

FASST: A FULLY AGE-SIZE AND SPACE-TIME STRUCTURED STATISTICAL MODEL FOR THE ASSESSMENT OF TUNA POPULATIONS

Olivier Maury¹, Blaise Faugeras¹ and Victor Restrepo²

SUMMARY

This paper describes the model FASST. FASST is a fully structured age, size, space and time stock assessment model. It is expressed as a deterministic system of partial differential equations continuous in age, size and time. Space is represented using large discrete zones interconnected by exchange rates. All the parameterizations used to represent recruitment, growth, mortality and movements are chosen to be size-dependent and age independent. Consequently, for simplicity purposes, the age dimension can be simplified. The model is integrated numerically to simulate the stock dynamics. A complex process -and observation-error framework is used to estimate the model parameters, and hence the past stock status, in a Bayesian context. The model uses simultaneously most fishery data currently available for tuna stock assessment such as catch by fleet in number and weight, length-frequency samples, otolith analysis, tag-recapture data. It provides a new way to integrate that information into a consistent stock assessment framework.

RÉSUMÉ

Le présent document décrit le modèle FASST. FASST est un modèle d'évaluation du stock totalement structuré par âge, taille, espace et temps. Il est exprimé comme un système déterministe d'équations différentielles partielles continues en âge, en taille et en temps. L'espace est représenté à l'aide de grandes zones hétérogènes interconnectées par des taux d'échange. Tous les paramètres utilisés pour représenter le recrutement, la croissance, la mortalité et les déplacements sont choisis de sorte à être dépendants de la taille et indépendants de l'âge. Par conséquent, aux fins de simplicité, la dimension par âge peut être simplifiée. Le modèle est intégré numériquement afin de simuler la dynamique du stock. Un cadre complexe d'estimation d'erreurs de processus et d'observation est utilisé afin d'estimer les paramètres du modèle et donc l'état passé du stock, dans un contexte bayésien. Le modèle utilise simultanément la plupart des données sur les pêcheries actuellement disponibles pour l'évaluation des stocks de thonidés, telles que la prise par flottille en nombre et en poids, les échantillons de fréquence de taille, les analyses des otolithes et les données de marquage-récupération. Il fournit une nouvelle façon d'intégrer cette information dans un cadre cohérent d'évaluation du stock.

RESUMEN

Este documento describe el modelo FASST. El FASST es un modelo de evaluación de stock estructurado por edad, talla, espacio y tiempo. Se expresa como un sistema determinista de ecuaciones diferenciales parciales continuas en edad, talla y tiempo. El espacio se representa utilizando grandes zonas independientes interconectadas por tasas de intercambio. Todas las parametrizaciones utilizadas para representar el reclutamiento, el crecimiento, la mortalidad y los movimientos se eligen para que sean dependientes de la talla e independientes de la edad. Por consiguiente, para mayor sencillez, la dimensión edad puede ser simplificada. El modelo se integra numéricamente para simular la dinámica del stock. Se utiliza un marco complejo de estimación de errores de proceso y observación para estimar los parámetros del modelo, y por tanto, el estatus pasado del stock en un contexto Bayesiano. El modelo utiliza simultáneamente la mayoría de los datos de pesquerías disponibles actualmente para la evaluación de stocks de túnidos, como la captura por flota en números y peso, las muestras

¹IRD, CRHMT, av. Jean Monnet, B.P. 171, 34203 Sète cedex, FRANCE; tel: +33 (0) 499 57 32 28 fax: +33 (0) 499 57 32 95; Olivier.Maury@ird.fr

²ICCAT Corazon de Maria n°8 Madrid Espana.

de frecuencias de talla, análisis de otolitos, y datos de marcado-recaptura. Proporciona una nueva forma de integrar esta información en un marco coherente de evaluación de stock.

KEYWORDS

Stock-assessment, Population dynamic, Recruitment, Tuna fisheries, Fishery statistics, Fishing effort, Size distribution, Tagging, Growth curves, Migrations, Otoliths, Catch-effort, Catchability, Fishing mortality, Gear selectivity, Stochastic models, Mathematical models, Simulation, Body Size

1 Introduction

The goal of this paper is to present the FASST model (Fully Age-Size and Space-Time structured statistical model). FASST is a new model for tuna stock assessment which extends to a complex size-structured model the MULTIFAN-CL methodology presented in Fournier *et al.* (1998). Primarily designed for representing Atlantic bigeye tuna fisheries, the model is designed to be generic enough to be used for other species/regions.

