13th Meeting of the Standing Committee on Tuna and Billfish

SCTB 13
Noumea, New Caledonia
5-12 July 2000

Working Paper

YFT-2

Report of the MULTIFAN-CL Workshop

(East-West Center, University of Hawaii, Honolulu, 1-3 February, 2000)

Oceanic Fisheries Programme
Secretariat of the Pacific Community
Noumea, New Caledonia

Report of the MULTIFAN-CL Workshop

East-West Center, University of Hawaii, Honolulu

1-3 February, 2000

List of participants

John Sibert, JIMAR, Honolulu
John Hampton, SPC, Noumea
Dave Fournier, Otter Research, Victoria
Chi-lu Sun, NTU
Sheng-ping Wang, NTU
Pierre Kleiber, NMFS, Honolulu
Marc Labelle, NMFS, Honolulu
Gerard DiNardo, NMFS, Honolulu
Marti McCracken, NMFS, Honolulu
Jim Ianelli, NMFS, Seattle
Ray Conser, NMFS, La Jolla
Mark Maunder, IATTC, La Jolla
George Watters, IATTC, La Jolla
Naozumi Miyabe, NRIFSF, Shumizu
Hiroshi Shono, NRIFSF, Shumizu
Rob Campbell, CSIRO, Hobart

1 Introduction

1.1 Opening Comments

John Sibert (chairman) welcomed all participants, commented on the various facilities available, transport and parking arrangements, and other logistical matters. He noted that the idea for this workshop originated at the 12th meeting of the Standing Committee on Tuna and Billfish (Papeete, 1999). The purpose of the workshop is to provide a forum for discussion and review of the MULTIFAN-CL model (Fournier et al. 1998) as it is currently being applied to tuna fisheries assessments in the western and central Pacific region. This will be facilitated by descriptions of analyses of existing data sets and by analyses of simulated data.

1.2 Appointment of Chairs and Rapporteurs

Pierre Kleiber was asked to act as interim chair in the absence of John Sibert on 2 February. Marc Labelle was appointed rapporteur, assisted by John Hampton.

1.3 Meeting Arrangements

The chairman invited all participants to attend a reception at the end of the first day of meeting at the Western Pacific Regional Fishery Management Council office.

2 MULTIFAN-CL overview

2.1 Description of MULTIFAN-CL

John Hampton gave a presentation on the various features of the model, as described in the draft manuscript distributed to the workshop (Hampton et al., in prep.). Briefly, the model consists of the following components (Figure 1):

- a general population dynamics model to predict catch and its size distribution
- a tagged population model to predict the catch of tags
- likelihood functions for the observed data
- priors and penalties to constrain parameterization
- parameter estimation by minimization of an objective function

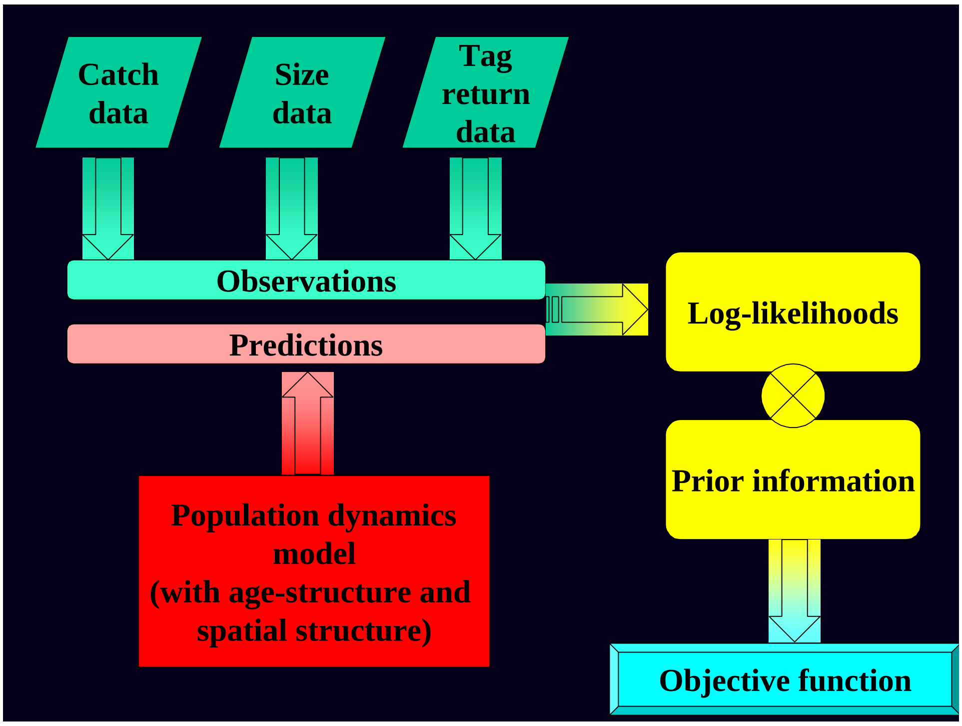

Figure 1. Schematic of the MULTIFAN-CL model.

