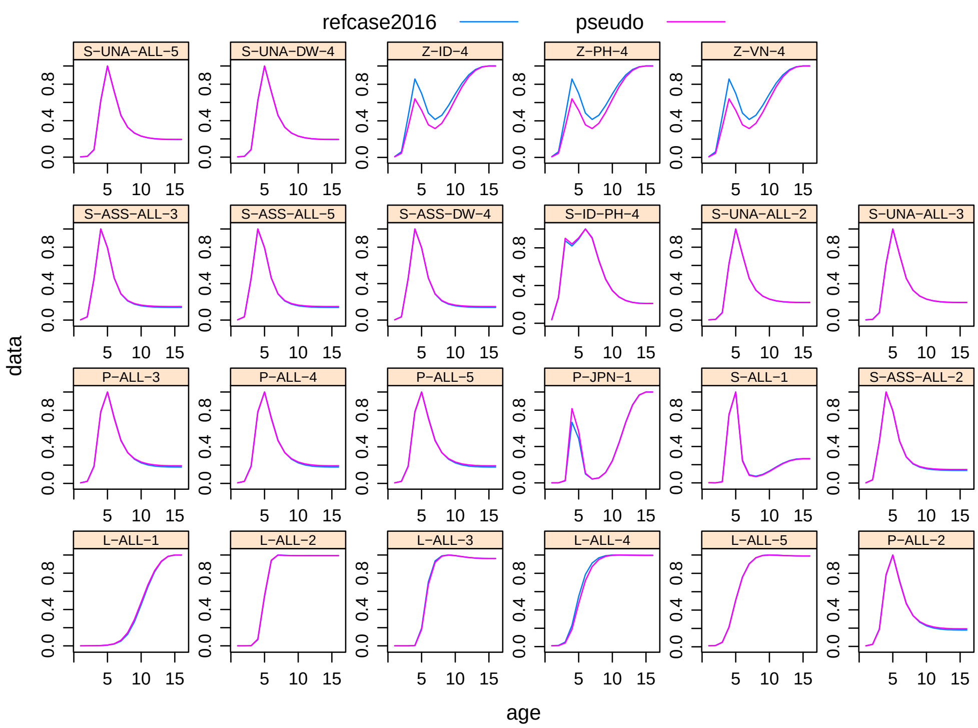

Figure 9: Selection at age - Reference case (blue) and refitted MFCL model based on 10000 function evaluations.

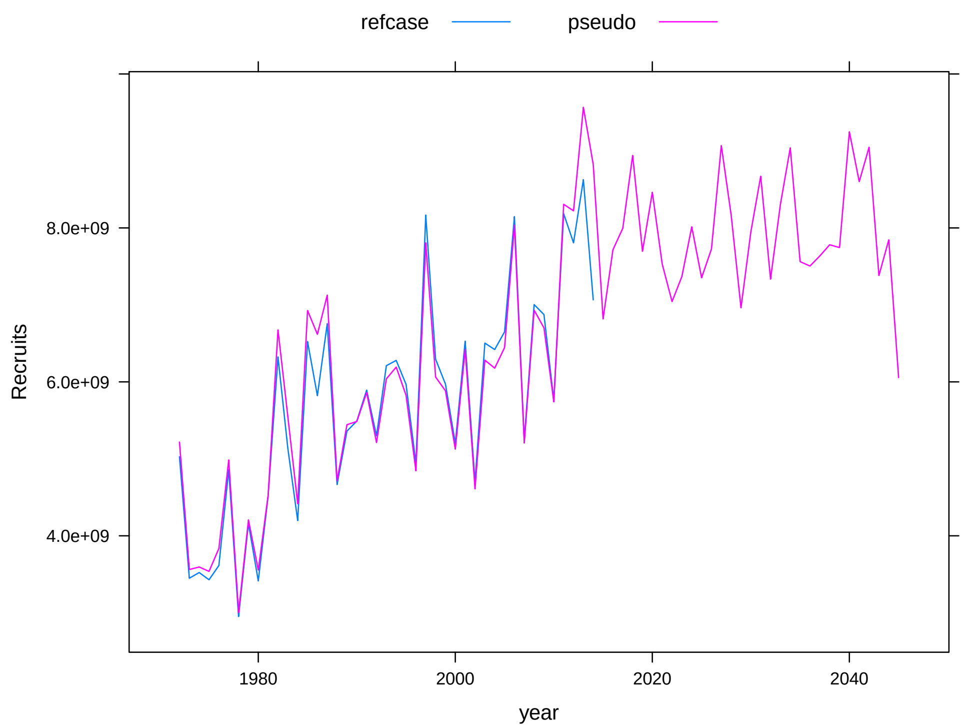

Figure 10: Annual recruitment for the 2016 reference case model and refitted MFCL model based on 10000 function evaluations.

4 Next Steps

This report presents preliminary results from initial investigations into the use of MULTIFAN-CL to generate pseudo data. We note that in some cases model settings have been based on relatively arbitrary assumptions and that more appropriate settings will need to be examined. Our initial attempts at generating pseudo data have, however, highlighted a number of issues. In some cases, resolving these issues will require further development of MULTIFAN-CL.

4.1 Catch, Effort and CPUE

There are substantial shifts in CPUE for some fleets between the historical, observed data and future generated data (Figure 3). This is particularly evident for some of the pole and line fisheries and, to a lesser extent, for some of the seine fisheries. We have assumed here a constant standard deviation of 0.1 for both catch and effort which was held constant through time and across all fisheries.

These initial settings have allowed for preliminary testing of the software but may not represent appropriate values to be carried forward when testing harvest control rules. Further work will be necessary to identify more appropriate fleet specific variance estimates for both catch and effort such that future generated data more appropriately represent historical observations. The assumed coefficient of variation of 0.1 (when generating pseudo catch and effort data) is relatively high with respect to catch and relatively low with respect to effort in comparison with standard assumptions for MULTIFAN-CL assessment models.

4.2 Length Composition Data

The characteristics of the pseudo length frequency data differ from those of the observed length frequencies. At low effective sample sizes the pseudo generated length frequencies show greater variability in the tails of the frequency distribution. As the effective sample size increases the variability in the tails decreases but continues to be greater than that of the observed data. In addition, throughout the range of effective sample sizes the modal distribution of the pseudo length frequency data remains invariant in comparison to the observed data for which greater variation in modal distribution is apparent.

The pseudo length frequency data are generated from multinomial re-sampling of the observed length frequency data. A more sophisticated approach that also takes account of the selectivity deviations may be a more appropriate method that would provide simulated length frequency data with greater modal variability.

4.3 Refitting MULTIFAN-CL

Currently it is only possible to produce pseudo data for the projection period and this makes it difficult to make detailed comparisons between observed and pseudo generated data sets. The ability to generate pseudo data for the historical time period in addition to the projection period has been tabled as a priority development requirement for MULTIFAN-CL.

In order to use MULTIFAN-CL inside the management procedure it will be necessary to find an appropriate model configuration that provides a sufficiently reliable estimation of stock status within an acceptable time frame. Achieving a reliable estimate of stock status is clearly the higher priority, however, doing so within a reasonable period of time is important if sufficient simulations are to be conducted and a sufficiently broad range of uncertainties are to be investigated.

Acknowledgements

We gratefully acknowledge funding for this work from the European Union through their funding support for the WCPFC "Simulation testing of reference points" and Project 75 (Technical Support for the Development of Harvest Strategies" and also from the New Zealand Ministry of Foreign Affairs and Trade (MFAT) funded project "Pacific Tuna Management Strategy Evaluation".

References

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.

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.

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.