

IOTC-2011-WPTT13-46 Rev_1

AD Model Builder Implemented Age-Structured Production Model (ASPM) software Users' Guide (ver. 1.0)

October, 2011

Rebecca Rademeyer
University of Cape Town, South Africa

Tom Nishida
(Corresponding author)
National Research Institute of Far Seas Fisheries, Fisheries Research Agency, Japan
(tnishida@affrc.go.jp)

Abstract and Notice (release of the software)

This is the users' manual describing how to use the AD Model Builder implemented Age-Structured Production Model (ASPM) software. **This software is free of charge. If someone wants to obtain this software, please contact the corresponding author.** After using this software and if any improvements are needed, please DO let us know. We will revise it and release the improved version in the (near) future. This software development project was funded by Fisheries Research Agency (FRA) of Japan (2008 and 2011) for the Tropical Tuna Group in the National Research Institute of Far Seas Fisheries (NRIFSF).

Contents

1. Introduction-----	01-04
2. How to run ASPM-----	04
3. Input files	
3.1 <i>Control file</i> -----	05
3.2 <i>Initial parameter values file</i> -----	06
3.3 <i>Biological data file</i> -----	06
3.4 <i>Index file</i> -----	07
3.5 <i>Fishery file</i> -----	07
3.6 <i>Projection file</i> -----	08
4. Running MCMC-----	08-09
5. Output files-----	09-12
6. Suggestions-----	12
Acknowledgement-----	13
References-----	13-14
Appendix A: Formation of the ASPM-----	15-20

Submitted to the IOTC WPTT13 (16–23 October 2011), Paradise Island, Lankanfinolhu, North Malé Atoll, Maldives

1. INTRODUCTION

Conceptually ASPMs fall somewhere between simple biomass-based production models (e.g., Schaefer 1957; Prager 1994) and the more data-demanding sequential age-structured population analyses (Megrey, 1989) and integrated models such as SS3, CASAL and MULTIFAN-CL. Typically, simple production models estimate parameters related to carrying capacity, rate of productivity, biomass at the start of the time series, and coefficients that scale indices of abundance to the absolute magnitude of biomass. ASPMs estimate similar parameters but make use of age-structured computations internally, rather than lumped-biomass ones and directly estimate parameters of a stock-recruitment relationship. Their main advantage over simpler production models is that they can make use of age-specific indices of relative abundance and the spawner-recruitment relationship.

There are a number of applications of ASPM for various species in the past. As our experience are mainly on tuna stock assessments, here we introduce a few application of ASPM on tuna in the past. In the International Commission for the Conservation of Atlantic Tunas (ICCAT), ASPM were applied for albacore tuna (*Thunnus alalunga*) in the south Atlantic and bluefin tuna (*Thunnus thynnus*) in the western Atlantic. In the Indian Ocean Tuna Commission (IOTC), ASPM were applied for bigeye tuna (*Thunnus obesus*) and yellowfin tuna (*Thunnus albacares*) (IOTC, 2002-2008).

The above mentioned ASPM software was first coded in FORTRAN by Restrepo (ICCAT, 1997). However, this FORTAN implemented ASPM has the following limitations:

- Very slow operating speed especially to conduct the bootstrap to estimate variances;
- It can only handle a maximum of 4 fleets;
- Steepness of the stock-recruitment curve is estimated and cannot be fixed. This has caused problems in past assessment as steepness was estimated to be unrealistically high (0.999) or low (less than 0.4). The ability to fix the steepness in ASPM runs and evaluating sensitivities could provide more reliable results.

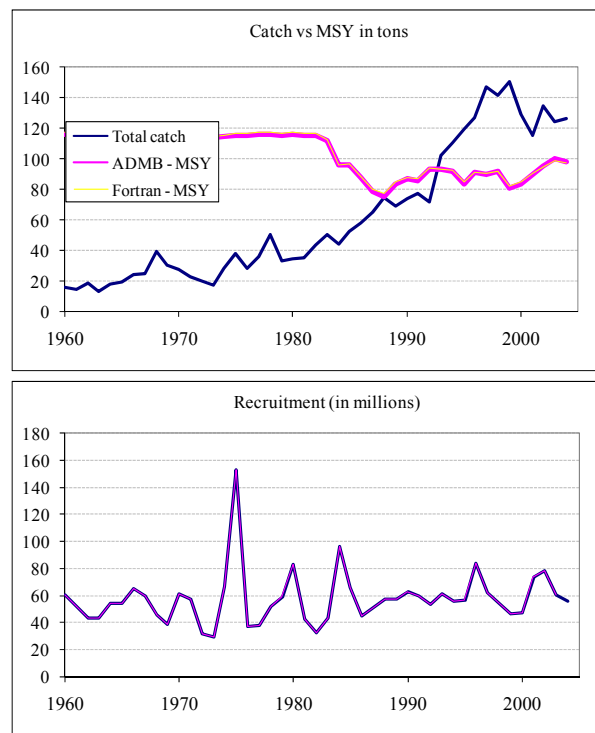
To improve these problems, we started to re-code the “FORTAN based ASPM” to “AD Model Builder implemented ASPM” in 2008 and completed in this time (2011). It took 4 years as we temporarily stopped development and testing works for 2 years. This software development has been funded by the Fisheries Research Agency (FRA), Japan. Table 1 summarizes the differences between the FORTAN and AD Model Builder implementations.

Table 1 Comparison between the FORTRAN and AD Model Builder ASPM implementations.

