Exploratory modelling of Indian Ocean tuna growth incorporating both mark-recapture data and otolith data

R.M. Hillary,
Division of Biology,
Imperial College,
London,
SW7 2BP,
UK.

J. Million and A. Anganuzzi,
IOTC Secretariat,
Victoria,
Seycbelles.

Abstract

This paper is an exploration of the potential growth characteristics of Indian Ocean tunas (yellowfin, bigeye and skipjack) which seeks to integrate the extensive mark-recapture data set gathered during the Regional Tuna Tagging Project – Indian Ocean (RTTP-IO), into the growth estimation procedure, where once only otolith and size frequency data existed for yellowfin and bigeye. This work explores first the integration of the tagging data with otolith data into a coherent estimation scheme and also explores a range of growth models and which might be most suitable for a given species. Key issues with respect to different signals in the otolith and tagging data for yellowfin and bigeye tuna are detailed and potential solutions are discussed.

Introduction

For Indian Ocean yellowfin and bigeye tuna, for many years the only data available with which to estimate growth was the otolith readings (Huang *et al.* 1973; Stequert *et al.*, 1996) and the size frequency collected in the different fisheries of the Indian Ocean, especially the one collected by on the purse-seine fleet operating in the Western Indian Ocean (Lumineau, 2002). Certainly for yellowfin tuna, the growth estimates arising from various growth curves fitted to the data were not considered satisfactory as the different studies based on length frequency analysis showed either that the growth was following a Von Bertalandy curve (Marcille and Stequert, 1976 ; Anderson, 1988), or a two stanza curves similar to the models used in the Atlantic and Pacific (Marsac and Lablache, 1985 ; Marsac, 1991, Lumineau, 2002).(). With the large wealth of growth increment data now available from the RTTP-IO for all three key tuna species there is an opportunity to either integrate these data with the otolith data (yellowfin and bigeye) or in the case of skipjack estimate a full Indian Ocean growth curve for the first time.

Data summary

The RTTP-IO is a project funded under the 9th European Development Fund (EDF) of the European Union, implemented by the Indian Ocean Commission (IOC) and Supervised by the Indian Ocean Tuna Commission (IOTC). The tagging activities started in 2005 with two pole-and-line vessels chartered for 31 months to operate in the Western Indian Ocean. In September 2007, the tagging activities arrived to an end after 168 163 tuna had been tagged and released (54 663 YFT, 34 570 BET, 78 324 SKJ and 606 unidentified tuna. So far, more than 24 000 of these fish have been recaptured and reported to the headquarters of the project based with the IOTC Secretariat in Seychelles, from more than 15 countries. However, most of the reported recoveries are coming from fish caught by the purse-seine fleet based in Seychelles.

Depending on where, during which process and who recovered the tagged fish, the quality of the data associated that will be used for the growth study is highly variable. However, quality indicators have been recorded during the measurement at tagging and at recovery in order to be able to discriminate the data. Furthermore, each recovery should ideally be associated to one position and date of catch, in order to calculate the time-at-liberty (time spent by the tagged fish between the release date and the recapture date). However, due to the process of storing and unloading the fish on the purse-seiners, this is not always possible. In fact, fish are loaded to different wells which can contained several sets, and so several date and location of catch and therefore, it is not always possible to have the real date of catch for all the recoveries. Around 20% of the recoveries are made during the fishing operation, when the fish is loaded into the wells. In this case, the finder is able to identify the exact date and position of the recapture. In this exploratory work, only this subset of recoveries – found during the fishing operation – have been analysed, and this subset have been filter in order to keep only the measurement data of good quality. This represents 1084 recovered yellowfin, 836 recovered bigeye and 1713 recovered skipjack.

