

## **Scientific questions and recommendations following the 2009 yellowfin IOTC stock assessment**

By Alain Fonteneau, IRD Scientist

### **Summary**

This document discusses the difficulties encountered by recent stock assessment conducted on yellowfin tuna by the IOTC WGs in 2008 and 2009. This assessment was done using the MFCL model, a complex statistical model often used to analyse stock status, but also a model that has been often providing rather questionable results. The paper discusses the typical problems and uncertainties faced by this type of statistical model, and it makes a wide range of recommendation allowing to run on yellowfin, skipjack and bigeye in the Indian Ocean more realistic tuna stock assessment models, that will be making a full use of the highly valuable IOTTP results. It is also concluded that the results obtained by these models (for instance recruitment, biomass and fishing mortality by area, and MSY) should necessarily be accepted by scientists expert in tuna stock assessments, as being valid ones.

### **INTRODUCTION**

The MFCL model used by the IOTC TT WG has been totally (2008) or quite unsuccessful (2009) to estimate in a consistent and convincing way the past & present status of the Indian Ocean yellowfin stock.

This unexpected failure is rather disappointing (1) as this model is a quite optimal and well founded stock assessment model, (2) as this MFCL model was ran by one of its best expert, Adam Langley, (3) as it was anxiously expected that the recent results of the large scale tagging program would widely improve the functioning and results of the MFCL model.

It is then very interesting to identify the causes explaining this failure, as this understanding should allow to solve these problems in the future yellowfin stock assessment done by the IOTC. These potential additive causes can be grouped in the following categories:

#### **(1) Increasing and major deficiencies in the basic catch at size and effort data:**

It should be kept in mind that statistical data (catch, effort and sizes) are always of fundamental importance in all tuna stock assessment models: when they are highly deficient or biased, even the best model cannot provide a realistic stock assessment. In the Indian Ocean the statistical data are + or - OK for most purse seine fleets and for some longline fleets, but they remain very poor or inexistent for various major longline fleets and for most artisanal countries, many of them being of major importance for yellowfin (Yemen, Sri Lanka, Indonesia, Oman, etc) (see figure 1).

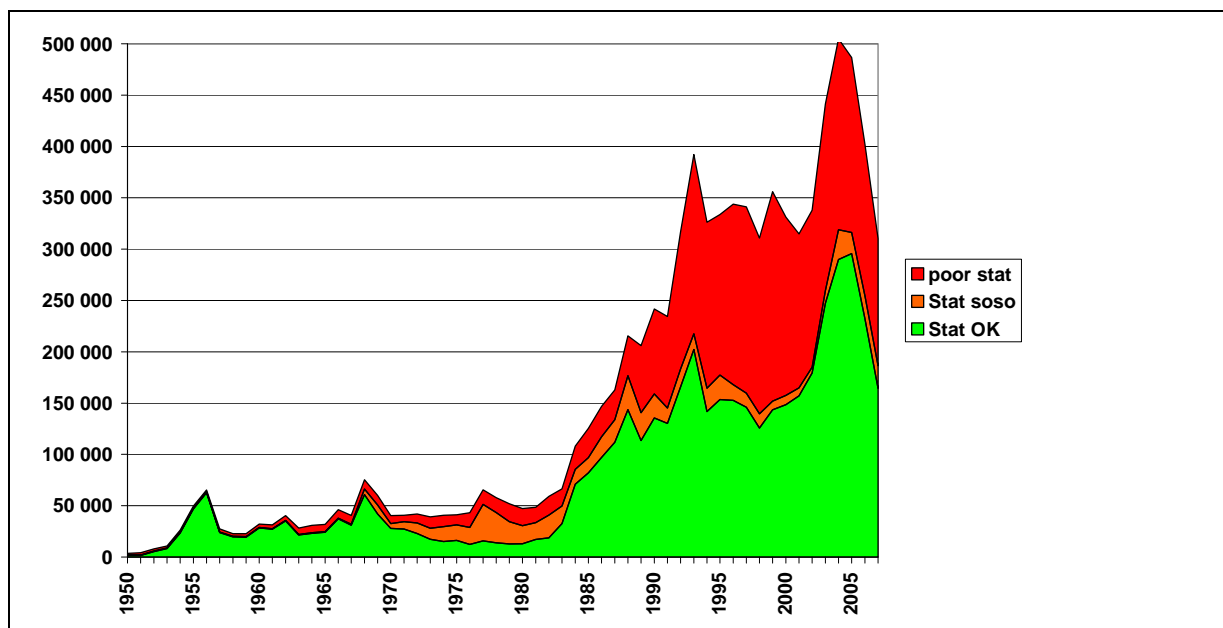


Figure 1: Yearly catches of yellowfin taken in the IO by countries providing good statistics to the IOTC (catch, effort and sizes by time and area strata) (South Africa, Australia, China Taipei, China, Spain, France, Japan, Maldives and Seychelles; green area) and other group of countries with limited statistics (Stat soso, orange area) or poor statistics (red area).