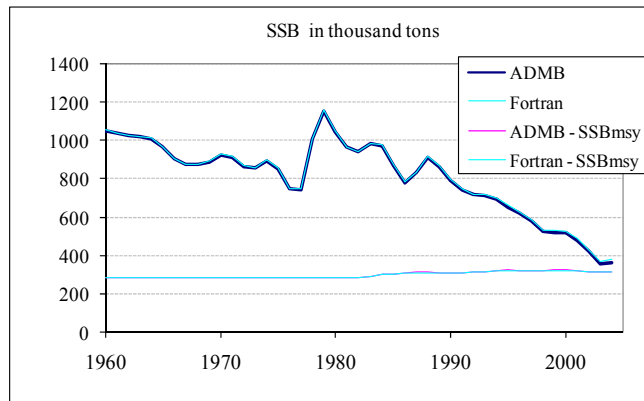
	FORATN ASPM	AD Model Builder ASPM <i>(refer to Appendix A: Formation)</i>
Fleets	4	No limits
steepness	Estimated	Estimated or Fixed between 0 and 0.95
computing speed	1	10 times or more faster
catch equations	Pope's approximation	
minimization routines	simplex algorithm	automatic differentiation
uncertainty (variances) estimation	Bootstrap	delta method MCMC
selectivity	Fixed	Fixed or Estimated (model free, ad hoc basis)
Others		Negative log likelihood is computed without the constants.

Performance tests

Runs with the FORTRAN and ADMB implementations for bigeye tuna in the IOTC were compared in WPTT08 in 2006. The SA results by FORTRAN ASPM were agreed by the WPTT8 and have included in the IOTC's BET Executive Summary in 2006. The three figures below show the results of the performance tests. The negligible discrepancies were caused by the different type of optimization methods used between two implementations. In addition CSIRO scientists in Australia evaluated the FORTRAN ASPM and it was concluded that results by the FORTRAN ASPM were robust (2004). This implies that ADMB ASPM is also robust as both are almost identical.



Results between two types of ASPM (two lines) are almost identical.

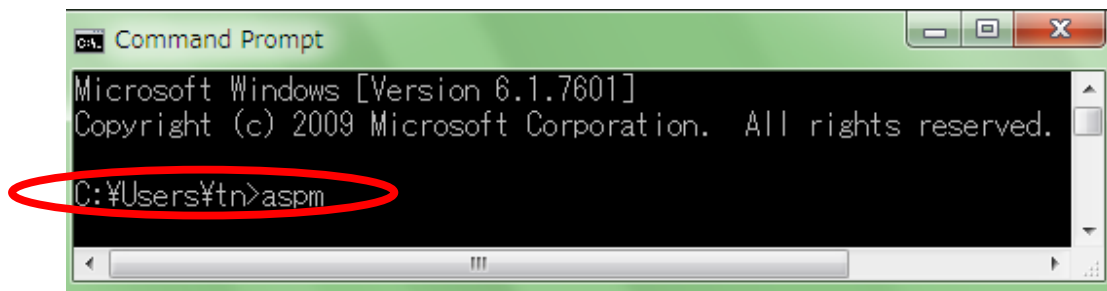


2. HOW TO RUN ASPM

To run ASPM, six input files are needed (see next Section 3: INPUR FILES). One of these is the *control* file, which controls some aspects of the run such as the problem's dimensions.

To run aspm.

in the command DOS prompt, type > **aspm** (see the example below)



The results will be sent to a series of output files which are described in Section 5 (OUTPUT FILES). To run with starting values other than those in the *parameter guesses* file (ASPM.pin, see below) (for example in a file named otherInitial.par), type

aspm -ainp otherInitial.par

3. INPUT FILES

Six input files are required, their names cannot be changed: A *control* file, a *parameter guesses* file, a *biological data* file, an *index* file, a *fishery* file, and a *projection* file. In all these files, “#” precedes a comment line. In all the input files, extra comment lines can be added without affecting the run, as long as these line starts with “#”. The example files contain annotations for input help. Their contents are as follows:

3.1 Control File (named: control.inp)

Line	Entry
1-3	Comments
4	First year of catch data
5	Last year of catch data
6	Comment
7	Number of fisheries or gear types for which yield data are available
8	Comment
9	Maximum age considered (taken to be a plus group)
10	Comment
11	Stock-recruitment curve (1= Beverton-Holt, other = Ricker).
12	Comment
13	First year with recruitment fluctuations (<i>see equation (21), p16</i>)
14	Last year with recruitment fluctuations
15	Comment
16	Standard deviation (σ_R) for the stock recruitment fluctuations [<i>see eq. (4), p12 and (21), p16</i>]
17	Comment
18	Deterministic (= -1) or stochastic recruitment (>0)
19	Comment
20	Phase for dummy parameter. Always -1 except if running with ALL the parameters fixed, in that case it needs to be >0.
21	Comment
22	Phase for estimation of virgin spawning biomass (K^{sp}), i.e. whether virgin biomass is estimated (>0) or fixed (to value read in aspm.pin) (-1)
23	Comment
24	Phase for estimation of steepness (h), i.e. whether steepness is estimated (>0) or fixed (to value read in aspm.pin) (-1)
25	Comment
26	Phase for estimation of initial biomass as a fraction of virgin biomass [θ , <i>see equation (8), p12</i>], i.e. whether θ is estimated (>0) or fixed (to value read in aspm.pin) (-1)
27	Comment
28	Phase for estimation of deviation from equilibrium age-structure in the first year [ϕ , <i>see equation (11), p13</i>], i.e. whether ϕ is estimated (>0) or fixed (to value read in aspm.pin) (-1). ϕ characterizes the average fishing proportion over the years immediately preceding y_0 ... If there is no fishing before the start year, $\phi = 0$
29	Comment
30	Phase for estimation of selectivity, i.e. whether selectivity is estimated (>0) or fixed (to value read in aspm.pin) (-1)
31	Comment
32	Type of weighting for the CPUE indices: 1 = maximum likelihood, 2 = equal weights, -1 = inverse-variance weighting for each point.

Note

The inputs corresponding to “phases” (lines 22, 24, 28 and 30) allow the minimization to be carried out over a subset of the parameters, while the others are fixed. In a non-linear model it can be useful to estimate the different parameters during different phase.

3.2 Parameter guesses File (named: aspm.pin)

Line	Entry
1	Comment
2	Guess for ln of virgin spawning biomass ($\ln(SSB_0)$)
3	Comment
4	Guess for steepness
5	Comment
6	Guess for θ
7	Comment
8	Guess for ϕ
9	Comment
10	Comment
11-i	Guesses for recruitment deviations
i+1	Comment
i+2-ii	Guess for commercial selectivities

Notes:

If you want the biomass at the beginning of the first year to be the virgin biomass, line 6 must be 1 and line 8 must be zero, with the phase parameters for θ and ϕ (lines 26 and 28 of the control.inp file must be negative).