Tuna populations and their associated fisheries exhibit very specific characteristics which have to be taken into account to model realistically their dynamics:

- Tuna fisheries are highly heterogeneous in space and time and such heterogeneity has a high functional importance in their functioning. The mixing rate of fish between different regions is generally not total. Then, the stock biomass located in a given fishing area interacts more or less strongly with the biomass located in other areas or with an unavailable part of the population located outside the fishing area. Important migrations and movements of fishes occurs at various scales so that the different regions interact differently: Fish movements must be explicitly represented either with a continuous model (such as an advection-diffusion-reaction model) either with a bulk transfer rate compartment model;
- Recruitment is continuous in time but highly heterogeneous in space and time (Cayré 1986) so that singular annual cohorts are most often difficult to identify. It follows that either continuous recruitment either a discretization of it such as the concept of mensual micro-cohorts (Fonteneau 1998) should be used;
- Non uniform mortality over sizes (due to both size dependant natural mortality and selectivities) may lead to non Gaussian distribution of size at age. Hence, simply adding a gaussian size distribution to an age structured model could lead to biases on both growth and mortality estimates (**Figure 1**): An explicit age and size structured model should be tested.
- Growth is potentially variable in space (especially for skipjack) depending of the region considered (Bard, 1986) so that fishes of the same age may exhibit very different sizes depending of their various history: A spatialized approach taking explicitly into account the potential variability of growth in space should be used;
- Many different fleets with different selectivity patterns and heterogeneous and changing catchabilities (in general showing an increasing efficiency) due to technical progress or to changes in fishing strategies and tactics fish the same population so that the use of fishing effort is problematic. A process error structure can profitably be added to fishing effort/mortality relationship and catchability time series (Fournier *et al.* 1998; Maury 2001; Maunder and Watter 2003).
- The uncertainty of quantities needed for management needs to be estimated so that a statistical approach must be used (Hilborn and Walters 1992).

Conducting reliable stock assessments taking into account those problems requires the development of a complex spatially explicit model dealing with all the fleets present in the fishery and with the movements of fish.

Such a stock assessment tool should profitably integrate a deterministic modelling of major processes with a statistical structure for both observation and process components (Fournier *et al.* 1990 and 1998). Space may be either considered using continuous representations such as advection-diffusion-reaction models (Bertignac *et al.* 1998; Sibert *et al.* 1999; Maury *et al.* 1999) either using discrete compartment models with fish transfer rates

between zones (Fournier *et al.* 1998). Statistical models provide confidence intervals for the estimated parameters and inferred management policies (Hilborn and Walters 1992; Fournier *et al.* 1998; Maunder and Watters 2003). The parameter estimation may be conducted using maximum likelihood methods in a Bayesian framework to easily handle the uncertainty concerning reference points useful for management (McAllister and Ianelli 1997; Punt and Hilborn 1997).

FASST is an age-size and time-space structured synthetic model of tuna population dynamics designed to use simultaneously all the information available (catches, effort, size frequencies, tagging³, otolith increments, length/weight relationships, ...) in a Bayesian framework to estimate the posterior probability distribution of the parameters (recruitment, growth, movement, natural mortality and catchability). Such statistical approach has already proven to be very powerful (Fournier *et al.* 1990 and 1998; Maunder and Watters 2003) and should be an important improvement for tuna stock assessment in ICCAT.

2 Model structure

2.1 Notations

2.1.1 Variables

N , the fish number

C , the catches

R , the recruitment

f , the fishing effort

q , the catchability

F , the fishing mortality

γ , the growth rate

D , the dispersion parameter of size

Q , the proportion of fish in a given data set lying in a given length interval

$matur$ indicating maturity

Θ , the percentage of females in a size class

W , the body weight

2.1.2 Subscripts

t , the time in months

a , the age in months

l , the length in cm

j , the zone j

k , the fleet k

α , indicates the length frequency data set

g , indicates the tag group (all the fish tagged the same month)

o , concerns otolith

2.1.3 Parameters

A , the maximal age

n , the total number of zones

T , the fish transfer rates between zones

M , the natural mortality

P , the fishing power

s , the selectivity

L_∞, K, a_0, X , the generalized von Bertalanfy growth parameters

β, χ, δ , the original von Bertalanfy growth parameters

³ The reliable parameterization of such complex spatially explicit statistical models has proven to very difficult without using tagging data (Hilborn 1990; Fournier *et al.* 1998; Sibert *et al.* 1999) which bring to the model an important amount of information concerning growth and length-specific natural mortality (Hampton 2000), catchabilities and selectivities (Anganuzzi *et al.* 1994) and age-dependent movements. Such reliable parameterization should be profitably conducted by using simultaneously fishery data and tagging data and potential auxiliary information for integrating the fishing effort standardization into the stock assessment model.

θ, m , the parameters for the length/weight relationship
 i , relates length at maturity to asymptotic length
 E , is the total number of length-frequency samples
 \mathcal{Q} , is the total number of otolith samples
 σ , the standard errors

2.1.4 Random variables

ε , the catchability random walk error
 η , the fishing mortality/ effort error
 μ , the selectivity random walk error
 ω , the natural mortality random walk error

2.2 Population dynamics

The dynamics of the fish population is described through the use of fish density functions $n_j(l, t, a)$ which j being the indice for spatial zones, $t \in (0, T)$ being the time dimension, $a \in (0, A)$ being the age dimension, and $l \in (0, L)$ the length dimension. The number of fish N of age between a_1 and a_2 , length between l_1 and l_2 at time

t in zone i is given by the integral:
$$\int_{a_1}^{a_2} \int_{l_1}^{l_2} n_i(l, t, a) dl da .$$