Models & methods

For the growth work we worked almost exclusively with three candidate growth models. The first was the well known von Bertalanffy growth equation relating length, *l*, to age, *a*, as follows:

$$l(a) = L_\infty\left(1 - \exp\left(-k(a - t_0)\right)\right). \quad (1)$$

Here the parameters are the maximum growth, L_∞, the growth rate, *k*, and the ‘age’ at zero length, t_0. A second model, more complex and also more flexible, called the Richards model was also employed:

$$l(a) = L_0 + \frac{L_\infty - L_0}{\left(1 + \beta \exp\left(-k(a - t_{\max})\right)\right)^{1/\beta}}. \quad (2)$$

The parameters, that are familiar with the von Bertalanffy model, have essentially the same interpretation but L_0 is a minimum length parameter, t_{max} is the age at maximum growth rate and β is a parameter that allows the growth rate to slow down then speed up as the fish ages, if required. A final model, with one less parameter than the Richards model but allowing this two-stage growth rate, was also considered:

$$l(a) = L_0 + \frac{L_\infty - L_0}{\mu^\gamma + a^\gamma} \quad (3)$$

Here, the common parameters mean the same but maximum growth occurs at $\mu/2$ and if $\gamma \leq 1$ then the model acts like a von Bertalanffy model, and if greater than 1 acts more like the Richards model, allowing for the two-stage growth behaviour. The Gascuel growth model (Gascuel et al, 1992), with one extra parameter than the Richard's model, was initially included in the model set but, given it possesses no closed-form inverse, it was found to be computationally infeasible to fit this model to the data.

The otolith readings of length-at-age, (the Stequert data primarily) for the yellowfin and bigeye were assumed to be normally distributed about their model-predicted counterparts in the likelihood function:

$$L^{oto}(l \mid a, \theta) = \frac{1}{\sqrt{2\pi\sigma_{oto}^2}} \exp\left(-\frac{(l(a) - \hat{l}(a, \theta))^2}{2\sigma_{oto}^2}\right). \quad (4)$$

Here, $\hat{l}(a, \theta)$ is simply the model-predicted length of the fish given it is aged a. The variance of this relationship is modelled as follows:

$$\sigma_{oto}^2 = \sigma_{age}^2 + \sigma_{PE}^2, \quad (5)$$

where σ_{age}^2 is an ageing error fixed term and σ_{PE}^2 is a process error term that is estimated along with the parameters of the model, denoted henceforth by θ.

With respect to the integration of the tagging data, we have measurements of length-at-release, length-at-recapture and time-at-liberty, so for each recovered fish we have an increment of growth and a time over which this increment occurred. Also, given a growth model, we can predict this growth increment given the length-at-release and the corresponding time-at-liberty so the likelihood of the observed growth increment, χ, given the model-predicted growth increment, $\hat{\chi}$, is assumed to be normal:

$$L^{tag}(\chi \mid \hat{\chi}) = \frac{1}{\sqrt{2\pi\sigma_{tag}^2}} \exp\left(-\frac{(\chi - \hat{\chi})^2}{2\sigma_{tag}^2}\right). \quad (6)$$

The form of the variance term in Eq. (6) is as follows:

$$\sigma_{tag}^2 = \sigma_l^2 + \sigma_\chi^2, \quad (7)$$

where

$$\sigma_\chi^2 = \varphi \times l_{rel}^{\mu} \times t_{lib}^{\nu}. \quad (8)$$

The reason for the complex nature of the process error term, σ_χ^2, is to attempt to model the fact that the natural variation in the growth increment is related to both the length-at-release and the time-at-liberty – one school of thought is that given variation in growth seems to increase with age and/or length then there is clearly going to be such an influence on the growth increment for both longer release lengths and longer times at liberty. For computational purposes, the variance term was expressed as an exponential log-sum. The total likelihood of all the data given the model would be the product of Eq (4) and Eq. (6) but for numerical purposes, we maximise the log-likelihood function with respect to the growth and process error parameters.

As a final possibility we allowed for a specific penalty term to give a certain weight (via a normal density function) to an arbitrary age at which the length was assumed known. We point out that such a term was not included in the analyses presented here but, depending on considerations and recommendations, could be used in future work.

All analyses were undertaken in R, using the optim(…) function. The authors experienced significant problems getting a gradient-based optimiser to work (the L-BFGS-B algorithm) and, as such, approximate variance-covariance information that can normally be derived from such methods was not available. We found that the only reliable optimisation algorithm (other than the simulated annealing option) was the Nelder-Mead algorithm (Nelder and Mead, 1965) which cannot produce variance-covariance information and, as such, we cannot produce any standard errors for the current parameter estimates.