Guesses for the commercial selectivities go in the order: fleet, periods, ages (minus to plus). E.g. two fleets, the first fleet has two selectivity periods and age minus=0, age plus=5, second fleet has only one period and goes from age minus=2 to age plus=7:

#	0	1	2	3	4	5	6	7	
0.1	0.2	0.3	1.0	0.6	0.3				# fleet1, period 1
0.15	0.3	0.4	1.0	0.5	0.2				# fleet1, period 2
		0.6	0.8	0.9	1.0	0.8	0.6		# fleet2, period 1

Guesses for recruitment deviations are not really necessary.

3.3 Biological Data File (named: biological.inp)

Line	Entry
1-5	Comments
6	Weights at age at the start of the year for all ages
7-9	Comments
10	Weights at age at the middle of the year
11-14	Comments
15	Natural mortality at age
16-18	Comments
19	Proportion maturity at age

3.4 Index File (named: Index.inp)

Line	Entry
1-3	Comments
4	Number of indices
5	Comments
6	Index type for each index (1 = biomass, other = numbers).
7	Comments
8	Index timing: Month of the year or -1 for mid-year.
9	Comments
10	Fleet the index corresponds to
11-13	Comments
14..Nyears:	Year, index value (0 for missing), CV(index),
...	start again for next index

NOTE:

The CV (index) values will be ignored if ML or equal weighting options are selected in the control file. The indices must start from the first year of the assessment period, if index doesn't exist for a particular year, set to zero.

3.5 Fishery File (named: fishery.inp)

Line	Entry
1-4	Comments
5	Number of selectivity periods for each fleet
6	Comments
	For each fleet
	Comments
	First year and last year for each selectivity period
x	Comments
x	Minus group for each fleet
x	Comments
x	Plus group for each fleet
x	Comments
x	Weight given to each commercial catch-at-age data set in the likelihood (for each fleet)
x	Comments
x	for each year: year, catch for each fleet
	For each fleet
x	Comments
x	for each year: catch-at-age for each age

NOTE:

If the selectivities are fixed (line 30 in control.inp is set to -1), then have the minus groups for each fleet be 0 and the plus groups for each fleet be the maximum age.

If the model is not fitted to the catch-at-age information (CAA weight=0), the program will still compare the observed and predicted CAA (but it won't be included in the likelihood). So if you don't have CAA information, rather fill the matrix with zeros.

3.6 Projections File (named: projection.inp)

Line	Entry
1-3	Comments
4	Number of years to project forward
5	Comments
6	Project with constant catch (=1) or constant F (=2)
7	Comments
8	Future catch by fleet if project with constant catch
9	Comments
10	Future F by fleet if project with constant F

Note

For each fleet, selectivity in the future is assumed to be the average of the last 5 years. In a maximum likelihood run, the projections are deterministic, i.e. the recruitment is exactly determined by the stock-recruitment curve. If MCMC are run however, variations around the stock-recruitment curves are generated for each year, i.e. stochastic projections.

4. RUNNING MCMCS

ADMB includes a Markov Chain Monte Carlo (MCMC) routine for Bayesian analysis. The objective is to represent the posterior distributions by means of a (large) number of vectors of parameters. The basic idea is to set up a (long) “chain” which starts at a pre-specified parameter vector and that then traverses the posterior distribution. The contribution from *Equation 21 (p16)* correspond to a prior on the distribution of the recruitment residuals, while priors on the other estimable parameters (SSB_0 , θ , ϕ and the selectivity parameters) are taken to be uniform over wide and/or feasible ranges with the intent that they be uninformative. The initial parameter vector used to start the MCMC computational process is the mode of the posterior. A chain of N iterations is run and the chain is “thinned” by taking every mth value in the chain. The results of the first iterations (5-50% of the total chain length) should be discarded to allow for a “burn-in” period, i.e. reduce the impact of the initial parameter vector.

To run MCMCs, type

aspm -mcmc N -mcsave m

where N is the number of simulations performed and every mth simulations are saved.

To get the desired output, type

aspm -mceval

See the ADMB manual for further MCMC options.

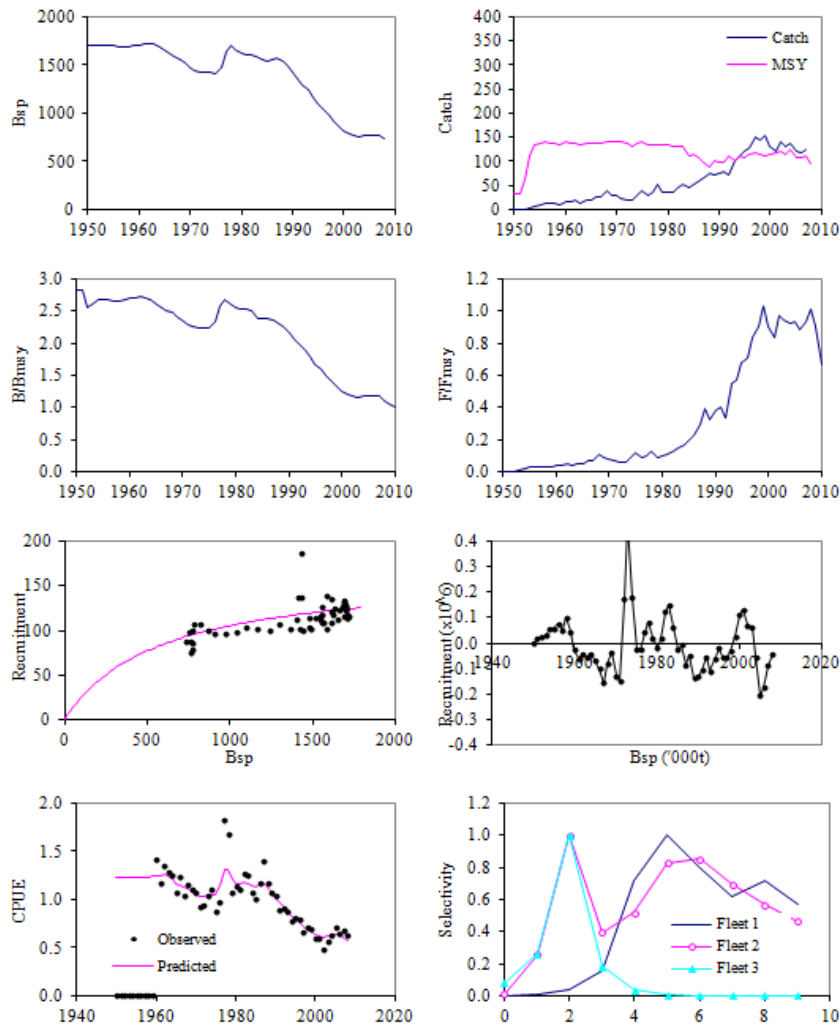
The output file **ASPM.hst** contains the means, standard deviations and observed distributions for all parameters included in the **aspm.std** file (after the estimable parameters, i.e. starting from K downwards). Once MCMC for a particular model has been run, stochastic projections can be run for different future catches (or F) without rerunning the MCMC, just changing the future catches and rerun **'aspm -mceval'**.

