Report of the Methods Working Group

15th Meeting of the Standing Committee on Tuna and Billfish

18 – 19 July 2002

Honolulu

Introduction

The SCTB Methods Working Group (MWG) held its second meeting on July 18 – 19, 2002 at the East-West Center in Honolulu, prior to 15th meeting of the Standing Committee on Tuna and Billfish.

John Sibert served as Chair and rapporteur. The agenda is presented in Attachment 1. The following individuals participated in the July 18 – 19 sessions: Shiham Adam, Melesia Aho, Keith Bigelow, Robert Campbell, Ray Conser, Paul Crone, Dave Fournier, John Hampton, Shelton Harley, Pierre Kleiber, Dale Kolody, Marc Labelle, Jacek Majkowski, Mark Maunder, Talbot Murray, John Sibert, Neville Smith, Chi-Lu Sun, and Chien-Hsiung Wang.

Simulation Results

MWG activity during the intercessional period between SCTB14 and SCTB15 was directed towards the work plan adopted at SCTB14. The SPC/OFP operational model was used to generate "simulated data" for evaluating the performance of several stock assessment models. Marc Labelle presented a brief overview of the operational model and preliminary results of testing against MULTIFAN-CL with data generated by the model (see MWG-1). Four basic scenarios were simulated. Each simulation models process and observation error by generating pseudo-random numbers to mimic these sources of variation. Multiple "realizations" of some scenarios sets were generated using different series of random numbers so that the statistical distribution of parameter estimates could be evaluated. Four fishery scenarios, intended to mimic aspects to the WCPO yellowfin fishery, were simulated:

1. One fishery, one region, one realization
2. One fishery, one region, ten realizations
3. Two fisheries, two regions, ten realizations
4. Sixteen fisheries, seven regions, ten realizations

Data sets one through three were circulated by email to MWG participants beginning in February 2002. Analysts sent their results to the MWG coordinator beginning in April, and the final results were collated during the July 18-19 MWG meeting.

The following stock assessment methods were applied to the data produced by the OFP operational model:

1. MULTIFAN-CL, Marc Labelle (MWG-1)
2. MULTIFAN-CL, Pierre Kleiber
3. SCALIA, Dale Kolody (MWG-5)

4. Age-structured (ASPM) and Fox production models, Dale Kolody (MWG-8)
5. ASCALA, Mark Maunder (MWG-7)
6. ADAPT, Keith Bigelow (MWG-6) and Ray Conser

The results of these analyses are summarized in Table 1. All models were not tested with all scenarios because of model limitations and time constraints. The table reports "actual" values of parameters as input to the simulations and the ratio of the values of the parameter estimated by each assessment method. Definitions of the model parameters can be found in MWG-1. Ratios less than 1.0 indicate that the estimate was lower than the actual value; ratios greater than 1.0 indicate overestimates. MWG-1 offers two potential biomass-based reference points: "BFe/BFs" is the ratio of fished biomass at the end of the time series to the biomass at the beginning of the time series; "BFe/Be" is the ratio of the fished biomass at the end of the time series to the biomass at the end of the time series in the absence of fishing.

Results vary considerably between both assessment method and scenarios. MWG participants considered the results to be preliminary, but felt that the results were "encouraging". In the simple scenario, the MFCL parameter estimates generally ranged between 0.5 and 2.0 times the actual values. The accuracy of some methods appeared to increase in more complex scenarios with higher exploitation rates. Production models appeared to be less suitable than more complex models. The preliminary results indicate that MFCL may provide more accurate estimates for some parameters than models without spatial structure, however this issue will be a focus of further testing. In general, ratios of parameter estimates (e.g. BFe/NFs, BFe/Be) appeared to be more accurate than the estimates of the individual parameters regardless of the model used. Chien-Hsiung Wang pointed out that in the context of production models, there are strong linkages between these ratios, biomass trends and levels of fishing mortality

MWG members agreed to continue the multi-model testing using simulated data sets during the coming intercessional period. The preferred scenarios consist of 1 region by 1 or 2 fisheries and 2 regions by 4 fisheries. The primary focus of the analysis will be on the effects of spatial stratification and use of mark-recapture data under moderate exploitation rates (F ~ M). Efforts will be made to generate and analyze 50-100 realizations of each scenario.

