PROCEAN : A PRODUCTION CATCH / EFFORT ANALYSIS FRAMEWORK TO ESTIMATE CATCHABILITY TRENDS AND FISHERY DYNAMICS IN A BAYESIAN CONTEXT.

Olivier MAURY

IRD (Institut de Recherche pour le Développement) UR THETIS

Centre de Recherche Halieutique Méditerranéenne et Tropicale

Avenue Jean Monnet - BP 171, 34203 Sète cedex

E-mail: olivier.maury@mpl.ird.fr

ABSTRACT

The model PROCEAN (PROduction Catch-Effort Analysis) is a bayesian statistical catch/effort analysis framework based on a generalized production model. The use of such a production model could be usefull in IOTC where reliable size data are missing for stock assessment.

The aim of this paper is to present the PROCEAN model. PROCEAN is a multi-fleet non equilibrium generalized production model which includes process error for both catchability time series and carrying capacity of the stock. PROCEAN assumes that fluctuations of the stock surface may only have consequences on fleets catchability and on the stock carrying capacity.

Our objective is not to propose a very realistic representation of the fishery. We propose a tool to extract the maximum amount of information from the data set by structuring it given a simple and well established theoretical model. Then, modeling is used here as a mean to explore data sets according to various hypothesis.

INTRODUCTION

For their flexibility, production models have proven for long to be very interesting tools for tuna stock assessment. But two main problems make difficult their use for investigating the status of the stock:

The first problem is the calculation of the effective fishing effort targeting a given stock. Indeed, many different fleets with heterogeneous and changing catchabilities (in general showing an increasing efficiency) are often exploiting the same population. Consequently, scientists in charge of tuna stock assessment are rarely in a position to estimate explicitely the effective fishing effort (*e.g.* an effort proportional to the fishing mortality).

The other important problem concerning the use of production models for tuna stock assessment is the fluctuation of the overall size of the exploited area. Indeed, it is well known for tuna fisheries that the estimated production curve and its associated MSY are closely linked to the exploited surface (Fonteneau and Marcille, 1988). To overide this problem, different production models incorporating the exploited surface have been proposed (Laloë, 1989 ; Die *et al*., 1990 ; Maury, 2000).

In this paper, we present the PROCEAN (PRoduction Catch / Effort ANalysis) modeling framework which adress both categories of problems in a bayesian context. PROCEAN is a multi-fleet non equilibrium generalized production model which includes process error for both catchability time series and carrying capacity of the stock. PROCEAN assumes that fluctuations of the stock surface may only have consequences on fleets catchability and on the stock carrying capacity.

Our objective is not to propose a very realistic representation of the fishery. We propose a tool to extract the maximum amount of information from the data set by structuring it given a simple and well established theoretical model. Then, modeling is used here as a mean to explore data sets according to various hypothesis.

THE BASIC MODEL

The PROCEAN model is based on the classic Pella and Tomlinson (1969) generalized production model which links the stock biomass B to the fishing mortality F by the mean of an ordinary differential equation continuous in time:

$$\frac{dB_t}{dt} = rB_t\left(1-\left(\frac{B_t}{K_t}\right)^{m-1}\right) - F_tB_t \qquad with \quad m>1 \qquad (1)$$

With B_t, the biomass at time t; F_t, the instantaneous fishing mortality rate; K, the carrying capacity of the stock; r, the per capita intrinsic growth rate of the population and m, the shape parameter (the model becomes a simple Schaefer model when $m=2$).

To introduce catches and effort for multiple fleets into the model, the fishing mortality F_t is expressed as the sum of each fleet's instantaneous fishing mortality:

$$\begin{cases} \dfrac{dB_t}{dt} = rB_t\left(1-\left(\dfrac{B_t}{K_t}\right)^{m-1}\right) - \sum_{i=1}^{n-1} q_{i,t} f_{i,t} B_t - C_{n,t} \\ \dfrac{dC_{i,t}}{dt} = q_{i,t} f_{i,t} B_t \quad 1 \le i < n \end{cases} \qquad (2)$$

With $n-1$, the total number of fleets for which fishing effort is available; $q_{i,t}$, the catchability coefficient for fleet i at time t, $f_{i,t}$ the mesured nominal fishing effort for fleet i at time t and $C_{i,t}$, the catches for fleet i at time t. $C_{n,t}$ represents the catches for all the fleets non documented in term of effort.