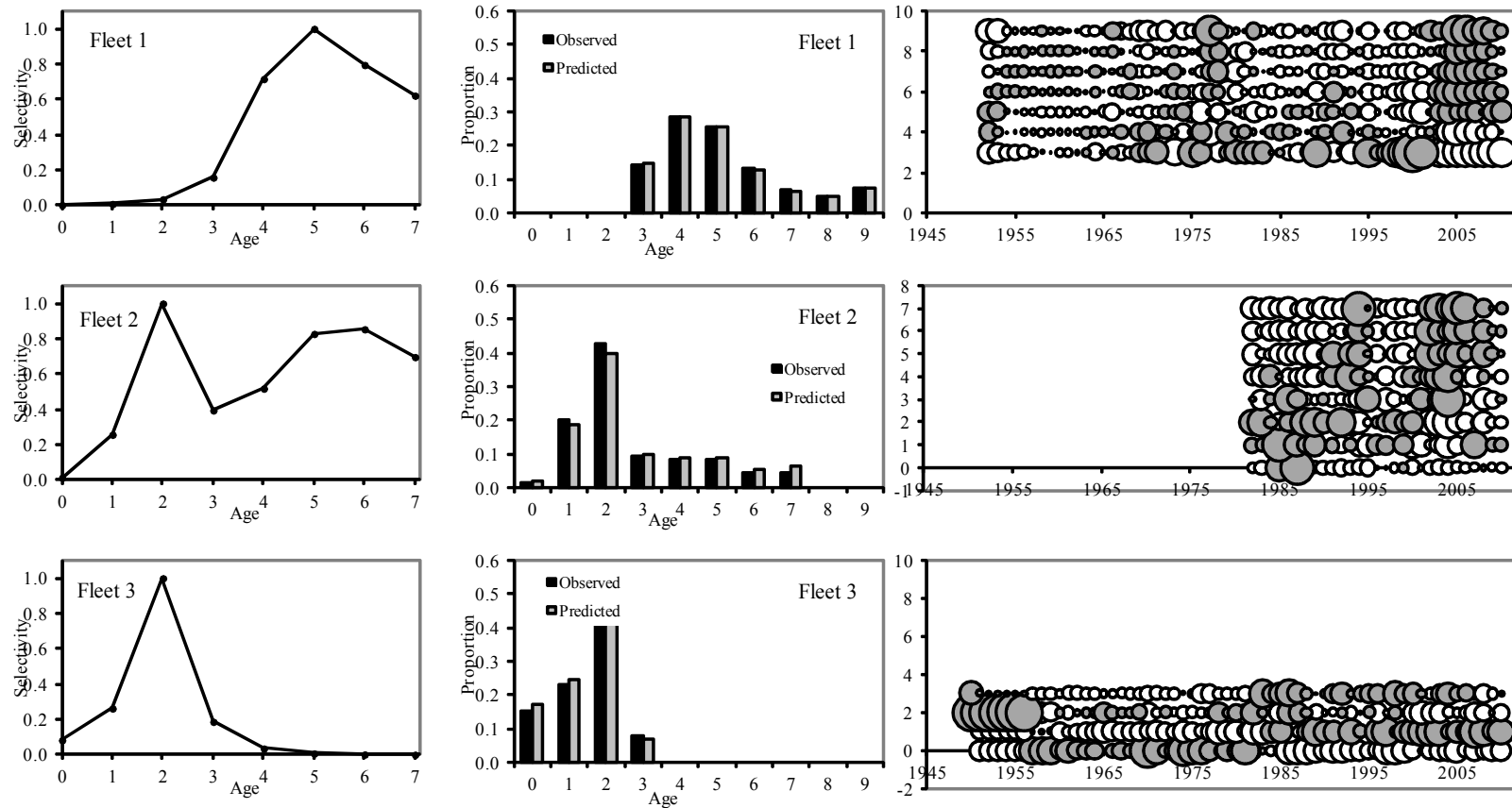
5. OUTPUT FILES

ASPM creates a series of output files, most of them are standard ADMB output files.

1. **ASPM.rep**: Results for the run, including spawning biomass, numbers-at-age, recruitment, fits to the data, fishing mortality, MSY and related quantities, etc. **"Bexp"** refers to the exploitable biomass (*see equation 7, p12*). **"Fmort"** is the fishing mortality.

Next two panels show some examples (1950-2010) of the results obtained by ASPM.rep.