Observation error terms

Estimates of variance were provided for the Stequert otolith data and also for the measurements of length (REF JULIEN/ALE?). For the length data, the precision of the measured length was said to be around +/- 1cm so we assumed that $\sigma_l^2 = 0.25$, so that the 95% confidence interval envelopes this range of length-measurement error. With respect to the otolith data, the situation is a little more complex – we have measurements for the error in the ageing reading from the otliths, but based only on 4 different readings. Also, it was not clear how to translate this ageing error into an observation error for the observed length-at-age, without already invoking a growth curve or something similar. Another issue that arose in the early exploratory runs was that there appears to be some critical value of this observation error for length-at-age at which the estimation algorithm appears to abandon these otolith data altogether, in the sense that the optimal parameter estimates ignore the age data altogether and fit only to the tag data, producing clearly nonsense answers. This is to be expected: the growth increment data alone cannot possibly hold information relating to the age-specific model parameters (t_0 for example). This was a dilemma – without the otolith information we can only fit reduced versions of the growth curves (as we do for the skipjack) or we must begin to fix values of the age-specific parameters, or apply the penalty term already mentioned in order to supply this vital age-specific anchor in the data.

For the yellowfin and bigeye data, only the Richards and the alternative two-stage model in Eq. (3) were fitted given the apparent evidence for the two-stage growth in the tag growth increment data and using model selection criteria detailed later on. Figure 1 shows the fits to both the Stequert otolith data and the tag growth increment data for yellowfin. Clearly, there is a mismatch between what the otolith data suggests about length-at-age and what the tagging data predicts in terms of how the fish grow. At the youngest ages the tagging data expect the fish to be longer at the lower ages than the otolith data. Also, the asymptotic length, L_∞, is much lower than that predicted by the otolith data alone. A similar problem appears when fitting the Richards model to the bigeye data, Figure 2, albeit not a significantly as is the case with the fits to the yellowfin data. When trying the sigmoidal model from Eq. (3) a similar problem appeared for both species, and similar answers were obtained that were, statistically speaking, no better than those obtained using the Richards model and, given the familiarity of the group with the Richards model, this model was used for the purposes of plots etc. Also, in all cases there was significant evidence for increased variation in the growth increment with both time-at-liberty and length-at-release – the longer a fish was at release and the longer it was out at sea the more variable the growth increment appears to be. This is interesting as some studies have seen evidence that yes we do see increased variation with length-at-release but not necessarily time-at-liberty (the argument being that and environmental changes affecting growth rates should 'average out' better over longer periods at sea). Perhaps this happens but it is no means true that it must occur this way – to look at this result from another perspective we might consider that increases in growth variation with both length-at-release and time-at-liberty are proxies for an increase in the variation of growth rate with increasing age and length, which is something we might well expect.