The biomass equation (equation 2) is an ordinary differential equation. It is integrated using a first order in time semi-implicit numerical approximation to have a better numerical stability than with a fully explicit scheme. This provides a time serie of predicted catches given a set of parameters (including biomass at time 0) :

$$\begin{cases} \dfrac{dB_t}{dt} = rB_t\left(1-\left(\dfrac{B_t}{K_t}\right)^{m-1}\right) - \sum_{i=1}^{n} q_{i,t} f_{i,t} B_t - C_{n,t} \\ \qquad \approx \dfrac{B_{t+dt} - B_t}{dt} = rB_t - rB_{t+dt}\left(\dfrac{B_t}{K_t}\right)^{m-1} - \sum_{i=1}^{n-1} q_{i,t} f_{i,t} B_{t+dt} - C_{n,t} \\ \dfrac{dC_{i,t}}{dt} = q_{i,t} f_{i,t} B_{t+dt} \quad 1 \le i < n \end{cases}$$

$$\Leftrightarrow \begin{cases} B_{t+dt} = \dfrac{B_t(1+rdt) - dtC_{n,t}}{1+\left(r\left(\dfrac{B_t}{K_t}\right)^{m-1} + \sum_{i=1}^{n} q_{i,t} f_{i,t}\right)dt} \\ C_{i,t+dt} = q_{i,t} f_{i,t} B_{i,t+dt} dt \quad 1 \le i < n \end{cases}$$

A STATISTICAL STRUCTURE FOR THE KEY PARAMETERS

To take into account potential fluctuations of the carrying capacity due to environmental fluctuations or to modifications of the fishery configuration such as stock surface (process errors), the parameter K is assumed to be dependent of time. We assume that the parameters $log(K_t)$ has the structure of a random walk which allows to model slow variations over time (Fournier, 1996) :

$$K_{t+1} = K_t e^{J_t - \frac{s_{q_i}^2}{2}} \qquad J \sim N(0, s_J)$$

The local catchability by fleet is also supposed to vary slowly each year to take into account potential fluctuations of fishing power for each fleet (process errors). We assume a random walk structure to the catchability time series for each fleet (Fournier *et al.*, 1998):

$$q_{i,t+1} = q_{i,t}.e^{e_{i,t} - \frac{s_{e_i}^2}{2}} \text{ with } e \in N(0, s_{e_i})$$

To address high-frequency variability of the catchability coefficient, a lognormal process-error structure is assumed for the fishing mortality. Then, the fishing mortality of fleet k at time t is written Concerning the catchability coefficient, the fishing mortality error structure is assumed to be lognormal. Then, the fishing mortality of fleet i in year t is written $F_{i,t} = q_{i,t} f_{i,t} e^{h_{i,t} - \frac{s_i^2}{2}}$ where the h_i are robustified normally distributed random variables with mean 0.

FITTING THE MODEL IN A BAYESIAN CONTEXT

To estimate the parameters in a bayesian context, we use the method of the maximum of posterior distribution (Bard, 1974) by maximizing the sum of the log-likelihood of the data plus the log of the prior density function. Then, given the data, the bayesian posterior distribution function for the model parameters has 4 components (one for the likelihood of the catch by fleet estimates L_C, one for the process errors for the carrying capacity L_K, one for the process errors concerning the effort-fishing mortality relationship L_q, and one for prior assumptions on the parameters r, m, and B_0).