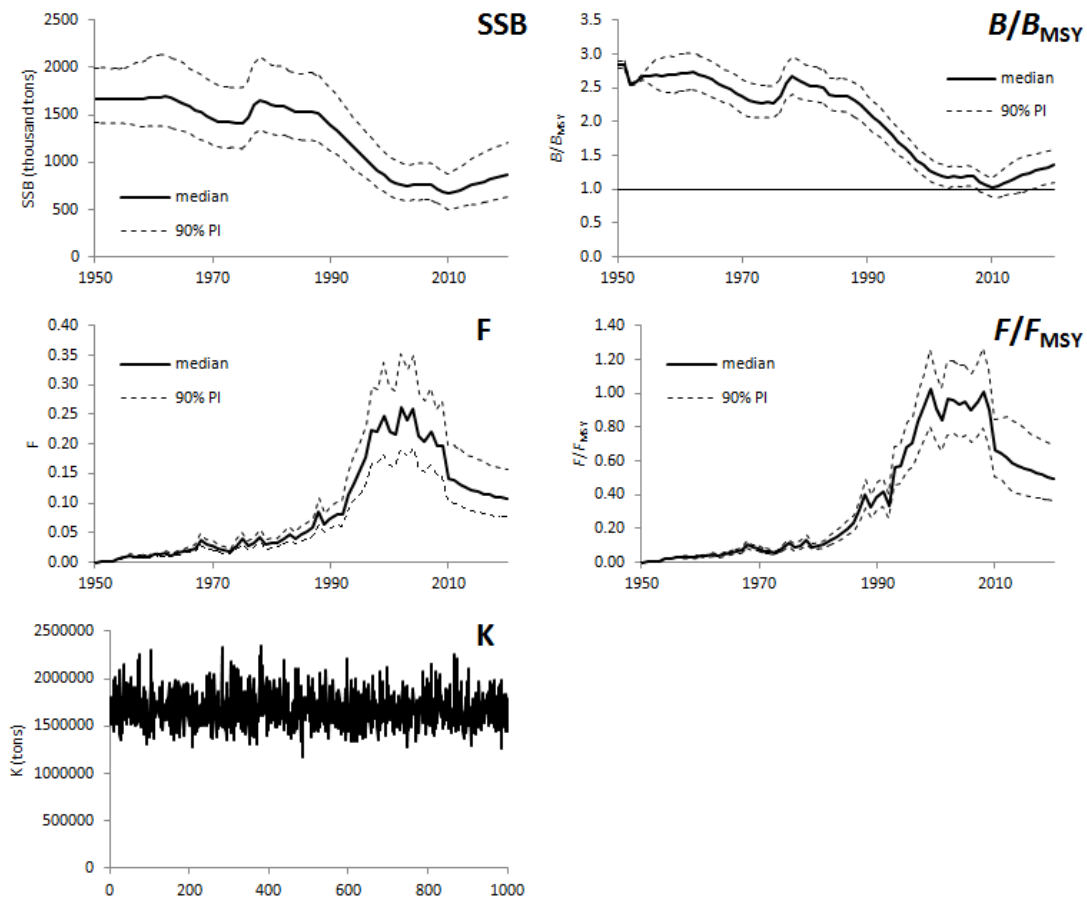
Some sample output of residuals plots by flee obtained by the MCMC run

2. **proj.out**: Projected spawning biomass and projected SSB_{MSY} .
3. **ASPM.par**: a standard ADMB output file, giving the objective function value, its gradient (this should be very small if the model has converged) and the parameters estimated/fixed for that run.
4. **ASPM.std**: a standard ADMB output file, with the parameters estimated for that run and their estimated Hessian-based standard deviation (SD) included those for the projected years. Annual SSB, Btot, SSB/SSBmsy, F, F/Fmsy and MSY their estimated standard deviations are also included (see sample table below). The table below shows the list of output parameters in the case of fixed steepness. Parameters with (*) will provide annual figures. If you have 50 years data and 10 years projected years, you will get 60 data sets for the parameters with (*). Values and std dev show just some examples. With this information you can compute the Confidence Intervals (CI), e.g. **95% CI** by $1.96 \pm SD$. Please note that lnK and K (carrying capacity) will produce just one figure each for value and std dev.

	name	value	std dev
	lnK	1.42E+01	9.79E-02
*	Selpar	8.70E-04	1.01E-03
	K	1.47E+06	1.44E+05
*	Bsp	1.47E+06	1.44E+05
*	Btot	1.74E+06	1.71E+05
*	Fmort_overall	1.65E-04	1.98E-05
*	MSY	1.27E+05	1.16E+04
*	FoverFmsy	2.27E-04	2.73E-05
*	BoverBmsy	3.64E+00	1.70E-01

5. Another useful ADMB standard output files is **ASPM.cor** for the correlations of the parameter estimates.
6. **ASPM.hst**: a standard ADMB output file when MCMCs are run (see above).
7. **MCMC_B.csv**: SSB, Btot and SSB/SSBmsy for each year. If MCMC has been run, one vector for each iteration. Probability intervals for these parameters can then be computed, taken the burn-in period into account (see above).
8. **MCMC_F.csv**: F, F/Fmsy and MSY for each year. If MCMC has been run, one vector for each iteration. Probability intervals for these parameters can then be computed, taken the burn-in period into account (see above).

Figures below show some examples of the results obtained by MCMC run
(1950-2020 and 2011-2020 projection years)



6. SUGGESTIONS

- Always try different starting guesses to see if the program converges on the same solution.
- If the model does not converge, you can try restarting it with the parameter that have just been estimated, to do this type:

ASPM -ainp aspm.par

i.e., it will use the values in aspm.par as initial guesses.

- Since the input and output file names are fixed, it's easier to keep each run in a separate folder.

ACKNOWLEDGEMENTS

We sincerely thank to Fisheries Research Agency (FRA) of Japan to provide the fund for this ASPM software development project in 2008 and 2011.