Introduction of more extensive errors into the simulations, particularly in relation to observational errors in the tagging data, was discussed extensively. Some participants felt that if such errors were found to have minimal impact on the accuracy of MFCL assessments, models that do not include tagging data and spatial structure should be dropped from the list of models to be considered. Other participants felt that overemphasis of a single model could stifle future model development. The current round of simulations was conducted with relatively low levels of error. Future simulations should incorporate more realistic process and observational errors in critical model components (e.g. natural mortality growth, effort, and tag reporting).

Other Matters

John Hampton reported on recent enhancements to MFCL (MWG-2). The model now provides ratio estimateors of biomass- and fishing mortality-based biological reference points required in some fishery management jurisdictions. **MWG participants were cautiously optimistic about these estimates given the preliminary results from the simulation studies on the accuracy of estimates of ratios.** Weight frequency data is often the usual information collected at fish auctions and landing ports, and in some instances, is the only available information on size. MFCL is now able to utilize these data in addition to the usual length-frequency data.

Pierre Kleiber summarized the results of a workshop on the use of oceanographic data in longline effort standardization (MWG-4). The workshop focused on the problem of trying to determine the effects of current shear on the depth of hook deployment.

MWG participants noted the importance of determining hook depth in longline sets and encourage the use of time-depth recorders and hook timers in longline sets.

Dave Foley reported on a project to make oceanographic data more easily available to fisheries scientists. In the near future, estimates of current shear and mixed layer depth will be made available at scales of resolution applicable to fisheries modelers. Problems with estimating depth of the oxycline and the deep scattering layer were discussed.

Mark Maunder also reported on a statistical method to evaluate habitat-based effort standardization. Maunder pointed out that there is no currently available means to determine if habitat standardization of effort has actually improved measures of fishing effort. He attempted to address this question by computing the likelihood of observing the catch given either the nominal effort or the standardized effort and concluded that standardized effort is a significantly better predictor of catch than nominal effort (MWG-7).

The MWG briefly reviewed its terms of reference in developing a work plan for the next intercessional period. Participants were encouraged by the simulations to date, and felt that the exercise should continue. However, in order to provide better advice to fishery managers, MWG participants felt that the simulations could be more comprehensive. As a longer-term goal, models should be developed that include management interventions based on the results of stock assessments and projections of the subsequent fisheries.

MWG recommends development of "management strategy evaluation models" over the longer term.

Ongoing review of the application of stock assessment models is identified as an important activity in the MWG terms of reference. MWG Participants agreed that one particular SCTB species assessment should be selected by SCTB15 for in-depth scrutiny prior to SCTB16 with the intention of reviewing a different species assessment each year. Examination of "diagnostic" statistics is a critical important aspect of assessment reviews that may reveal more subtle aspects of model performance. Ray Conser offered a list of useful diagnostics that might be applied to some of the models under evaluation (see Attachment 2).

MWG Work plan for 2002 – 2003:

The primary emphasis for the coming intercessional period will be a more systematic exploration of the effects of spatial structure. If time permits, effects of higher levels of process and observational error will be explored.

1. Continue simulations focusing on
   - 1 region x 1 fishery
   - 1 region x 2 fisheries
   - 2 regions x 4 fisheries
   - All at moderate exploitation rates (F~M) with > 50 realizations
   - Increase levels of process and observational error
2. Conduct an in-depth review of the MFCL yellowfin assessment for SCTB 16

Research Priorities for population modeling:

- Develop other biological reference points with less restrictive assumptions than B/B_{MSY} and F/F_{MSY}Begin development of "management strategy evaluation model", such as currently under development by CSIRO for the Australian swordfish fishery, that will evaluate the performance of the whole fishery management system including fisheries, data collection, stock assessment and regulation
- Begin development of multi-species models that will explore potential ecosystem effects species interactions
- Enhancements to MFCL