A compartment model which takes explicitly into account movement between homogeneous zones, growth in length (and potential growth spatial variability), natural and fishing mortality is used for population dynamics. It is based on the following system of partial differential equations:

$$\left\{ \begin{array}{l} \frac{\partial n_1(l, t, a)}{\partial t} + \frac{\partial n_1(l, t, a)}{\partial a} = \sum_{i=2}^n T_{i \rightarrow 1}(l, t) n_i(l, t, a) - \left(F_1(l, t) + M(l) + \sum_{i=2}^n T_{1 \rightarrow i}(l, t) \right) n_1(l, t, a) \\ \quad - \frac{\partial(\gamma_1(l, t) n_1(l, t, a))}{\partial l} + \frac{\partial \left(D_1(l, t) \frac{\partial n_1(l, t, a)}{\partial l} \right)}{\partial l} \\ \quad \vdots \\ \frac{\partial n_n(l, t, a)}{\partial t} + \frac{\partial n_n(l, t, a)}{\partial a} = \sum_{i=1}^{n-1} T_{i \rightarrow n}(l, t) n_i(l, t, a) - \left(F_n(l, t) + M(l) + \sum_{i=1}^{n-1} T_{n \rightarrow i}(l, t) \right) n_n(l, t, a) \\ \quad - \frac{\partial(\gamma_n(l, t) n_n(l, t, a))}{\partial l} + \frac{\partial \left(D_n(l, t) \frac{\partial n_n(l, t, a)}{\partial l} \right)}{\partial l} \\ \frac{dc_{1,k}(l, t, a)}{dt} = q_{1,k}(l, t) f_{1,k}(t) n_1(l, t, a) \\ \quad \vdots \\ \frac{dc_{n,k}(l, t, a)}{dt} = q_{n,k}(l, t) f_{n,k}(t) n_n(l, t, a) \end{array} \right. \quad (1.)$$

with $T_{i \rightarrow j}$, the transfer rate from zone i to zone j , q_i the catchability coefficient in zone i , M the natural mortality rate and $f_{i,k}$ the fishing effort exerted by fleet k in zone i .

2.3 Simplifying the model

In the model, according to basic ecological theory, all the parameterizations used to represent recruitment, growth, mortality and movements are chosen to be size-dependant and age-independent. Consequently, for simplicity purposes, the age dimension can be simplified without losing anything from the complexity of the

full model dynamics. With this simplification, the number of fish N of length between l_1 and l_2 at time t in zone i is given by the integral: $\int_{l_1}^{l_2} n_i(l, t) dl$.

The simplified model is written as follows:

$$\left\{ \begin{array}{l} \frac{\partial n_1(l, t)}{\partial t} = \sum_{i=2}^n T_{i \rightarrow 1}(l, t) n_i(l, t) - \left(F_1(l, t) + M(l) + \sum_{i=2}^n T_{1 \rightarrow i}(l, t) \right) n_1(l, t) - \frac{\partial(\gamma_1(l, t) n_1(l, t))}{\partial l} + \frac{\partial \left(D_1(l, t) \frac{\partial n_1(l, t)}{\partial l} \right)}{\partial l} \\ \vdots \\ \frac{\partial n_n(l, t)}{\partial t} = \sum_{i=1}^{n-1} T_{i \rightarrow n}(l, t) n_i(l, t) - \left(F_n(l, t) + M(l) + \sum_{i=1}^{n-1} T_{n \rightarrow i}(l, t) \right) n_n(l, t) - \frac{\partial(\gamma_n(l, t) n_n(l, t))}{\partial l} + \frac{\partial \left(D_n(l, t) \frac{\partial n_n(l, t)}{\partial l} \right)}{\partial l} \\ \left\{ \begin{array}{l} \frac{dc_{1,k}(l, t)}{dt} = q_{1,k}(l, t) f_{1,k}(t) n_1(l, t) \\ \vdots \\ \frac{dc_{n,k}(l, t)}{dt} = q_{n,k}(l, t) f_{n,k}(t) n_n(l, t) \end{array} \right. \end{array} \right. \quad (2.)$$

with $T_{i \rightarrow j}$, the transfer rate from zone i to zone j , q_i the catchability coefficient in zone i , M the natural mortality rate and $f_{i,k}$ the fishing effort exerted by fleet k in zone i .

2.4 Spatial structure of the model

The geographical structure of the model is based on the heterogeneity of the fishery and the data available (**Figure 2**). Considering fleets to be zone-specific enables the model equations to be fully nested if zones are grouped. Then, different level of resolution (from a model with no spatial structure to a model with a fully detailed spatial structure) can be compared by using usual statistical tests (AIC, PBF, likelihood ratios) to select the resolution according to the amount of information contained into the data.

2.5 Recruitment and boundary conditions

Homogeneous Neumann boundary conditions at $l=0$ and $l=L$ express the fact that the length of individuals can not reach negative values or values larger than L :

$$\frac{\partial n_j(0, t, a)}{\partial l} = \frac{\partial n_j(L, t, a)}{\partial l} = 0$$

The initial size distribution is fixed:

$$n_j(l, 0, a) = n_j^0(l, a)$$

We also need a boundary condition for $a=0$. In the model, the only source of individuals is the recruitment which corresponds to the entrance of age 0 (in the full model) or size 0 (in the simplified model) individuals into the system.