REFERENCES

- Anonymous (2001) Report of the IOTC ad hoc working party on methods, Sète, France 23-27, April, 2001: 20pp.
- Beddington, J. R., and J. G. Cooke. 1983. The potential yield of fish stocks. FAO Fish. Tech. Pap. 242. 50pp.
- Beverton, R. J. H., and S. Holt. 1957. On the dynamics of exploited fish populations. Reprinted in 1993 by Chapman and Hall, London. 553 pp.
- Deriso, R. B., T. J. Quinn, and P. R. Neal. 1985. Catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. 42:815-824.
- Francis, R.I.C.C. 1992. Use of risk analysis to assess fishery management strategies: a case study using orange roughy (*Hoplostethus atlanticus*) on the Chatham Rise, New Zealand. Can. J. Fish. Aquat. Sci. 49: 922-930.
- Goodyear, C. P. 1993. Spawning stock biomass per recruit in fisheries management: foundation and current use. p.67-81 in S. J. Smith, J. J. Hunt, and D. Rivard (eds.) Risk evaluation and biological reference points for fisheries management. Can. Spec. Publ. Fish. Aquat. Sci. 120
- Hilborn, R. 1990. Estimating the parameters of full age structured models from catch and abundance data. Bull. Int. N. Pacific Fish Comm. 50: 207-213.
- ICCAT. 1997. Report for biennial period 1996-97. Part I (1996), Vol.2. Int. Int. Comm. Cons. Atl. Tunas. 204pp.
- IOTC. 2002-2008. Working documents in WPTT (Working party on Tropical Tuna) available in the IOTC web site at <http://www.iotc.org/>

- Nishida, T. 1995 Preliminary resource assessment of yellowfin tuna (*Thunnus albacares*) in the western Indian Ocean by the stock-fishery dynamic model. IPTP/Collective volume 6, 167-180.
- Myers, R.A., W. Blanchard, and K. R. Thompson. 1990. Summary of North Atlantic fish recruitment 1942-1987. Can. Tech. Rep. Fish. Aquat. Sci. 1743.108pp.
- Prager, M. H. 1994. A suite of extensions to a no equilibrium surplus-production model. Fish. Bull. 92:374-389.
- Punt, A. E. 1994. Assessments of the stocks of Cape hakes, *Merluccius* spp. Off South Africa. S. Afr. J. Mar. Sci. 14 : 159-186.
- Restrepo, V. 1997. A stochastic implementation of an Age-structured Production model (ICCAT/SCRS/97/59), 23pp. with Appendix
- Schaefer, M. B 1957. A study of dynamics of the fishery for yellowfin tuna in the eastern tropical Pacific Ocean. IATTC Bull. Vol. 2: 247-285.
- Seber, G.A.F., and C.J. Wild. 1989. Nonlinear Regression. J. Wiley, New York.
- Shepherd. J. G 1982. A versatile new stock-recruitment relationship for fisheries, and the construction of sustainable yield curves. J. Cons. Int. Explor. Mer 40 : 67-75.

APPENDIX A: FORMULATION OF THE ASPM

The deterministic formulation, for ease of presentation, precedes the formulation for the stochastic model. A Beverton and Holt (1957) type of stock recruitment relationship (SRR) is assumed here. Note, however, that other forms could be implemented following the same basic procedure outlined here.

Deterministic formulation

The deterministic model is essentially like that of (Punt 1994), which was based on ideas presented by Hilborn (1990). It consists of a forward population projection,

$$N_{1,t+1} = f(S_t) \quad \text{for age 1} \quad (1a)$$

$$N_{a+1,t+1} = N_{a,t} e^{-z_{a,t}} \quad \text{for other ages except the "plus" group, and} \quad (1b)$$

$$N_{p,t+1} = N_{p-1,t} e^{-z_{p-1,t}} + N_{p,t} e^{-z_{p,t}} \quad \text{for the plus group, } p, \quad (1c)$$

where $f(S)$ is a stock-recruitment function (explained below), a and t index age and year, and age 1 is, for simplicity, assumed here as the age of recruitment. Z denotes the total age and year-specific mortality rate, which is the sum of natural mortality (M_a , an assumed input value) and fishing mortality, F . In the (Restrepo *in press*) implementation, F is calculated based on total yields, weights at age ($\bar{W}_{a,t}$), and age –specific selectivities that are input and assumed exact, for up to five fisheries. This is accomplished by solving for the fishery-specific multipliers ($F_{g,t}$) of the input selectivities ($s_{g,a,t}$) that result in the observed yields (Y), given the estimates of stock sizes:

$$Y_{g,t} = \sum_{a=1}^p F_{g,t} s_{g,a,t} \bar{W}_{a,t} N_{a,t} U_{a,t} \quad \text{with} \quad (2)$$

$$U_{a,t} = \frac{\left[1 - e^{-\sum_g F_{g,t} s_{g,a,t} - M_a} \right]}{\sum_g F_{g,t} s_{g,a,t} + M_a}$$

Thus, the population projection is conditioned on known yields. The Beverton and Holt SRR can be described by the equation

$$R_{t+1} = f(S_t) = \frac{\alpha S_t}{\beta + S_t}, \quad (3)$$

where R is the number of recruits ($N_{1,t+1}$ in eq.1a) and S is the reproductive output, namely the product of numbers times maturity times fecundity, summed over all ages. For simplicity, we hereafter refer to S as “spawning biomass”, which is often used as a proxy for reproductive output.

Formulation (3) is not very desirable for estimation because starting values of the parameters α and β are not easy to guess. For this reason, the ASPM uses a different parameterization, following (Francis 1992). It consists of defining a “steepness” parameter, τ , which is the fraction of the virgin recruitment (R_0) that is expected when S has been reduced to 20% of its maximum (i.e., $R = \tau R_0$ when $S = \gamma/5$,

where γ is the virgin biomass). The SRR can thus be defined in terms of steepness and virgin biomass, two parameters that are somewhat easier to guess initial values. For a Beverton-Holt relationship, virgin biomass should generally be of similar magnitude to the largest observed yields, while steepness should fall somewhere between 0.2 and 1.0, with higher values indicating higher capacity for the population to compensate for losses in spawning biomass with increases in the survival of recruit. Nothing that equilibrium recruitment at virgin biomass can be computed as the ratio of virgin spawning biomass to spawning biomass per recruit in the absence of fishing $(S/R)_{F=0}$,

$$R_0 = \frac{\gamma}{(S/R)_{F=0}} \quad (4)$$

α and β are given by

$$\alpha = \frac{4\tau R_0}{5\tau - 1} \quad (5)$$

and

$$\beta = \frac{\gamma(1-\tau)}{5\tau - 1} \quad (6)$$

The spawning potential ratio, SPR , is measured by the spawning biomass per recruit obtained under a given F , divided by that under $F=0$ (Goodyear 1993). A useful benchmark for management is the SPR corresponding to the slope of the SRR at the origin, i.e., at the point when the stock is expected to “crash”. From equations (4) to (6) it follows that this SPR_{crash} is given by

$$SPR_{crash} = \frac{(S/R)_{crash}}{(S/R)_{F=0}} = \frac{\beta/\alpha}{\gamma/R_0} = \frac{1-\tau}{4\tau} \quad (7)$$

Hence, in a deterministic sense, any fishing mortality that results in an SPR lower than SPR_{crash} is not sustainable.