  –Seasonal variability movement (currently in progress)

  –Differential growth and mortality of males and females

  –Species misspecification, e.g. reporting of bigeye as yellwfin

  –Population projections ~ 5 years

Table 1. "Actual" values of parameters as input to the simulations and the ratio of the actual values to the values of selected parameters estimated by each assessment method. Definitions of the model parameters can be found in MWG-1. Ratios less than 1.0 indicate that the estimate was lower than the actual value; ratios greater than 1.0 indicate overestimates. "BFe/BFs" is the ratio of biomass at the end of the time series to the biomass at the beginning of the time series; "BFe/Be" is the ratio of the biomass at the end of the time series to the biomass at the end of the time series in the absence of fishing.

Model Parameter	**Actual**	**MFCL MWG-1**	**MFCL PK**	**A-SCALA**	**SCALIA**	**ADAPT**	**ASPM**	**Fox**
Scenario 1Fx1R (1)								
FL-at-age (range)	36 - 140	1.03	1.02					
M-at-age (range)	0.48 - 0.24	0.84	0.65			0.08†		
Mean biomass (mt, F=ON)	2.07E+06	2.46	1.83			0.27	0.42	
Mean biomass start	2.71E+06	2.16	1.71		0.60	0.27	0.61	2.44
Mean biomass end	1.98E+06	2.46	1.57		0.60	0.31	0.30	1.06
Mean S_biomass	8.53E+05	2.47	1.94			0.36	0.45	
Mean S_biomass start	1.34E+06	1.80	1.52		0.75	0.39	0.83	
Mean S_biomass end	1.12E+06	2.23	1.67		0.63	0.34	0.25	
Mean recruitment	6.96E+07	1.19	1.12		2.51	0.13	0.20	
Overall F by quarter	0.030	0.60	0.67			5.08		
Mean F start	0.035	0.59	0.67			5.14		
Mean F end	0.022	0.55	0.78			4.66		
BFe/BFs	0.732	1.14	0.94		0.97	0.98	0.50	0.42
BFe/Be	0.771	1.14	1.05		0.70		0.40	0.40
Scenario 1Fx1R (10)								
FL-at-age (range)	32 - 140	1.00	1.02					
M-at-age (range)	0.47 - 0.24	0.83	0.42					
Mean biomass (mt, F=ON)	2.32E+06	1.82	1.83				0.55	
Mean biomass start	2.98E+06	1.64	1.71		0.69		0.81	6.36
Mean biomass end	2.57E+06	1.68	1.57		0.68		0.46	4.94
Mean S_biomass	9.01E+05	2.05	1.94	12.19			0.64	
Mean S_biomass start	1.36E+06	1.52	1.52	9.96	0.84		1.18	
Mean S_biomass end	9.26E+05	1.94	1.67	11.71	0.80		0.59	
Mean recruitment	8.80E+07	1.61	1.12		0.35		0.22	
Overall F by quarter	0.033	0.72	0.67					
Mean F start	0.036	0.68	0.67					
Mean F end	0.021	0.80	0.78					
BFe/BFs	0.870	1.04	0.94	0.96	0.74		0.51	0.35
BFe/Be	0.796	1.05	1.05		0.76		0.50	0.89
Scenario 2Fx2R (10)								
FL-at-age	31 - 140	1.01	1.02					
M-at-age (range/Q)	0.53 - 0.13	0.89	0.55					
Mean biomass (w_fishery)	4.31E+06	1.32	1.20				117.77	
Mean biomass start	8.16E+06	0.67	0.61		0.90		62.58	3.84

† Assumed annual value; not estimated.