Size sampling has been also totally poor for Japanese longliners during at least since 2002: only 1/1000 of yellowfin caught being sampled and most strata fished without any size sampling (figure 2), at a widely insufficient level, thus producing an unrealistic catch at size table for this Japanese fleet.

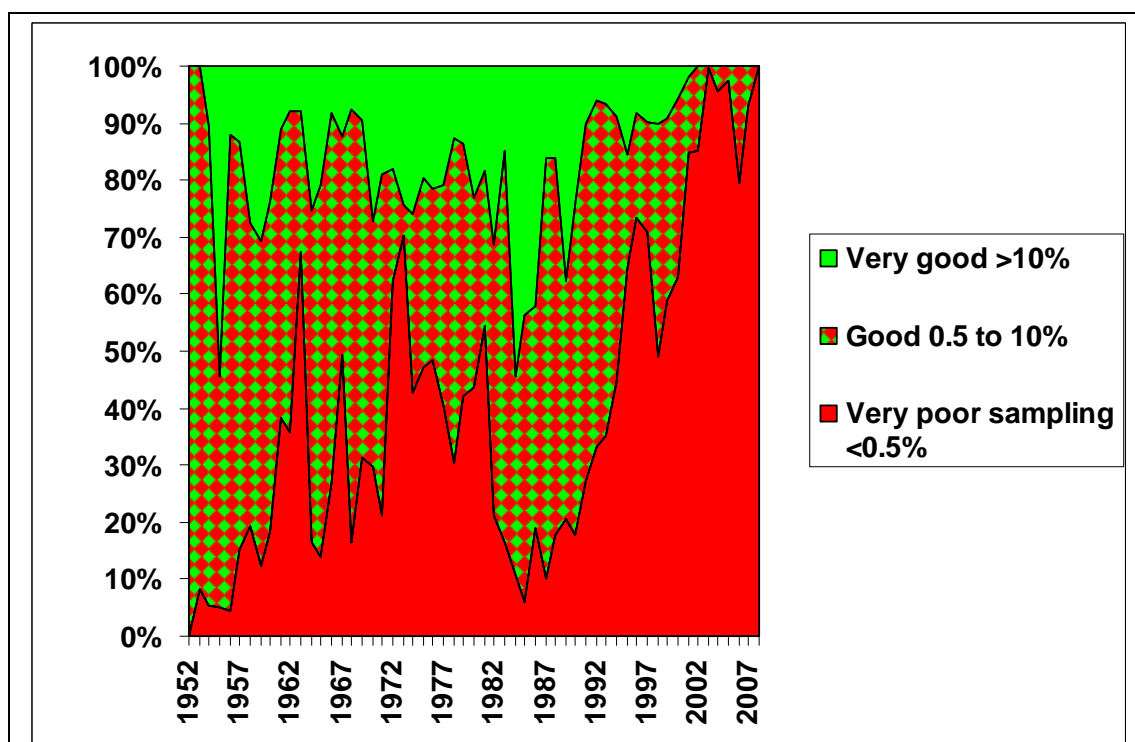


Figure 2: Levels of yearly size sampling of yellowfin taken by Japanese longliners, red area showing a very poor sampling rate < 0.5 %: percentage of sampled strata have been calculated by quarters and geographical cell of 10° lat\*20° longitude

The IOTC Scientific Committee and Commission should make a strong pressure on Japan to improve its sampling: Japan being still a major fishing country, and these yellowfin size data being very important in the present MFCL stock assessments. This basic statistical deficiency remains a source of major uncertainty in all the IO yellowfin stock assessment, as it may have been producing a totally artificial decline in recently estimated recruitments and in recent figures of biomass.

Following the increasing & major statistical deficiencies in the yellowfin fisheries, more than 50% of total yellowfin catches have been taken without basic statistics during the last 20 years and since 2002, a great majority of the Japanese catches are not sampled at all. As a consequence, it would appear that this decline of sample size may have been producing a severe change in the MFCL “virtual stocks”: when a given range of sizes/ages disappears from the Japanese catches, the model tends to conclude that these sizes are absent from the underlying population. This type of problem is qualitatively clear, but it should be further analysed in details and with great care: the reduced size of the samples should not produce a collapse of the modelled stock!