Fitting the model requires finding the values of the **SRR** parameters that best explain the trends in indices of abundance, given the observed yields and other inputs. For a set of initial conditions ($N_{a,t}$ for all ages in $t=1$), equations (1) and (3) are used to project the population forward, with the fishing mortalities being calculated conditional on observed yields, by equation (2). Values of the parameters γ and τ are chosen to minimize the negative log-likelihood,

$$-\ln(L_1) = \sum_i \left[\frac{n_i}{2} \sum \ln(\sigma_{i,t}^2) + \sum_t \frac{1}{2\sigma_{i,t}^2} (I_{i,t} - \hat{I}_{i,t})^2 \right] \quad (8)$$

where i denotes each available index. The last term is for the squared differences between observed and predicted indices (these could be in logarithmic units if a lognormal error is assumed), and $\sigma_{i,t}^2$ are variances whose computation is explained below. The predicted indices are obtained as the summation of stock sizes, times an input index selectivity, u , over all ages:

$$\hat{I}_{i,t} = q_i \sum_a N_{a,t} u_{a,i} \omega_i \quad (9)$$

where ω indicates some input control as to whether the index is in numbers or biomass (in which case the product being summed include weight at age), and whether computations are for the start or middle of the year. The parameters q_i scale each index to absolute population numbers (or biomass) and their maximum likelihood values can be obtained analytically by setting the derivative of equation (8) with respect to q_i equal to zero, and solving for the q_i .

There are several options for handling the variances, $\sigma_{i,t}^2$. If all the values for all indices are given equal weight, they can be set to

$$\sigma_{i,t}^2 = \sum_i \left[\frac{1}{n_i} \sum_t (I_{i,t} - \hat{I}_{i,t})^2 \right] \quad (10)$$

or, if all values within an index are to have equal weights but each index is weighted depending on how it is fitted by the model (maximum likelihood weighting) then:

$$\sigma_{i,t}^2 = \frac{l}{n_i} \sum_t (I_{i,t} - \hat{I}_{i,t})^2 \quad (11)$$

Alternatively, the variances could be input for each value, based on external information.

So far, the presentation of the method has indicated that parameters γ and τ (or, equivalently, α and β) are estimated directly in the search, and the parameters q_i and $\sigma_{i,t}^2$ are obtained indirectly or externally. The remaining requirement to complete the estimation procedure has to do with the initial conditions. This can be handled in various ways and perhaps the easiest is to assume that the initial age composition corresponds to inequilibrium one in virgin state. For this to be approximately valid, the time series of yield data should be extended as far back in time as possible, preferably to the onset of fishing. In this case,

$$N_{1,1} = R_0 \quad (12a)$$

$$N_{a,1} = N_{a-1,1} e^{-M_{a-1}} \quad \text{for ages } a = 2 \text{ to } p-1, \text{ and} \quad (12b)$$

$$N_{p,1} = \frac{N_{p-1,1} e^{-M_{p-1}}}{(1 - e^{-M_p})} \quad \text{for the plus group.} \quad (12c)$$

An alternative consists of estimating the equilibrium recruitment in year $t=1$ as an additional parameter and solving for the initial age composition that produces a spawning biomass that results in that recruitment given τ and γ . Several other options exist, but it appears that none will generally be superior unless there is adequate relative abundance information for the start of the time series. A useful option may be to “fix” the initial age composition at same scaled fraction of the virgin one, and to conduct sensitivity trials for that choice.

The computation of statistics such as maximum sustainable yield (MSY) and related benchmarks (e.g. S_{MSY} , F_{MSY}) is straightforward once the parameters for the SRR have been obtained. Shepherd (1982) describes the procedure used to compute equilibrium yield curves from a SRR , together with yield-per-recruit and spawning biomass-per-recruit calculations. Conditional on a given F (including an overall

selectivity pattern), equilibrium spawning biomass, recruitment and yield are computed as (for the Beverton and Holt SRR)

$$S_F = \alpha(S/R)_F - \beta \quad , \quad (13a)$$

$$R_F = \frac{S_F}{(S/R)_F} \quad , \text{ and} \quad (13b)$$

$$Y_F = R_F(Y/R)_F \quad (13c)$$

where $(S/R)_F$ and $(Y/R)_F$ are the spawning biomass and yield per recruit values resulting from exploitation at F . To search for MSY -related statistics, this procedure is built into an algorithm to obtain the desired target, e.g. to find the maximum Y_F as the estimates of MSY . Note that, if the selectivity pattern changes over time, then the computed MSY -related values will also change as a result of changes in the per-recruit computations.

Stochastic formulation

A stochastic ASPM requires that a recruitment value be estimated for every year. If this were attempted without constraints on the possible recruitment values, while simultaneously estimating the SRR, the application would be over-parameterized in most real situations. In this work, we have chosen to estimate the recruitments as lognormal deviations from the equilibrium SRR, assuming that these deviations follow a first-order autoregressive process.