Then, the posterior distribution is equal to L :

$$L = L_C \times L_K \times L_q \times L_{prior}$$

Catch component

We assume that the log of the predicted catches are the expected values of a random variable with a normal distribution:

$$L_C = \prod_{i=1}^{n-1}\prod_{t=0}^{t_{max}}\left[\frac{1}{C_{i,t}\, s_{C_i}\sqrt{2p}}\, e^{-\frac{\left(\log(C_{i,t})-\log(\hat{C}_{i,t})\right)^2}{2s_{C_i}^2}}\right]$$

With $\hat{C}$, the observed catches and C, the predicted catches.

Carrying capacity process error component

This component corresponds to the log-normal structured random walks for carrying capacity over time:

$$L_K = \prod_{t=0}^{t_{max}}\left[\frac{1}{s_J\sqrt{2p}}\, e^{-\frac{J^2}{2s_J^2}}\right]$$

Catchability process error component

This component combines the log-normal structured random walks for fishing power trends for each fleets and the effort/fishing mortality process error which has a robustified normal structure. This robustified normal distribution assumes a probability p for unlikely events (events which are more than e times the variance from the mean) and $1-p$ for the standard normal distribution (Fournier *et al.*, 1996) (Fig.1):

$$L_q = \prod_{i=1}^{n-1}\prod_{t=0}^{t_{max}} \left[\frac{1}{s_{e_t}\sqrt{2p}} e^{-\frac{e_{i,t}^2}{2s_{e_i}^2}} \right] \times$$

$$\prod_{i=1}^{n-1}\prod_{t=0}^{t_{max}} \left[(1-p)\left(\frac{1}{s_{h_i}\sqrt{2p}} e^{-\frac{h_{i,t}^2}{2s_{h_i}^2}}\right) + p\left(\frac{\sqrt{2}}{p s_{h_i} e \left(1+\frac{h_{i,t}^4}{\left(s_{h_i} e\right)^4}\right)}\right)\right]$$

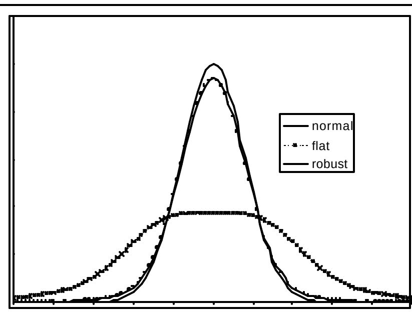

Fig. 1 : the robust distribution function used is a combination of a flat and a normal distribution function. Here, e=3 and p=0.1.

Priors and penalties

Informative priors can be added to the likelihood to take into account potential external informations concerning the parameters *r*, *m* and B_0. In the present version of the software, these three parameters are assumed to follow either a normal distribution, either a lognormal distribution either a beta distribution.

Estimating at the same time the variances for observation and process errors often lead to very unstable behaviors of the estimation process. In PROCEAN, only the standard errors s_C for the catches by fleet observation errors and the standard errors for the carrying capacity process errors s_q can be estimated simultaneously. The standard errors for the catchability process errors, s_e and s_h are considered to be proportional to s_C with fixed proportionality coefficients p :

$$\begin{cases} s_{e_i} = p_1 s_{C_i} \\ s_{h_i} = p_2 s_{C_i} \end{cases} \quad \forall i \in [1, n]$$

Thus, important additional information is provided through the use of the coefficients p which fix the strength of the constraints on the catchabilities variability.

The parameters of the model are estimated by finding the values of the parameters which minimize the opposite of *log(L)*. This minimization is performed with a quasi-Newton numerical function minimizer using exact derivatives with respect to the model parameters with the AD model builder software (ADMB © 1993-1996 by Otter Research Ltd). ADMB calculates the exact derivatives with a technique named automatic differentiation (Griewank and Corliss, 1991) and provides the variance of the parameter estimates by computing the Hessian matrix, ***H***, the elements of which are:

$$H_{i,j} = \frac{\partial^2(-\log L)}{\partial q_i . \partial q_j}$$

Where θ_i and θ_j are any two model parameters. Covariance matrix of the model parameters are estimated by computing the inverse of the hessian at the minimum.