Model Parameter	Actual	MFCL MWG-1	MFCL PK	A-SCALA	SCALIA	ADAPT	ASPM	Fox
Mean biomass end	4.88E+06	1.18	1.16		0.83		109.12	4.59
Mean S_biomass	2.38E+06	1.33	1.15	1.78			144.43	
Mean S_biomass start	5.79E+06	0.49	0.42	1.24	0.93		59.97	
Mean S_biomass end	2.94E+06	1.19	1.11	1.65	0.83		124.76	
Mean recruitment	9.70E+07	1.37	0.85		0.53		35.68	
Overall F by quarter	0.070	0.90	0.99					
Mean F start	0.100	0.99	1.04					
Mean F end	0.050	0.98	0.87					
BFe/BFs	0.598	1.80	1.93	1.04	0.76		1.48	1.15
BFe/Be	0.486	0.92	0.93		0.74		1.55	1.35
Scenario 16Fx7R (10)								
FL-at-age	31 - 144	1.02						
M-at-age (range/Q)	0.53 - 0.13	1.35						
Mean biomass (w_fishery)	6.18E+06	1.05						
Mean biomass start	7.74E+06	1.20						
Mean biomass end	5.19E+06	1.60						
Mean S_biomass	4.64E+06	0.88						
Mean S_biomass start	5.98E+06	0.81						
Mean S_biomass end	3.78E+06	0.89						
Mean recruitment	9.45E+07	1.57						
Overall F by quarter	0.027							
Mean F start	0.004							
Mean F end	0.047							
BFe/BFs	0.670	1.10						
BFe/Be	0.660	1.03						

Standing Committee on Tuna and Billfish

Methods Working Group

Agenda

July 18, Imin Conference Center

1. INTRODUCTION – John Sibert
2. RESULTS OF SIMULATION STUDIES
   a. Operational model and "true" parameter values; MFCL example fits – Labelle; MWG – 1.
   b. SCALIA fits – Kolody; MWG - 5
   c. Production model fits – Kolody (tentative)
   d. ASCALA fits – Maunder
   e. ADAPT fits – Bigelow; MWG - 6
   f. Synthesis of results

July 19, Imin Conference Center

3. RECENT DEVELOPMENTS
   g. Reference points and incorporation of weight-frequency information in MFCL – Hampton; MWG-2
   h. Report of workshop on longline effort standardization – Kleiber
   i. A statistical method to integrate habitat based and GLM CPUE standardization - Maunder
   j. Availability of oceanographic data – Foley
4. FUTURE PLANS
   k. Review MWG terms of reference – Sibert
   l. Future plans: whether and how to do the simulation study correctly- group
5. OTHER BUSINESS
6. PREPARATION OF REPORT

July 23, Hawaii Convention Center – SCTB Plenary

7. MWG Report - Sibert
8. Access to Oceanographic Data for Fishery Managers – Foley
9. Fundamental limitation of TAC as the goal of fishery management – C. H. Wang; MWG - 3.
10. Other Business

Attachment 2

Intermediate Results and Diagnostics Commonly Used in
Reviewing Stock Assessment Modeling Results
(This is a draft list – additions are encouraged)

[1] Matrix of predicted catch numbers by age and time period. Similar matrices for stock numbers and instantaneous fishing mortality rates.

[2] Table of parameters estimated, values at the global solution, CVs, flags identifying parameters that hit constraints or significant penalties.

[3] Details of the phased estimation of parameters. Trace of initial parameter values at the beginning of each phase plus values of parameters (estimated and fixed) and likelihood components at the end of each stage of estimation (including at the global solution).

[4] Correlation among selected parameter estimates, namely those directly related to management advice, e.g. recent-period estimates of recruitment, catchability, selectivity, spawning biomass, etc.

[5] Examination of the response surface at the global solution - especially with respect to changes in key management parameters. For example, convergence checks using different initial value vectors. Likelihood profiling on key management parameters can also be informative here.

[6] Residuals summarized and plotted by various types (including but not limited to size composition residuals).

[7] Influence of priors. Plot priors vs. their respective posterior, including implied priors for key management parameters.

[8] Predictive capability of environmental factors. Develop predictive relationships and appropriate lags (e.g. for recruitment) using half of the available time series. Examine the utility of environmental factors when applied to the other half of the time series.

[9] Compare and contrast results obtained from other assessment methods, e.g. by applying commonly used age-structured models to the predicted catch-at-age data from the A-SCALA model.