This major structural problem in the stock assessment cannot be solved for historical data, but at least major sampling efforts should be done by longline countries (Japan & Taiwan) to improve their yellowfin size sampling in the near future and the MFCL problems due to this decline of sampling rates should be better analyzed by scientists.

## (2) Conflicting trends in catches & longline CPUEs during the 1950-2008 period:

It was clear before the Mombasa 2009 stock assessment, as in 2008, that this yellowfin stock assessment would be very difficult to run: simply & basically because of the conflicting trends in total yellowfin catches and the yellowfin longline CPUEs used to evaluate the biomass of the adult stock (figure 3):

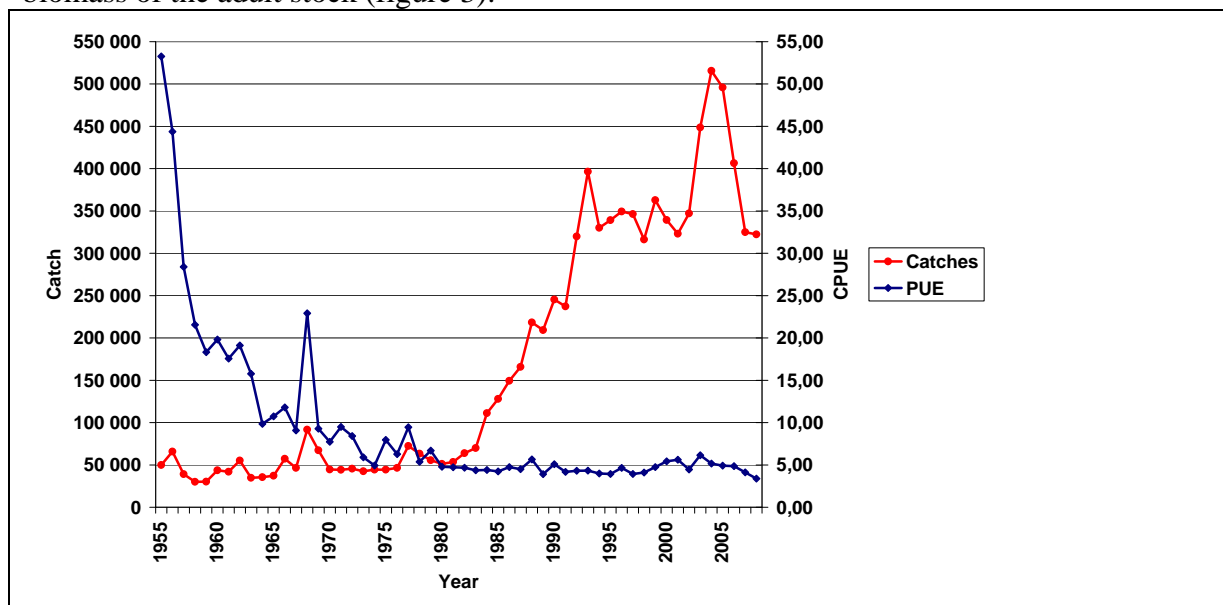


Figure 3: Yearly levels of total yellowfin catches and of nominal Japanese CPUEs (average of the 3 main areas: 2, 3 and 5)