2.6 Numerical approximation

The systems (1) and (2) are integrated numerically on a monthly time basis. The method of the characteristics is used for the full model (1). Explicit approximations are used for the growth, movements, mortality and catch components with a time splitting method using the Lie formula (Marchuk 1990). A 2nd order MUSCL (Monotone Upstream-centered Schemes for Conservation Laws) explicit numerical scheme is used for the growth component. Such a MUSCL scheme has been chosen because it ensures the positiveness of the solutions, it is conservative, very stable and has a much lower numerical dispersivity than simpler difference schemes.

In the full model (1), a closed boundary is used for the age dimension at age $a=A$, the maximal age which accumulates old fishes (A is a “+” group).

The recruitment do correspond with the boundary condition in age 0 and length 0. It is assumed to vary in each zone. A Beverton and Holt type stock-recruitment relationship is used in each zone with a lognormal process error allowing for random fluctuations of recruitment around its mean:

$$N_{j,l,t,a=0} = R_{j,t} = \frac{\Psi SSB_{j,t}}{\zeta_j + SSB_{j,t}} e^{\nu_{j,t} - \frac{\sigma_\nu^2}{2}} \quad \text{with} \quad SSB_{j,t} \propto \int_{a=0}^A \int_{l_{matur}}^{l_{max}} \Theta_l W_l n_j(l,t,a) dl da \quad (3.)$$

Bigeye are assumed to be recruited when their fork length reach 30cm. In the full model, between age 0 and the age corresponding to the recruitment, the growth and movement processes are applied without mortality in order to obtain a realistic distribution of size at age when the micro-cohorts enter the fishery.

Both models (1) and (2) are integrated. The simplified model (2) is used during the parameter estimation process to decrease calculation time and memory requirements. The full model (1) is used only for simulations, when the parameters have been estimated with (2) and when both age and size distributions are useful for interpreting the results.

3 Parameterizations

3.1 Growth

We consider that, in a given zone, fish growth follows a generalized von Bertalanffy curve which allows for slower growth rates for young fish as they are frequently observed for tunas, especially for yellowfin tuna (Pauly and Moreau 1997; Gascuel *et al.* 1992). From a process point of view⁴, this growth curve is based on a length-dependent growth rate, independent of age:

$$\frac{\partial w}{\partial t} = 3l^2 \frac{\partial l}{\partial t} = \text{anabolism} - \text{catabolism} = 3\beta_{j,t} l^{\chi_{j,t}} - 3\delta_{j,t} l^{m_{j,t}} \quad \left| \begin{array}{l} \chi \approx 2 \\ m \approx 3 \end{array} \right. \quad (4.)$$

$$\gamma_{j,l,t} = \frac{\partial l}{\partial t} = \beta_{j,t} l^{\chi_{j,t}-2} - \delta_{j,t} l^{m_{j,t}-2} \approx \beta_{j,t} - \delta_{j,t} l$$

The parameter m can be estimated externally with a length-weight relationship. In the general case where $\chi=2$ and $m=3$, equation (3) reduces to a von Bertalanffy growth equation with $L_\infty = \beta_{j,t} / \delta_{j,t}$ and $K = \delta_{j,t}$.

According with that formulation, age a fish belonging to the same cohort will have a different growth rate if they have a different size. This will cause a non-gaussian distribution of size frequencies for a given age⁵, even if mortality applies homogeneously over sizes.

Different assumptions can be made about the dispersion rate over the length dimension. They will have implications on the standard deviations of the fish length at age distributions. The *a priori* more realistic assumption is a dispersion coefficient being linearly related to growth rate:

$$D_{j,l,a,t} = u\gamma_{j,l,a,t} + v$$

which also allows for growth-independent dispersion rate when $v=0$.

3.2 Mortality

In the model, the natural mortality M is length-dependant (Hampton 2000). Its structure is assumed to be a random walk with respect to length:

$$M_{l+1} = M_l e^{\omega_l - \frac{\sigma_\omega^2}{2}} \quad \omega \sim N(0, \sigma_\omega)$$

⁴ From a processes point of view, growth results from the antagonism between anabolism proportional to a physiological surface and catabolism proportional to fish weight.

⁵ If a more standard analysis based on the usual generalized von Bertalanffy growth curve $\left(l_a = L_\infty \left(1 - e^{-KX(a-a_0)} \right)^{\frac{1}{X}} \right)$ is desired in the full model, the growth rate may be assumed to be age-dependent and written as follows:

$$\gamma_{j,a,t} = \frac{\partial l}{\partial a} = L_{\infty,j,t} K_{j,t} e^{-K_{j,t} X(a-a_0)} \left(1 - e^{-K_{j,t} X(a-a_0)} \right)^{\frac{1-X}{X}}$$

With homogeneous mortality over sizes, this will lead to a gaussian distribution of size frequencies for a given age.

The fishing mortality F is also length-dependent and defined as the sum of the fishing mortality of the n fleets k :

$F_{j,l,t}^* = \sum_{k=1}^n F_{k,j,l,t}^*$ with $F_{k,j,l,t}^* = q_{k,j,l,t} f_{k,j,t}$ (here, the * indicates that the statistical part of the fishing mortality has not yet been introduced).