The population projection equations are as in equation (1), except that recruitment is estimated as

$$N_{1,t} = R_0 e^v \quad (14)$$

That is, recruitment is estimated as deviations from a virgin level. Instead of estimating γ and τ directly as parameters, the model estimates γ and all the v_t . R_0 is computed from equation (4). These are essentially all parameters that would be needed to project the population forward and compute the log-likelihood in equation (8). The AR [1] process is incorporated by assuming that the recruitment estimates thus obtained vary around the expected stock recruitment relationship as

$$R_{t+1} = \frac{\alpha S_t}{\beta + S_t} e^{\varepsilon_{t+1}} \quad (15)$$

with $\varepsilon_{t+1} = \rho \varepsilon_t + \eta_{t+1}$, $|\rho| < 1$, the η have zero expectation and variance equal to σ_η^2 . In equations (14) and (15) we distinguish between recruitment values estimated as parameters ($N_{1,t}$) and those predicted from the estimated stock-recruitment relationship (R_t). The negative log-likelihood for these residuals would be (Seber and Wild 1989):

$$-\ln(L_2) = \frac{n_t}{2} \ln(\sigma_\eta^2) - \frac{1}{2} \ln(1 - \rho^2) + \frac{1}{2\sigma_\eta^2} \left[(1 - \rho^2) \varepsilon_1^2 + \sum_{t=2}^{n_t} (\varepsilon_t - \rho \varepsilon_{t-1})^2 \right] \quad (16)$$

where the residuals would be computed as

$$\varepsilon_{t+1} = \ln(N_{1,t+1}) - \ln(R_{t+1}) = \ln(N_{1,t+1}) - \ln\left(\frac{\alpha S_t}{\beta + S_t}\right) \quad (17)$$

Computation of the first residual would depend on the initial conditions. For example, in a virgin state, it would be

$$\varepsilon_1 = \ln(N_{1,1}) - \ln(R_0).$$

Note that α and β in equations (15) and (17) could be computed from knowledge of virgin biomass and steepness (see equations (5) and (6)). However, only the former is being estimated directly as a parameter. To include steepness as an additional parameter to be directly estimated by the search would confound the information contained in R_0 and γ (refer to equations (4), (5), and (6)). Our approach is to replace α and β in the *SRR* of equation (17) by a function of those parameters being estimated in the search, and steepness. From equations (5) and (6) it follows that

$$R_{t+1} = \left(\frac{4R_0 S_t \tau}{\tau(5S_t - \gamma) - S_t + \gamma} \right), \text{ such that} \quad (18)$$

$$\varepsilon_{t+1} = \ln(N_{1,t+1}) - \ln\left(\frac{4R_0 S_t \tau}{\tau(5S_t - \gamma) - S_t + \gamma}\right) \quad (19)$$

We take advantage of this relationship in order to solve for τ , noting that, for a given ρ and σ_η^2 , equation (16) will be at a minimum when

$$\sum_{t=2}^{n_t-1} \left[\ln(N_{1,t+1}) - \ln\left(\frac{4R_0 S_t \tau}{\tau(5S_t - \gamma) - S_t + \gamma}\right) - \rho \ln(N_{1,t}) + \rho \ln\left(\frac{4R_0 S_{t-1} \tau}{\tau(5S_{t-1} - \gamma) - S_{t-1} + \gamma}\right) \right]^2 \quad (20)$$

is also at a minimum. Thus, in every iteration in the search, a subprocedure is invoked to minimize (20) with respect to τ . Having thus calculated the steepness (and, consequently, α and β), the log-likelihood of equation (16) is added to the overall objective function.

It remains to be mentioned what to do about the parameters ρ and σ_η^2 . In theory, there is a potential for these to also be estimated. In practice, however, it is unlikely that data will contain so much information as to determine the relative contribution from recruitment variability with respect to the variability in the index values (see equations (8) and (16)). In our limited experience with this model, it appears that these values should be controlled by the analyst in much the same way as contributions to the likelihood from different data sources are weighted externally in other assessment methods (e.g., Deriso et al. 1985). Lower σ_η^2 values will result in lower stochasticity in recruitment, while higher σ_η^2 values will allow recruitment to fluctuate more widely in order to better fit the index data. A value of $\rho=0$ would assume no autocorrelation between successive recruitment deviations. Empirical studies such as those of Beddington and Cooke (1983) and Myers et al. (1990) may yield information about likely ranges of values for ρ and σ_η^2 for species groups. Reported values for these parameters (Myers et al. 1990) are quite variable across species.

Estimating the initial conditions for the stochastic model can be problematic, as with the deterministic model. Estimating the age structure in year 1 would not generally be an option as the model would easily become highly over-parameterized unless there were age-specific relative abundance data for the start of the series. Thus, using a long time series of data extending to the onset of fishing, and assuming an initial

equilibrium state at γ , remains a useful option. Other alternatives are also possible. In this paper we examine one in which we calculate a stable age structure (with only natural mortality) resulting from a pre-series recruitment that is fixed. That is, we fix $v_{t=0}$ and set the starting population sizes as

$$N_{2,1} = R_0 e^{v_0} e^{-M_1} \quad (21 \text{ a})$$

$$N_{a,1} = N_{a-1,1} e^{-M_{a-1}} \quad \text{for ages } a = 3 \text{ to } P-1, \text{ and} \quad (21 \text{ b})$$

the plus group is calculated as in equation (12c). This alternative allows the initial age structure to be either higher or lower than that corresponding to an equilibrium virgin state. The parameter $v_{t=0}$ could potentially be estimated in the search procedure as well. If it is, it may be desirable to place a penalty on how much it can alter the initial biomass, say, away from γ . This could be accomplished with the term

$$-\ln(L_3) = \frac{\ln(\sigma_v^2)}{2} + \frac{(\ln(S_1) - \ln(\gamma))^2}{2\sigma_v^2} \quad (22)$$

where σ_v^2 is a variance value to be fixed by the analyst.

Estimation of the stochastic model parameters for any given data set then requires several choices associated with how much recruitment can fluctuate around its deterministic predictions and about the initial conditions. In addition to choices about variances (σ_η^2 , σ_v^2 and possibly $\sigma_{i,l}^2$), the log-likelihood components could be given different emphases (λ) to obtain model estimates by minimizing:

$$-\ln(L_T) = -\ln(L_1) - \lambda_2 \ln(L_2) - \lambda_3 \ln(L_3) \quad (23)$$