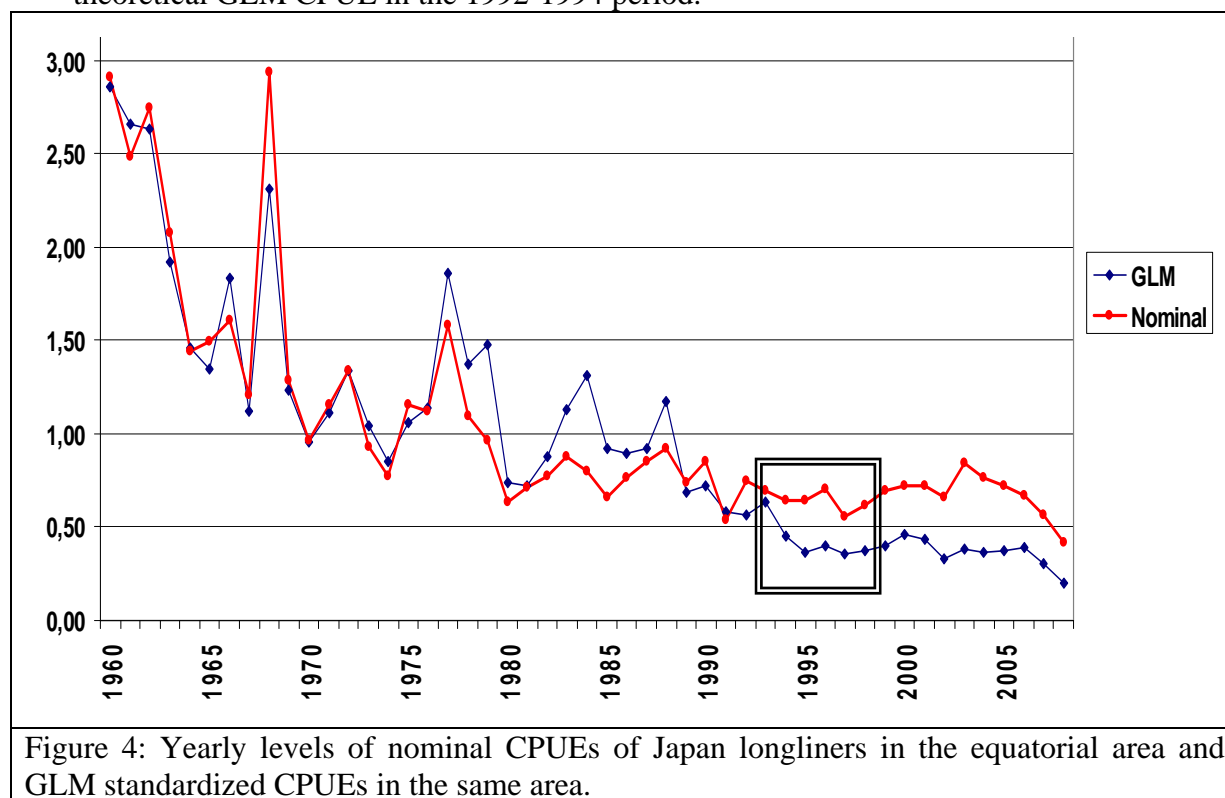
It should be very clear, for every tuna scientists as well as for every IOTC commissioners, *“that these observed trends in yellowfin catches and cpues are not consistent with any*

*known theory of fishing. For any fished stock, to give larger and larger yields with no significant decline in abundance, cannot be explained, unless there are some unexplained factors”<sup>1</sup>.*

The major early decline of longline CPUE was clearly due to a major decline in the stock catchability to longliners (not to an early stock collapse!), but surprisingly the reasons of this major decline are still entirely hypothetical ones. On the opposite, the “excessive” stability of the recent longline nominal catch rates was probably the consequence of the permanent improvements developed by this fishery (as for the purse seine fleet showing permanent increase of its catch rates!), then masking the decline of stock densities. Furthermore, a major structural problem was identified this year by the WG in the trend of the GLM Japanese CPUE that has been heavily used in the present MFCL stock assessment: when nominal CPUEs of Japanese longline tend to be flat in most areas during the last 20 years, the GLM CPUEs estimating biomass trend show a sudden marked decline in 1993-1995 (figure 4).

### **(3) Why the sudden decline of Japanese standardized CPUEs in 1992-1994?**

It is quite strange to note that when the nominal equatorial CPUEs of Japanese longliners have been quite stable during the period 1980-2006, there was a sudden drop of the theoretical GLM CPUE in the 1992-1994 period.



This surprising sudden decline of the GLM CPUEs (assumed to be indicative of the adult biomass) had a major impact on the pessimistic results of the MFCL stock assessment. This result is very surprising for all tuna experts: nominal and GLM CPUE tends to be most often very similar or nearly identical. When they show a divergence, this event should be easy to understand: for instance the effects of deep longline in the late seventies producing a major

<sup>1</sup> A conclusion given by Michael Prager, an external expert in stock assessment, after examining the yellowfin catch and CPUEs trends presently used for the IO 2008 yellowfin stock assessment.

increase of bigeye CPUEs. In the case of the anomaly in the yellowfin GLM CPUEs during the 1993-1995 period was explained by Japanese scientists in the following way:

- (1) New monofilament lines were introduced by Japanese longline during this period.
- (2) It has been estimated by Japanese scientists that these monofilament lines should widely increase the fishing efficiency of longline on yellowfin.
- (3) When nominal CPUEs were stable during the 1990-1995 period, despite of these new lines,
- (4) The GLM model tends to conclude that there was a major & sudden decline of biomass during this period, simply because these new monofilament lines should have produced an increase of the yellowfin CPUEs.
- (5) This uncertainty on the Japanese GLM yellowfin CPUEs are also confirmed by the divergence observed during this period between the declining Japanese CPUEs and the stable Taiwanese CPUEs (shown by figure 5).