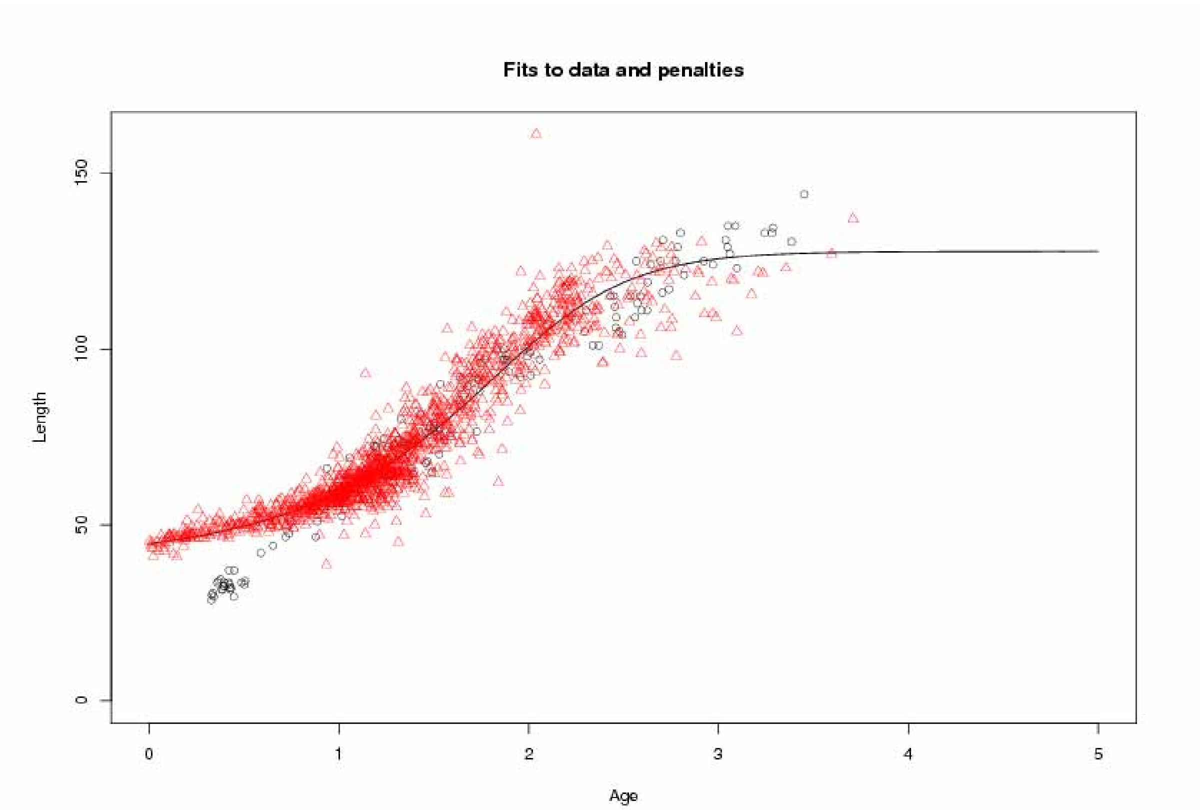

Figure 1: Fits to the otolith and tagging growth increment data for yellowfin tuna using the Richards growth model. Otolith data are the black circles and the observed length at recapture is given by the red triangles. Not the clear two-stage growth rate in the tagging data and the disparity between the younger age otolith data and the tagging data.