Catch for all fisheries is currently expressed in numbers of fish, but could be expressed in weight if appropriate. Catch data must be available for each time period where a fishery exists, and are assumed to be relatively accurately measured. Length frequency data are required for all fisheries covered in the model, but missing data are allowed. Tagging data, if available, may also be used to fit the model. Effort data are not "data" in the sense of contributing to the likelihood function, but are used as an independent variable. Missing effort data can be accommodated by the model.

In terms of the population dynamics component, standard catch equations are used to describe changes in the age structure of the population. The parameterization used in the model allows for auto-correlation in the recruitment time series. Recruitment may be allowed to vary across regions and in time, with various constraints to such variation applied as appropriate, e.g. time-series variation in recruitment is assumed to be log normally distributed with specified variance. The periodicity of recruitment must be assumed in the model. For the yellowfin tuna analysis, initial attempts to fit a model with recruitment assumed to occur once per year were unsatisfactory in that the length data could not be adequately fitted. With recruitment assumed to occur four times per year (as an approximation to continuous recruitment), much better fits to the length data were obtained.

The model uses a 'plus' group to aggregate all age groups above a certain value. Generally speaking, the plus group comprises age groups that have essentially stopped growing and that have uniform selectivity characteristics for particular fisheries. The initial age structure for each region is generally computed to reflect the total mortality plus movement for some specified number of initial years of the analysis. While the initial age structure can be parameterized independently, it is generally very poorly determined, necessitating such an "equilibrium" assumption.

Movement is assumed to occur instantaneously at the beginning of a time period between spatially adjacent regions only. A fully implicit scheme is used to solve the equation. Movement rates can be allowed to be age dependent if appropriate.

The parameterization of fishing mortality conforms to the "separability" assumption (catchability distinct from selectivity). Selectivity for a particular fishery is generally considered to be constant over time (variable selectivity has been incorporated into the MULTIFAN-CL software, but previous analyses of simulated data indicate that estimation is not feasible). While the selectivity coefficients are age-specific, the variability among age classes is constrained according to overlap in the length distribution of the age classes. This reflects the notion that selectivity is likely in most cases to be a length-based rather than age-based process.

Catchability may be allowed to vary over time according to a time-series random walk process. The prior on the catchability deviations (i.e. the deviation in catchability that occurs at a step in the random walk) is zero, implying that systematic and known changes in the effectiveness of effort have been removed in a prior analysis. Seasonal changes in catchability are also allowed for in the model.

Several assumptions are applied in modeling the length data. First, lengths are assumed to be normally distributed for each age class. Secondly, the mean lengths at age are assumed to conform to a von Bertalanffy growth curve. Thirdly, the standard deviations of the lengths-at-age are assumed to be a simple linear function of the mean lengths. For species that show growth characteristics that differ systematically from von Bertalanffy growth, the mean lengths of a specified number of initial age classes may be estimated independently. This proved to be necessary for the yellowfin tuna analysis.

Standard catch equations are used to model the tag recovery patterns over time. Tagged "cohorts" (a group of releases made in the same region and time period) may be treated separately or pooled. Separating the cohorts for some period of time after release can potentially provide more information to the analysis, but comes at a high computational cost. For the yellowfin tuna analysis, tag reporting rates have been estimated for some fisheries through tag seeding experiments, and the estimates are used in the present yellowfin model. Note that, in principal, the model can estimate tag reporting rates independently, although such estimates may not be well determined unless independent information on tag reporting is available for some fisheries.

Log-likelihood contributions are obtained for the total catch data (by time period and fishery), the length data (by time period and fishery) and the tagging data (by cohort, age class, time period and fishery or fishery group). For the catch likelihood, the differences between the logged true and observed catches are assumed to be normally distributed with relatively low variance (CV less than about 10%). This implies that the catch data are accurately measured, which is a fundamental requirement for this type of assessment model. An "implicit" version of MULTIFAN-CL, in which the catches are assumed to have no error, is currently being developed.