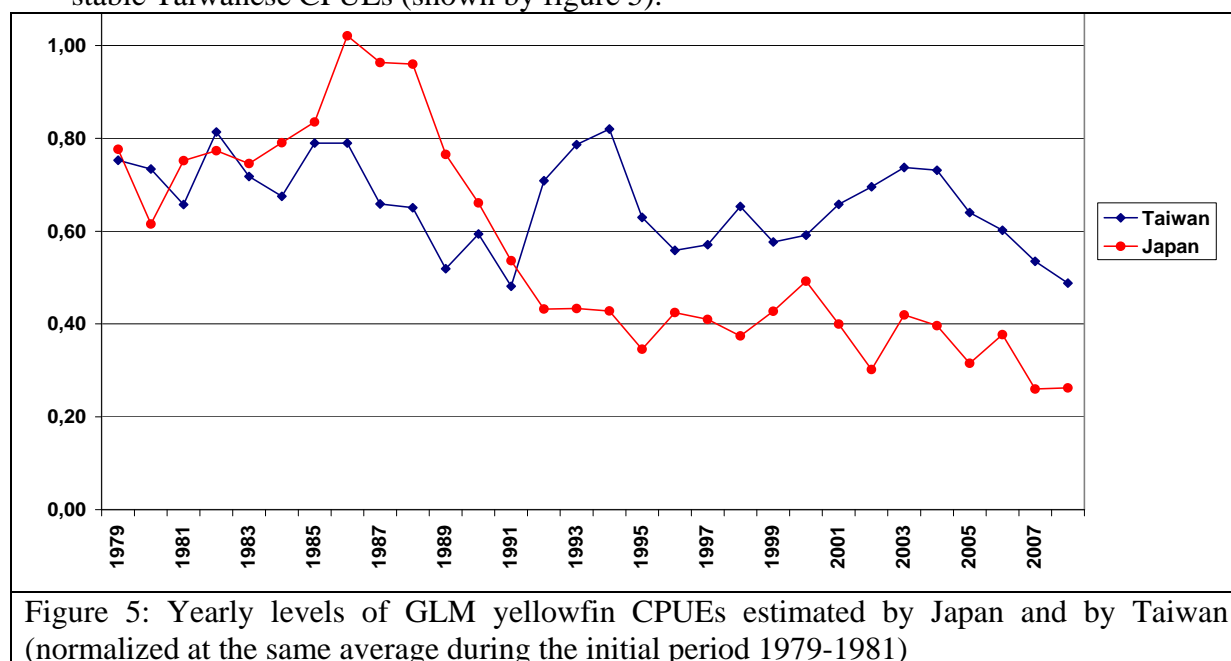


Figure 5: Yearly levels of GLM yellowfin CPUEs estimated by Japan and by Taiwan (normalized at the same average during the initial period 1979-1981)

*A firm conclusion, expressed during the WG, but unfortunately quite lost in its report, is that the collapse of the GLM CPUE in the early 90<sup>ies</sup> is possibly/probably widely artificial and due to a technical problem in the present GLM standardization, giving an improper weight to the “type of line” parameter.*

#### (4) Severe impact on the model of the tagging & recovery data:

Two types of problems have been faced by the MFCL model in its use of the tagging/recovery results.

- (a) **Insufficient analysis of IOTTP results:** the tagging/recovery data recently obtained by the IOTC are highly complex ones and very far to be fully analyzed: none of the basic results –growth, movement patterns, exploitation rates at size or natural mortality at age & by sex- have been fully analyzed today by scientists. Small scale tagging have not yet been analysed. Present MFCL model has been working using only the results of the very preliminary 2008 analysis by the IOTC tagging WG on the RTTP-IO data. The results of the various small scale and pilot tagging operations should also be analyzed in conjunction with the RTTP-IO results. An urgent scientific effort should be devoted to do an in depth and urgent analysis of these updated and corrected recovery data. The weakness of the present analysis