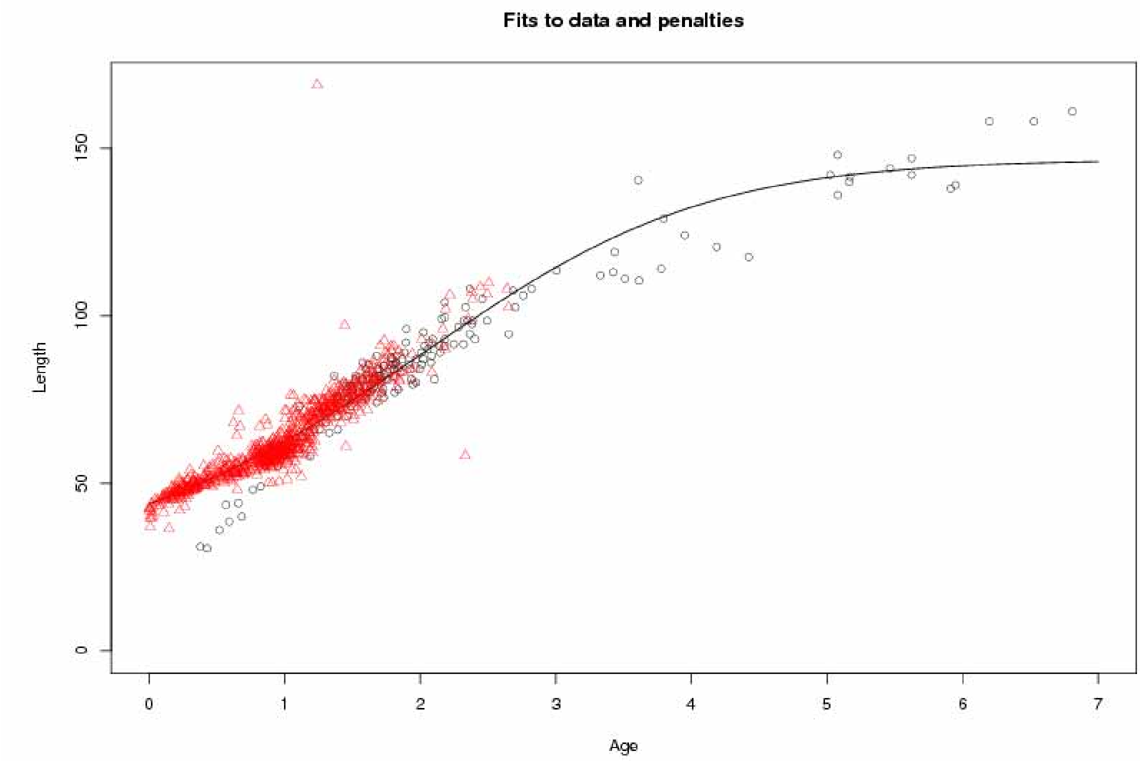

Figure 2: Fits to the otolith and tagging growth increment data for bigeye tuna using the Richards growth model. Otolith data are the black circles and the observed length at recapture is given by the red triangles. Not the clear two-stage growth rate in the tagging data and the disparity between the younger age otolith data and the tagging data.

One very important issue is also the stability of the results – by this we mean the robustness of the estimates of growth to alternate starting points for the R-based optimiser. In an ideal world situation we would always arrive at the same optimal answer (*i.e.* that which maximises the log-likelihood). This situation does not appear to apply in our case:

Multiple starting points for the different parameters to estimate were used to try and see if there was indeed any multi-modality in the log-likelihood and it appeared that there was. The main issue is that we lack observations of length-at-age or growth increments at the longer lengths and higher ages, which would naturally give us information on the asymptotic length. However, this parameter also strongly correlates (levels close to 1 in absolute terms) with some of the other parameters, so the observations at lower ages are heavily influencing the estimate of asymptotic length – something they would in reality hold little information on; as a result, the estimates of asymptotic length tend to 'jump-around' quite a lot (from 120-150cm for yellowfin but less so for the bieye). There are many local optima for the parameters so we cannot really give a definitive estimate for the growth parameters in this present formulation for yellowfin and also for bigeye – the simpler von Bertalanffy model simply does not fit the tagging data at all but the data lack the information to fully separate the parameters in the more complex models that can accommodate the two-stage growth rate. It was then study to fix one of the criteria, especially the asymptotic length. (maybe we could give a couple of fits with realistic Linf (160 for example and 180) and see how this is affecting the curve. I did not manage to get fits for these one that could results in a plot (in all my attemps with the fix Linf, the curve is way outside, and doesn't appear).

Model selection criteria

Seen as we are comparing models, we need some kind of model selection criteria. The one we chose to use, given models with differing numbers of parameters but the same composite normal likelihood function and data, uses the AIC and the chi-squared distribution. For a given model M_1 and a more parsimonious model (*i.e.* with less parameters) M_2, then for a given likelihood function, $L(\theta \mid X)$, and associated MLE estimates θ_1 and θ_2, respectively, the following statistic:

$$\Delta_{1,2} = -2 \times \ln\left(L(\theta_2 \mid X) / L(\theta_1 \mid X) \right) \tag{9}$$

is a chi-squared random variable, with degrees of freedom equal to the number of extra parameters in model M_1. The probability that model M_1 is no better than model M_2 is then easy to assess using the chi-squared distribution thus giving us a significance test for whether a more complex model improves on a more parsimonious one. It should be noted that this test is strictly only for nested models, and neither the Richards nor the sigmoidal model reduce down to the von Bertalanffy model.

When comparing the von Bertalanffy model with the Richards or the sigmoidal model, for both bigeye and yellowfin tuna, there was a significant improvement in moving to either of the more complex models – even though we have stated that we do not believe there is a unique MLE for these more complex models, any of the two to three 'local' MLEs all suggested a highly significant improvement in the fits to the data with these models. This is not a major surprise, given that the von Bertalanffy model simply cannot accommodate the apparent two-stage growth behaviour seen in the yellowfin and bigeye tagging data. With respect to choosing between the sigmoidal model of Eq (3) and the Richards model, depending on the choice of MLE we used, the Richards model either presented a significant improvement ($p < 0.05$) or did not. Both models can reproduce the two-stanza growth curve and in the plots we showed the fits to the data using the Richards model, given that this has already been applied to these stocks and the Richards model has already been tested (Lumineau, 2002) whereas the model in Eq (3) has not.