According with the separability assumption, the catchability for each fleet is split into a length component (varying slowly over time), the selectivity s , and two time-varying components, the fleet fishing power p and the fish disponibility d for a given fleet:

$$F_{k,j,l,t}^* = \sum_k q_{k,j,l,t} f_{k,j,t} = \sum_k s_{k,l,t} p_{k,t} d_{k,j,t} f_{k,j,t}$$

To take into account a potential seasonality due to fish behavior (reproductive concentrations, ...), the disponibility d may be allowed to have a sinusoidal structure :

$$d_{k,j,t} = d_{k,j}^r + d_{k,j}^* \left(\zeta_{k,j} + \sin \left(\frac{2\pi}{\tau_{k,j}} + \phi_{d,k,j} \right) \right)^+ \quad \begin{cases} (d_{k,j}^r, d_{k,j}^*) \in R^{+2} \\ \zeta \in [-1;1] \\ \phi \in [0;2\pi] \end{cases} \quad \begin{cases} \tau \in [0;12] \\ \phi \in [0;2\pi] \end{cases}$$

With the positive part of a real number noted $()^+$.

Catchability may also be related to external factors such as environmental factors (thermocline depth, ...) with a deterministic function.

Stochasticity is added to the deterministic F^* parameterization to take into account time variability of the key parameters due to non-explicit external processes. A random walk structure is assumed for selectivity with respect to length:

$$s_{k,l+1,t} = s_{k,l,t} e^{\mu_{k,l,t} - \frac{\sigma_\mu^2}{2}} \quad \mu \sim N(0, \sigma_\mu)$$

Selectivities by size may be allowed to vary in time on a slow time basis (e.g. every 5 years) for each fleet.

To account for potential fluctuations of fishing power, we assume a random walk structure for the annual fishing power time series for each fleet. Such trends in fishing power may be due to changes in targeting or due to technological progress for instance. A normal random walk structure for $\log(p_t)$ allows the catchability to vary slowly over time without *a priori* assumption on its trend (increase or decrease) (Fournier *et al.*, 1998; Maury, 2001). For that purpose, we assume that the yearly variability of p has the following structure:

$$p_{k,t+1} = p_{k,t} e^{\varepsilon_{k,t} - \frac{\sigma_\varepsilon^2}{2}} \quad \varepsilon \sim N(0, \sigma_\varepsilon)$$

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

$F_{k,j,l,t} = F_{k,j,l,t}^* e^{\eta_{k,j,t} - \frac{\sigma_\eta^2}{2}}$ where the $\eta_{k,j,t}$ are robustified normally-distributed random variables with mean 0 (Fournier *et al.*, 1998).

3.3 Movements

Fish movement rates are restricted to occur only between spatially adjacent zones. They are supposed to have two components. A diffusive one which does not vary in time for a given adjacent pair of zones and an advective one which has a truncated sinusoidal structure which enables for seasonal advective movements. To account for length-dependent variation of fish speed, movement rates are supposed to be proportional to a power of fish length:

$$\begin{cases} T_{i \rightarrow j,l,t} = \left(T_{i \rightarrow j,l}^r + T_{i \rightarrow j,l}^* \left(Y_{i \rightarrow j,l} + \sin \left(\frac{2\pi}{\Lambda} + \phi_{T,i \rightarrow j,l} \right) \right) \right)^+ \\ T_{j \rightarrow i,l,t} = \left(T_{j \rightarrow i,l}^r + T_{j \rightarrow i,l}^* \left(Y_{i \rightarrow j,l} + \sin \left(\frac{2\pi}{\Lambda} + \phi_{T,i \rightarrow j,l} \right) \right) \right)^- \end{cases} \begin{cases} (T_{i \rightarrow j,l}^r, T_{i \rightarrow j,l}^*) \in R^{+2} \\ Y \in [-1;1] \end{cases} \begin{cases} \Lambda \in [0;12] \\ \phi \in [0;2\pi] \end{cases}$$

Because movement patterns are likely to be different for adults which may experience reproductive migration and stay more in equatorial reproductive waters than juveniles which exhibit home range movements, two sets of movement parameters will be estimated and tested: a set for juveniles whose length is smaller than the length at first maturity l_{matur} and a set for adults whose length is greater. Length at first maturity is defined as being equals to a fixed fraction of the asymptotic length L_∞ :

$$l_{matur} = iL_\infty = i \left(\frac{\beta}{\delta} \right)^{\frac{1}{m-x}} \quad i \in [0;1[$$

3.4 Tagging

Tagged fish are modelled by using exactly the same population dynamics equations and parameters that for non tagged fish. Additional mortality, tag shedding and tag reporting rates can be added if some information is available about it.

4 Fitting the model in a Bayesian context

To estimate the parameters in a Bayesian context, we will 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 8 components (one for the log-likelihood of the catch by fleet estimates L_C , one for the length frequency distributions of catches L_l , one for the age-length relationship from otolith reading L_o , one for the catchability statistical structure L_q , one for the recruitment deviations L_R , one for the tag recovery estimates L_{tag} , one for the natural mortality random walk over length L_M , and one for priors and penalties L_{penal}).

Then, the posterior distribution is equal to L :

$$L = L_C \times L_l \times L_o \times L_q \times L_M \times L_R \times L_{tag} \times L_{penal}$$

The parameters of the model will be estimated by finding the values of the parameters which minimize the negative log of this posterior distribution function. This minimization will be 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).