For the length data, the log-likelihood function is based on a normal distribution, with terms added to improve robustness. The formulation used for the yellowfin analysis assumes that the variance of the length proportions is weakly dependent on the observed length proportions. Variance is also related to sample size where n<1000. The robust term used in the yellowfin analysis defines an outlier to be more than about three standard deviations from the mean (compared to two standard deviations previously used for the albacore analysis).

For tagging data, a Poisson log-likelihood function is used. A function that allows a more flexible specification of the variance, such as the negative binomial, may be preferable because the variance of tag returns generally seems to be larger than that allowed by a strictly Poisson process.

The model requires that Bayesian priors (assumed underlying distributions of parameters) be provided. For most of the parameters, the priors are uninformative, i.e. the estimates derived are determined only by the data. However, weak priors may be applied to a variety of parameters on a case by case basis (see Hampton and Fournier (in prep.) for details of the yellowfin tuna example). Somewhat informative priors need to be provided for the effort and catchability deviations. For the effort deviations, a CV of 20% is routinely used, although this can be varied in cases where the effort data for a particular fishery are believed to index fishing mortality either very poorly or very well. Catchability deviations are generally assumed to have lower variance than effort deviations (given the random walk assumption), typically with a CV of 0.1. Again, this variance can be specified for individual fisheries according to some prior belief about the likely magnitude of catchability changes.

2.2 IATTC model – a simplified version of MULTIFAN-CL

Mark Maunder described a length-based, age-structured model being developed at the IATTC for tuna stock assessment. This model, termed Age-Structured Statistical Catch-at-Length Analysis (A-SCALA), is largely a simplified version of MULTIFAN-CL (coded in AD Model Builder) with some new features added.

A-SCALA differs from MFCL in the following ways:

- initial F function of average survival
- selectivity curvature function of first + 2^{nd} difference and average length
- standard deviation of effort deviations is a function of actual effort
- catchability has seasonal, environmental and temporal components, and is modeled as a positive random walk
- recruitment is function of average recruitment, and an environmental component with an automatic scaling parameter

- anomalies in initial conditions are corrected for lognormal bias, and adjusted to reduce chance of over-parameterization.
- constraints on recruitment anomalies. Recruitment index are used to determine if environmental effects are significantly different than zero
- catch can be expressed in weights or numbers
- no spatial structure in the population, only in fishing operations.

Issues that were considered during development of the model included the efficiency of the code, the selectivity parameterization, constraints on parameterization, the number of estimation phases required, the best initial values to use, which parameters should be estimated or fixed, and the type and amount of simulated data required for testing.

2.3 Model performance diagnostics

Ray Conser noted that modelers are generally in the best position to determine which diagnostic features are most indicative of model performance. He asked the modelers what features they pay most attention to when assessing the performance of the model. John Hampton noted that he looks at how the model fits various time series of observations, including ancillary data sets that are not used directly during the fitting procedure. He noted that running the model many times under various conditions sometimes can reveal problems with particular constraints or assumptions. Mark Maunder noted that large departures between observed and predicted trends in catchability, effort, and other features help identify potential anomalies.

The issues of parameter confounding and the need for visual displays of final output and intermediate results to reveal such problems were raised. It was noted that work was in progress to create a graphical back end to MULTIFAN-CL (extending the existing java program that displays length data fits) that would be useful for assessment of model results. Additional work in this area, as well as model diagnostics more generally, was thought to be quite important for future model development.

3 Assessment of MULTIFAN-CL using simulated data

Jim Ianelli described the simulation model used to conduct a preliminary assessment of the performance of MFCL model (Ianelli, in prep.). Basically, the simulation model generates data sets with known parameter values and specified error structures. The model includes 2 fisheries (surface, longline) that operate for 40 years. The simulation model generates numbers at age, catch at age, mean lengths, biomass levels, and catch at length. Observation error is added to the true effort and a process error to the mean lengths-at-age, while the length data are generated by random sampling from specified distributions. However, the simulation model did not allow for movement, spatial structure, tagging data. The data sets generated for testing included cases characterized by greater/lesser effort variability, effort levels, recruitment variability, growth variability, catchability changes, length frequency sample sizes, and assumed error in catch data (see Ianelli (in prep.) for a complete description of cases).