CONCLUSION

The statistical structure of the model enables to compare different levels of statistical complexity. The simplest model should be first fitted to the data. Then, more and more complexity can be progressively added and tested.

Concerning the determination of priors, McAllister *et al*. (2000) proposes a method for estimating priors on *r* and *m* parameters based on life history traits. This method could be profitably used for tropical tunas keeping in mind that priors on parameters distribution may have important consequences on the estimated values of the parameters (McAllister and Kirkwood, 1998). This should be carefully studied and can be estimated by comparing prior distribution with posteriors empirical distributions of the parameters (McAllister and Ianelli, 1997; Punt and Hilborn, 1997).

Preliminary trials seems to indicate that the absolute values of the estimated parameters are sensitive to the priors used but the value of the MSY and the trends obtained for the carrying capacity, the fishing mortality and the catchability by fleet seems to be robust. If confirmed with simulation trials, this could indicate that the method proposed is adapted to study the catchability evolution by fleet due to technical progress or to changes in fishing strategy and tactics such as changes in targeting practices.

Finally, it should be kept in mind that the use of such a multifleet model depends on the catches and effort time series availability and is based on an adhoc fishing fleet stratification (gears and periods) which has to be determined.

REFERENCES

DIE D.J., RESTREPO V.R. AND W.W. FOX, JR. 1990. Equilibrium production models that incorporate fished area. Transaction of the American Fisheries Society. **119**: 445-454, 1990, pp:445-454.

BARD Y., 1974. Nonlinear parameter estimation. Academic Press, New York. 341pp.

FONTENEAU A. AND J. MARCILLE, 1988. *Ressources, pêche et biologie des thonidés tropicaux de l'Atlantique Centre-Est*. FAO document technique sur les pêches. Fonteneau A. Et Marcille J. (eds.). *FAO Tech. Doc.* 292. 391p.

FOURNIER D., 1996. An introduction to AD MODEL BUILDER for use in nonlinear modeling and statistics. Otter research Ltd.

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.

FOURNIER D.A., J.R. SIBERT, AND M. TERCEIRO, 1990. Analysis of length frequency samples with relative abundance data for the Gulf of Maine northern shrimp (*Pandalus borealis*) by the MULTIFAN method. *Can. J. Fish. Aquat Sci.* **48** : 591-598.

MAURY O., 2000. Multi-fleet non-equilibrium production models including stock surface to estimate catchability trends and fishery dynamics in a bayesian context. Application to the skipjack tuna's fishery (Katsuwonus pelamis) in the Atlantic Ocean. ICCAT meeting on assessment methods, Madrid 8-11/05/00. ICCAT SCRS/00/37.

MCALLISTER M.K. AND IANELLI J.N., 1997. Bayesian stock assessment using catch-age data and the sampling/importance resamplin g algorithm. *Can. J. Fish. Aquat Sci.* **54** : 284-300.

MCALLISTER M.K. AND KIRKWOOD G.P., 1998. Bayesian stock assessment: a review and example application using the logistic model. *ICES Jour. of Mar Sci.* **55**: 1031-1060.

GRIEWANK A. AND G.F. CORLISS, 1991. Automatic differientiation algorithms: theory, practice and application. *SIAM*, Philadelphia.

LALOE F. 1989. Un modèle global avec quantité de biomasse inaccessible dépendant de la surface de pêche. Application aux données de la pêche d'albacores (Thunnus albacares) de l'Atlantique est. *Aquat. Living Resour.*, **2**, 231-239.

PELLA J.J. AND P.K. TOMLINSON, 1969. A generalized stock production model. Bull. Inter. Am. Trop. Tuna Com., **13**: 420-496.

PUNT A.E. AND HILBORN R., 1997. Fisheries stock assessment and decision analysis: the bayesian approach. *Rev. of Fish Biol. and Fisheries* **7**: 35-63.