4.1 Catches component

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

$$L_C = \prod_{j=1}^n \prod_{k=1}^m \prod_{l=0}^{l_{\max}} \prod_{t=0}^{t_{\max}} \left[\frac{1}{C_{j,k,l,t} \sigma_{C_k} \sqrt{2\pi}} e^{-\frac{(\log(C_{j,k,l,t}) - \log(\hat{C}_{j,k,l,t}))^2}{2\sigma_{C_k}^2}} \right]$$

With \hat{C} , the observed catches and C , the predicted catches. Important additional information may be provided by fixing the variances σ by fleet which fix the weights of the corresponding likelihood components.

4.2 Length-frequencies component

The robust normal likelihood of Fournier et al. (1990 and 1998) modified by Maunder and Watters (2000) is used for the contribution of the length frequency distribution to the likelihood:

$$L_l = \prod_{\alpha=1}^E \prod_{l=0}^{l_{\max}} \left[\frac{1}{\tau_{\alpha} \sqrt{2\pi \left(\xi_{\alpha,l} + \frac{1}{N_l} \right)}} e^{-\frac{(Q_{\alpha,l}^{obs} - Q_{\alpha,l}^{pred})^2}{2 \left(\xi_{\alpha,l} + \frac{1}{N_l} \right) \tau_{\alpha}^2}} + 0.01 \right]$$

With $\xi_{\alpha,l} = Q_{\alpha,l}^{obs} (1 - Q_{\alpha,l}^{obs})$ and $\tau_{\alpha}^2 = \frac{\delta}{\min(S_{\alpha}, 1000)}$ with S_{α} the sample size. N_l is the number of length intervals in the sample.

4.3 Age-length relationship from otoliths component

With a simple assumption of gaussian distribution of length measurement errors (and no error on age measures) the otoliths component is written as follows:

$$L_o = \prod_{o=1}^{\Omega} \left[\frac{1}{\sigma_o \sqrt{2\pi}} e^{-\frac{(\hat{l}(a_o) - l(a_o))^2}{2\sigma_o^2}} \right]$$

In the case where parameters can be estimated with the full model, a multinomial likelihood function can be used to model the error between predicted von Bertalanffy length at age and length at age derived from otolith increments counting (preliminary trials indicate that this robust function gives the same results that a likelihood assuming normal error for age estimations but brings information about potential spatial variability of growth).

$$L_o = \prod_{o=1}^{\Omega} \frac{C_{z(o),f(o),l(o),t(o),a(o)}}{\sum_{a=0}^A C_{z(o),f(o),l(o),t(o),a}}$$

4.4 Catchability structure components

This component combines the log-normal structured random walks for selectivities and 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):

$$L_q = \prod_{k=1}^{nfleet} \prod_{t=0}^{t_{\max}} \left[\frac{1}{\sigma_{\varepsilon_k} \sqrt{2\pi}} e^{-\frac{\varepsilon_{k,t}^2}{2\sigma_{\varepsilon_k}^2}} \right] \times \prod_{k=1}^{nfleet} \prod_{t=0}^{t_{\max}} \prod_{l=0}^{l_{\max}} \left[\frac{1}{\sigma_{\mu_k} \sqrt{2\pi}} e^{-\frac{\mu_{k,t,l}^2}{2\sigma_{\mu_k}^2}} \right]$$

$$\times \prod_{k=1}^{nfleet} \prod_{t=0}^{t_{\max}} \left[(1-p) \left(\frac{1}{\sigma_{\eta_k} \sqrt{2\pi}} e^{-\frac{\eta_{k,t}^2}{2\sigma_{\eta_k}^2}} \right) + p \left(\frac{\sqrt{2}}{\pi \sigma_{\eta_k} e \left(1 + \frac{\eta_{k,t}^4}{(\sigma_{\eta_k} e)^4} \right)} \right) \right]$$

4.5 Natural mortality random walk component

This component corresponds to the log-normal structured random walks for natural mortality over lengths:

$$L_M = \prod_{l=0}^{l_{\max}} \left[\frac{1}{\sigma_{\omega} \sqrt{2\pi}} e^{-\frac{\omega_l^2}{2\sigma_{\omega}^2}} \right]$$

4.6 Recruitment deviations components

This component assumes a log-normal recruitment deviation from its mean in each zone and time strata:

$$L_R = \prod_{j=1}^n \prod_{t=0}^{t_{\max}} \left[\frac{1}{\sigma_v \sqrt{2\pi}} e^{-\frac{v_{j,t}^2}{2\sigma_v^2}} \right]$$

4.7 Tagging data component

Observed numbers of tag returns are related to predicted numbers of tag returns by a Poisson likelihood function (Sibert *et al.* 1999):

$$L_{tag} = \prod_{g=1}^G \prod_{l=1}^{l_{\max}} \prod_{j=1}^n \prod_{t=1}^{t_{\max}} \left(\frac{\hat{C}_{g,j,t}^{TC} e^{-\hat{C}_{g,j,t}^{TC}}}{C_{g,j,t}^{TC} !} \right)$$

4.8 Priors and penalties

Priors are added to the likelihood to take into account potential external information. A prior assumption based on the literature (*e.g.* Hampton 2000) will be made for the parameters M_l . Another based on fishbase (Froese and Pauly 1997) data will be developed for the parameter i relating length at first maturity to asymptotic length.

$$L_{penal} = \prod_{l=1}^{l_{\max}} \left[\frac{1}{\sigma_M \sqrt{2\pi}} e^{-\frac{\left(\log\left(\frac{M_l}{V_{M_l}}\right)\right)^2}{2\sigma_M^2}} \right] + \frac{1}{\sigma_i \sqrt{2\pi}} e^{-\frac{\left(\log\left(\frac{i}{V_i}\right)\right)^2}{2\sigma_i^2}}$$

Penalties on fishing mortality (*cf.* multifan-cl and A-scala) may be added (at least during the first phases of the estimation process) to keep results in a realistic range.