The test results obtained so far indicate that MULTIFAN-CL accurately predicted time series of biomass, relative biomass, recruitment and catchability for cases that are best described as being relatively simple compared to reality. Estimated trends in biomass agreed well with the true trends even for cases with low sample sizes and where catches were under-reported by 10%. The model tended to slightly under-estimate *M*, which was not surprising given that tagging data were not included in the simulations.

Discussions focused on the range of simulations that should be conducted in the future. It was noted that there was an operational model used at the NMFS Honolulu Laboratory that could also be used to generate length frequency data sets, with different types of data contamination including movement effects (Labelle, in prep.). Mark Maunder indicated that some of the problems IATTC staff intended to address concerned the effects of data gaps, larger changes in growth rates, cumulative changes in catchability, flat biomass trajectories, gear interactions, and errors in catch statistics. It was generally agreed that there was a need for more exhaustive testing of MULTIFAN-CL and similar types of models using more realistic simulation models that, in particular, incorporate spatial structure and tagging data.

4 Progress with current analyses supported by the PFRP

4.1 Yellowfin

John Hampton gave a description of the available data on the yellowfin fishery in the Western and Central Pacific Ocean (WCPO). He also reviewed (i) the periods/regions/fisheries for which effort and length data was lacking, (ii) the procedure used to account for the apparent two-stage growth pattern of yellowfin, (iii) trends in predicted catchability and selectivity by fishery/region. Some of the points discussed were as follows (see Hampton and Fournier (2000) for a complete description of the yellowfin analysis):

- The model's estimates of exploitable population for the longline fisheries showed similar time-series trends to longline CPUE in each region, reflecting the model assumption regarding longline effort data quality. The ratio of exploitable population to CPUE was also similar for each region. In this analysis, fairly strong assumptions have been made regarding longline CPUE indexing abundance, which has been justified by the pre-standardization of the effort data.
- Natural mortality trends follow a U-shaped pattern with size (or age), progressively decreasing from 25 cm to 75 cm, and then increasing again (*M* range: 1.0-3.0). It was hypothesized that the increase in the estimated *M* in later life might be real (e.g. due to the onset of reproductive maturity), or simply be an artifact caused by declining longline selectivity at large size (non-decreasing selectivity was assumed).
- Movement patterns are assumed to be constant over time; therefore, there is no attempt to model movement in great detail, rather to capture the general dispersal characteristics of the stock over its life history. The general pattern estimated for yellowfin indicates considerable variability in movement coefficients among regions and age classes.
- By far the highest fishing mortality rates (annual catch divided by average annual population for a region) have occurred in region 3 (Philippines/Indonesia area). Exploitation rates for tropical regions in particular show progressive increases since 1962.
- The overall estimated recruitment pattern indicates slightly greater mean recruitment levels since 1975, with high seasonal variability in all years. Plots of recruitment versus spawning stock size do not reveal a significant relationship.
- Total biomass shows no long-term trend (although various cycles are apparent), and vary within the range of about 3-5 million tons over 1962-96.
- Estimated catches conform very well to reported catches, as intended by the assumed catch variance. The length frequency fits for different fisheries are generally quite good, particularly when aggregated over time. Likewise, the tag data fits are also satisfactory. Overall, there is no indication from the data fits of any serious problems with the analysis.

- Virgin adult stock biomass was predicted based on the assumption that the fishery exploited a “virgin” biomass during the first five years. The predicted ratio of current to virgin adult biomass is about 0.72.

Ray Conser questioned the narrow confidence intervals obtained for some of the estimates from the yellowfin analysis. For example, the terminal year stock biomass estimate had a CV of approximately 10%. In view of the great deal of uncertainty in biological parameters coupled with the high variance inherent in fisheries data generally, this CV appears to be overly optimistic. John Hampton noted that the incorporation of tagging data into the yellowfin assessment would have an impact on confidence intervals because tagging data are informative about absolute population size. Mark Maunder suggested that the confidence intervals could be tested by determining the proportion of simulated data set runs for which the estimated confidence intervals included the true value.