This criterion was used to differentiate not just the growth models but also to see whether both the otolith data process error term and the more complex generalised variance function for the growth increment process error term were warranted. For both yellowfin and bigeye, the inclusion of the otolith process error term was highly significant ($p < 0.001$). For the growth increment process error term the inclusion of both the time-at-liberty and length-at-release terms were also highly significant ($p < 0.0005$) and this was the case for all three species.

For skipjack, there is no apparent evidence for this two-stage growth rate so a simple growth increment von Bertalanffy model was fitted to these data – it should be noted that in this case we can only estimate the k and L_∞ parameters and not t_0 because we have no age data at all to tell us what this parameter might be. We used the same likelihood as used before and detailed in Eq. (6) and the same observation/process error structure as defined in Eqs. (7) and (8). Figure 3 shows the fits to the growth increments for the skipjack tagging data. The estimates were k=0.289 and L_∞ = 76cm – slightly slower growing than in other oceans but nothing that would appear suspicious.

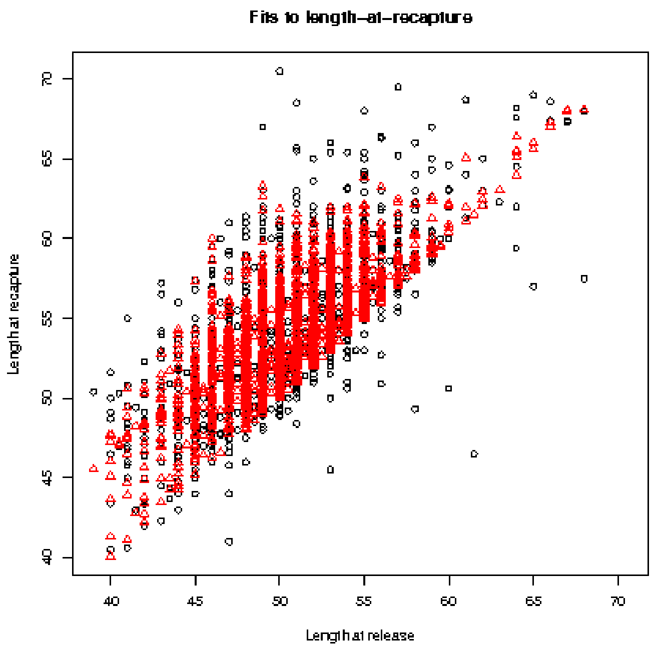

Figure 3: Fits to the skipjack length-at-recapture given length-at-release assuming a von Bertalanffy growth model. Obviously the negative increments cannot be fitted but overall the fit is reasonably good.

Discussion

The results detailed in this paper represent a first analysis of how we might use the wealth of tagging data to explore growth models for all three key tuna species in the Indian Ocean. With respect to yellowfin and bigeye tuna, when integrating the growth increments from the tagging data into a probability model with the otolith data it is clear that the von Bertalanffy model cannot fit the tagging information at all well. When moving to a more complex form of growth model that can accommodate the apparent two-stage growth behaviour of these two species we seem to lose the ability to estimate an unequivocal maximum likelihood estimate for the parameters. Also, there is a clear disparity at the younger ages between what the tagging and the otolith data think about length-at-age.

For the skipjack data, the growth rate and asymptotic length parameters of the von Bertalanffy model were estimated using a simple growth increment form of this model and using the tag data only, given no otolith data for this species. In this case the MLE was robust to different starting values and displayed none of the instability seen in the more complex yellowfin and bigeye models. The resultant parameter estimates did not seem to be massively different to those estimated for Maldivian skipjack (Adam and Kirkwood, *pers. com.*).

For all three models, there was strong evidence for an increase in the variation in growth with both time-at-liberty and length-at-release, both of which could be considered as proxy effects for an increased variation in growth with age and length, as one might expect. However, one of the main arguments levelled against the estimates of asymptotic length in this work and in previous such studies (IOTC, 2007) is that such an asymptotic length simply does not tally with observed catch-at-length data. One thing to consider though is this: is the observed strong variance effect enough to produce the types of lengths seen in the catches with a sufficiently high probability, even with such a low 'mean' asymptotic length? The point being that what you see in the catches might simply be highly variable but not improbable realisations from such an uncertain growth curve. Seen as these fish are seemingly never aged (by inspection of the spread in the otolith data) it is impossible to know whether we have an artificially small estimate of asymptotic length or if the larger fish seen out there are not-so-improbable super-individuals, which one might expect to see given the apparent variation in growth at the longer lengths and older ages.