References

- ANGANUZZI, A., R. Hilborn, J.R. Skalski. 1994. Estimation of size selectivity and movement rates from mark-recovery data. *Can. J. Fish. Aquat. Sci.* 51 : 734-742.
- BERTIGNAC, M., P. Lehodey, J. Hampton. 1998. A spatial population dynamics simulation model of tropical tunas using a habitat index based on environmental parameters. *Fisheries Oceanography*. 7:3/4, 326-334.
- FONTENEAU, A. and P. Pallarés. 1998. Considérations sur l'état du stock de listao de l'Atlantique. *Col. Vol. Sci. Pap. ICCAT*, 49(3): 446-458.
- 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. 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, J. Majkowski and J. Hampton. 1990. MULTIFAN a likelihood-based method for estimating growth parameters and age composition from multiple length frequency data sets illustrated using data for southern bluefin tuna (*Thunnus maccoyii*). *Can. J. Fish. Aquat. Sci.* 47 : 301-317.
- HAMPTON, J. 2000. Natural mortality rates in tropical tunas: size really does matter. *Can. J. Fish. Aquat. Sci.* 57 : 1002-1010.
- HILBORN, R. and C.J. Walters. 1992. *Quantitative Fisheries Stock Assessment. Choice, Dynamics and Uncertainty.* Chapman and Hall, New York, 570p.

- HILBORN, R. 1990. Determination of fish movement patterns from tag recoveries using maximum likelihood estimators. *Can. J. Fish. Aquat. Sci.* 47 : 635-643.
- MARCHUK, G.I. 1990. Splitting and alternating direction methods. *In* Handbook of numerical analysis, volume 1, p. 197-462. North Holland, Amsterdam, 1990.
- MAUNDER, M., G. Watters. 2000. A-SCALA : an Age-Structured Statistical Catch-At-Length Analysis for assessing Tuna Stocks in the Eastern Pacific Ocean. IATTC technical paper.
- MAURY, O., D. Gascuel and A. Fonteneau. 1999. Spatial modeling of Atlantic yellowfin tuna population dynamics: a habitat based advection-diffusion-reaction approach with application to the local overfishing study. *In*: Spatial Processes and Management of Fish Populations, 17th Lowell Wakefield Symposium, Alaska Sea Grant College Program, Anchorage 27-30/10/1999.
- MAURY, O. 2001. 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 Col. Vol. Sci. Pap. ICCAT, 52(5): 1739-1752.
- MCALLISTER, M.K. and J.N. Ianelli. 1997. Bayesian stock assessment using catch-age data and the sampling/importance resampling algorithm. *Can. J. Fish. Aquat. Sci.* 54 : 284-300.
- MEGREY, B.A. 1989. Review and comparison of age-structured stock assessment models from theoretical and applied points of view. *In* Mathematical analysis of fish stocks dynamics. Edited by E.F. Edwards and B.A. Megrey. *Am. Fish. Soc. Symp.* 6: 8-48.
- PUNT, A.E. and R. Hilborn. 1997. Fisheries stock assessment and decision analysis: the bayesian approach. *Rev. of Fish Biol. and Fisheries* 7: 35-63.
- SIBERT J. R. and D.A. Fournier. 1994. Evaluation of Advection-Diffusion Equations for Estimation of Movement Patterns from Tag Recapture Data. Proceedings of the First FAO Expert Consultation on Interactions of Pacific Ocean Tuna Fisheries. Vol. 1. R. S. Shomura, J. Majkowski and S. Langi (eds) FAO fisheries technical paper 336/1 326pp.
- SIBERT J., J. Hampton and D.A. Fournier. 1996. Skipjack movement and fisheries interaction in the western Pacific. *In* Status of interaction of Pacific tuna fisheries in 1995. Proceedings of the Second FAO Expert Consultation on Interactions of Pacific Ocean Tuna Fisheries. Shomura, Majkowski and Harman (eds.). FAO fisheries technical paper. 365: 402-424.
- SIBERT, J., J. Hampton, D.A. Fournier and P. Bills. 1999. An advection-diffusion-reaction model for the estimation of fish movement parameters from tagging data, with application to skipjack tuna (*Katsuwonus pelamis*). *Can. J. Fish. Aquat. Sci.* 56 : 925-938.