4.2 Bigeye

George Watters gave a presentation on the bigeye fishery in the EPO, and the state of the IATTC model for this fishery. Trends in effort for the various fisheries were shown. The catch length frequency data sets for each fishery were presented along with the predicted distributions. Mark Maunder noted that they were experiencing problems in fitting the length frequency data. The fits obtained with the latest version of the model were not considered to be very good. In some case there is considerable overlap, but in other cases, predicted and observed catches do not overlap. Length frequencies of catches are lacking for a substantial number of strata. Mark Maunder also noted that there were poor fits between the aggregated data sets as well. To improve the fits, they are attempting to (i) change the likelihood weights, (ii) discard length frequency data sets from the early years which were subject to sampling difficulties, (iii) fix the selectivity to free up variation in recruitment. Other options include adding parameters to re-define the fisheries in the model, and estimate time-specific selectivities and/or growth parameters. Other modeling problems to solve include estimated effort deviations from standardized effort series, and obtaining estimates of exploitable biomass trends that do not match the CPUE time series. Efforts will be made to separate the technological and environmental effects on catchability, and incorporate environmental indices in the recruitment and growth sub-models.

John Hampton described the ongoing collaborative study conducted by SPC/IATTC/NRIFSF to identify the preferred spatial stratification for bigeye assessment in the Pacific. For current assessment, the Pacific is divided into four regions that separate temperate and tropical fisheries in each hemisphere. Data time series cover the same period as in the yellowfin model. The limited tagging data available will be incorporated into the analysis. Longline effective effort will be estimated in a similar fashion to that used for the yellowfin analysis, and in particular will incorporate recent information on bigeye vertical distribution obtained from archival tag returns. Information was presented on the time series of catch, effort and size composition for specific time/area strata.

5 Conclusion

The workshop proved to be a valuable forum for discussion of the MULTIFAN-CL approach to stock assessment. The approach is fairly complex, and it is difficult for many scientists, particularly those not directly involved in using this type of model, to gain a thorough understanding of all of the technical issues involved. Hopefully, this workshop has played some part in facilitating a greater understanding of the methodology.

There appeared to be general agreement that the methodology was promising and further application of the approach to tuna stock assessment for this region was encouraged. Recommendations for further work were as follows:

- Further testing of MULTIFAN-CL and similar models using simulated data would be beneficial in understanding their characteristics. It would be desirable to extend the simulation model presented by Jim Ianelli to include spatial structure, movement and tagging data. Such an extended simulation model could be used to test simulated data sets having the same structure, data types and data coverage as some of the real data sets currently under analysis.
- Further work is required to determine the best method of obtaining estimates of uncertainty, conditional on the model, for important management state variables. Currently, MULTIFAN-CL uses the inverse-Hessian and delta methods to approximate the posterior distributions of these variables. The accuracy and efficiency of this method and of alternatives needs to be assessed using simulation analysis.
- Some additional features that could usefully be added to the MULTIFAN-CL software (based mainly on ideas stemming from the A-SCALA development and the needs of the bigeye application) include:
  - ⇒ incorporation of environmental effects on recruitment and catchability
  - ⇒ weight effort deviations by the actual level of effort
  - ⇒ incorporate sex-specific *M* and fitting of sex-ratio data
  - ⇒ seasonal recruitment
  - ⇒ seasonal movement
  - ⇒ implementation of the "implicit" method (no error in catch and computation of fishing mortality using the Newton Raphson) and Kalman filter approach to modelling time-series variation in catchability
  - ⇒ development of graphical tools for interpretation of model results
- The development of length-based, age-structured models (such as the IATTC effort with A-SCALA) from first principles is a useful exercise in helping scientists understand the statistical properties of complex models such as MULTIFAN-CL, and should be encouraged.

Acknowledgements

We are grateful to the Government of Taiwan and the Pelagic Fisheries Research Program (PFRP), University of Hawaii, for providing funding support for several participants to the workshop. We are also grateful to John Sibert and Dodie Lau of the PFRP for dealing with logistical arrangements and to Kitty Simonds and staff of the Western Pacific Regional Fisheries Management Council for their hospitality.

References

Fournier, D.A., J. Hampton, and J.R. Sibert. 1998. MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. Can. J. Fish. Aquat. Sci. 55: 2105–2116.

Hampton, J., and D. A. Fournier. (in prep.). Stock assessment of yellowfin tuna (*Thunnus albacares*) in the western and central Pacific Ocean using a spatially-disaggregated, length-based, age-structured model. Can. Jour. Fish. Aquat. Sci.

Ianelli, J. M. (in prep.). Some simulation analyses for evaluating length-based stock assessment methods. Progress Report. NMFS Alaska Fisheries Science Center.

Labelle, M. (in prep.). An operational model to evaluate assessment and management procedures for the north Pacific swordfish fishery. Technical Report. NMFS SW Fisheries Center.

Maunder, M., and G. Watters. (in prep.). A-SCALA: an age-structured statistical catch-at-length analysis. IATTC Bulletin.