One would have to conclude that these analyses in this paper are insufficient to establish a robust growth curve for yellowfin and also for bigeye tuna at this stage. For skipjack, we appear to have a robust and sensible growth model which can be used for the construction of a length transition matrix, which is vital for any sort of realistic attempt to integrate the tagging and catch data into an assessment model. How to progress with the growth modelling for yellowfin and bigeye is something that the Working Party will hopefully be able to clarify, but here are some potential options for consideration:

One obvious first choice is to fix one (or maybe more) of the key parameters – particularly the asymptotic length. If we abandon the idea of integrating the otolith and the tag data then we will have to fix the age-specific parameters of the more complex growth curves if we wish to estimate the other parameters from the tag data alone, given the growth increment data tell us nothing about age, *per se*. We have the option of choosing a specific set of points at which we specify the length at a given age, using the penalty option outlined previously, but this is obviously extremely subjective. Another option is to use Bayesian approaches – such multiple modality in the likelihood is less of a problem then (we are deriving a sample from the parameter's posterior distribution, not simply the MLE) and we can perhaps use informative priors for asymptotic length instead of simply fixing the values. With the Bayesian option we may also develop a better understanding of the variation in the estimated parameters, which was not possible using maximum likelihood techniques.

Acknowledgements

The tuna tagging data analysed in this publication were collected by the Regional Tuna Tagging Project of the Indian Ocean (RTTP-IO) funded under the 9th European Development Fund (9.ACP.RSA.005/006) of the European Union. The RTTP-IO was implemented by the Indian Ocean Commission under the technical supervision of the Indian Ocean Tuna Commission. We wish to acknowledge the contributions of the project staff and all the technicians, recovery officers and fishers that have been involved in the RTTP-IO.

References

Anderson, R.C., 1988. Growth and migration of juvenile yellowfin tuna (*Thunnus albacares*) in the central Indian Ocean. Coll. Vol. Work. Doc., IPTP, TWS/88/21 :28-39.

Bertalanfy, L. Von., 1938. A quantitative theory of organic growth (inquiries on growth laws II). Hum. Biol., 10 (2) : 181-213.

Huang, C.C., Sun, L., Yang, R.T., 1973. Age, growth and population structure of Indian yellowfin tuna. J. Fish. Soc. Taiwan 2:16-30.
IOTC, 2007. Notes on the estimation of a growth curve and catch-at-age for yellowfin tuna.

J.A. Nelder, J.A. and R. Mead, R., 1965. Computer Journal, 1965, vol 7, pp 308-313

Lumineau, O., 2002. Study of the growth of yellowfin tuna (*Thunnus albacares*) in the western Indian ocean based on length frequency data. Document WPTT-02-18, IOTC WPTT 2002, 3-11 June 2002, Shangai, China. 10p.

Marcille, J. and Stéquert, B., 1976. Croissance des jeunes albacores, *Thunnus albacares*, and patudos, *Thunnus obesus*, de la côte nord-ouest de Madagascar. Cah. Orstom, sér. Océanogr., 14(2), 153-162.

Marsac, F and Lablache, G., 1985. Preliminary study of the growth of yellowfin estimated of purse seine data in the Western Indian Ocean. Coll. Vol. Work. Doc., IPTP, TWS/85 :22 :91-110

Marsac, F., 1991. Growth of Indian Ocean yellowfin tuna estimated from size frequencies data collected on French purse seiners. Coll. Vol. Work Doc., IPTP, TWS/91/17 :35-39.

Richards, F.J., 1959. A flexible growth function for empirical use. J. Exp. Bot., 10 : 290-300.

Stéquert,B., Panfili, J., Dean, J.M., 1996. Age and growth of yellowfin tuna, Thunnus albacares, from the western Indian Ocean, based on otolith microstructure. Fish. Bull. 94, pp. 124-134.