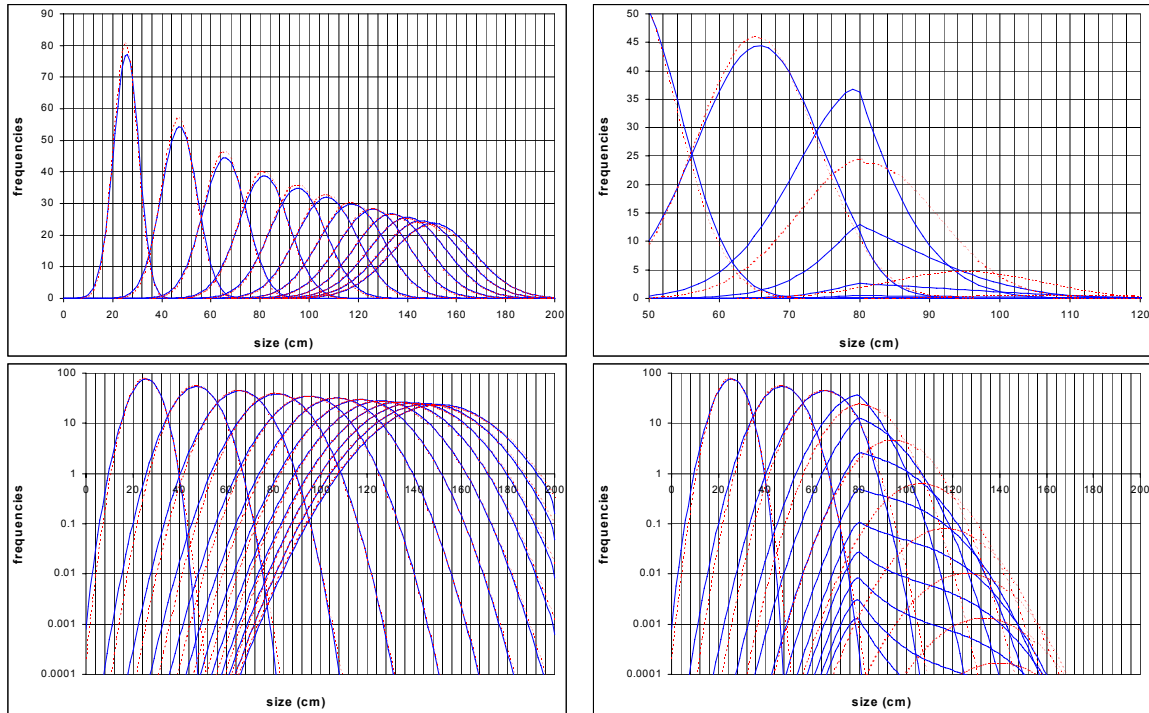


Figure 1. Simulated size frequencies obtained by using a gaussian size structure added to an age structured model (dashed line) or by using a fully size and age structured model such the one presented here (continuous line). First column, no fishing, homogeneous natural mortality over sizes. Second column, no fishing for sizes smaller than 80cm and strong fishing for longer sizes, homogeneous natural mortality over sizes. First line, absolute scale; second line, log scale. When the mortality is not homogeneously distributed over sizes (second column), it appears that simply adding a normal size structure on the age structured model may cause strong biases on abundance, position of the mode and shape of the size distribution for each cohort.

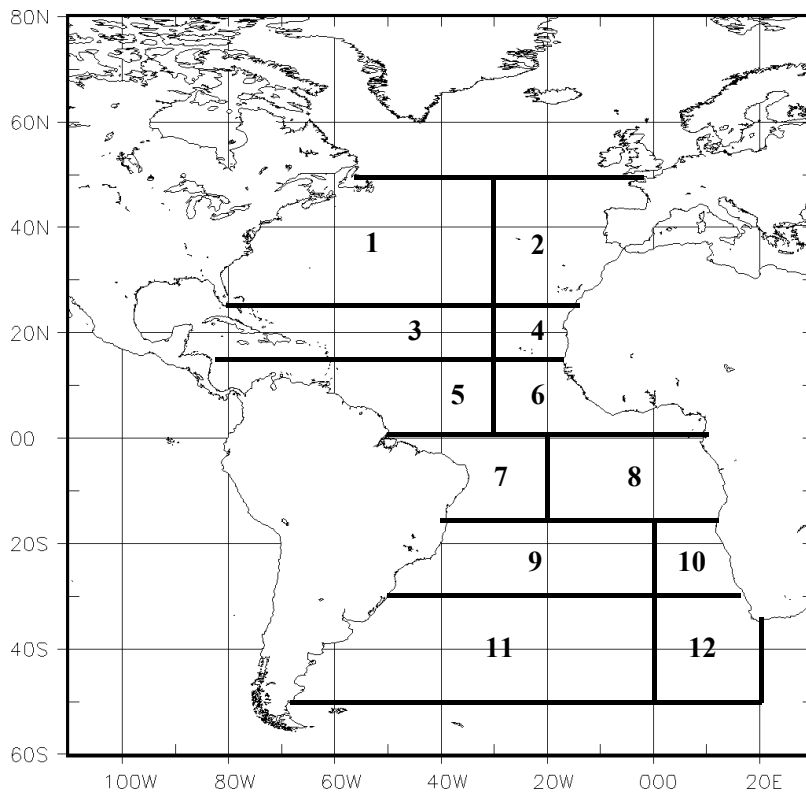


Figure 2. Detailed geographical structure of the model.