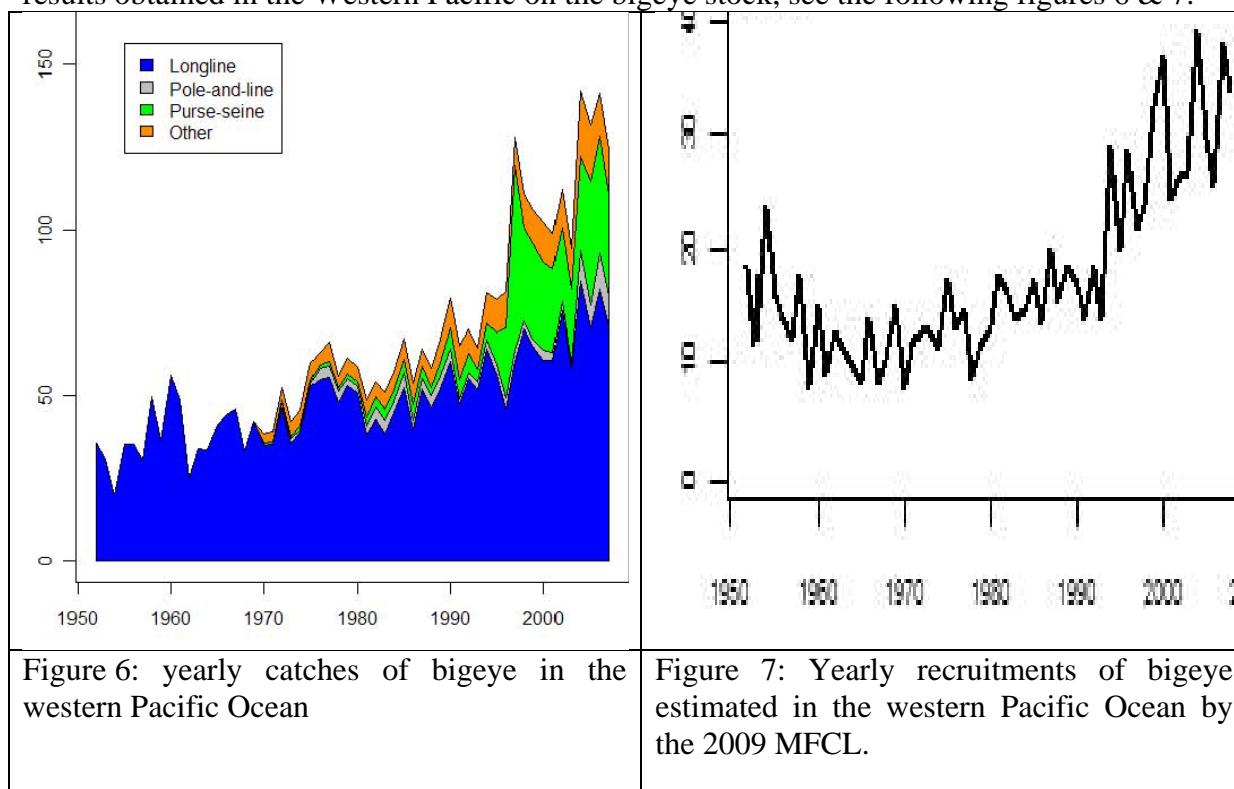
had clearly a negative, but unknown impact on the results of the present MFCL model!

- (b) **Major modelling constraints introduced by the tagging results:** when the MFCL model is based only on catch/effort/size data, this model has a full flexibility to search its best statistical fit among the multiple parameters handled in this model. These results can be good or bad, but these MFCL models have very few constraints to find their best statistical fits. On the other side, when strong results obtained from the tagging programme have been introduced in the new MFCL model, these tagging data put severe statistical constraints on the model basic parameters: on growth, natural mortality, exploitation rates, movement patterns. In my view, these new constraints are now a major source of “statistical disasters” in such models: today, all the MFCL model are + or – “prisoners” by their stability of the assumed age selectivity of each gear and of their stable catchabilities during long periods of time.

## (5) Typical uncertainties often found in the results of tuna MFCL models:

### (A) Unrealistic recruitment trends?

A typical & rather “absurd” result of these statistical constraints is the frequently observed basic result of these models: recruitment yearly levels following most often the trend in total catches (a well known problem discussed by Fonteneau & al 1998). This type of analytical potential bias is for instance well shown by the MFCL 2009 stock assessment results obtained in the Western Pacific on the bigeye stock, see the following figures 6 & 7.



On the opposite, the other family of SPA<sup>2</sup> models or the PROCEAN Bayesian model tend to obtain opposite conclusions: for instance in the hypothesis that recruitment

<sup>2</sup> SPA : Sequential Population Analysis, general term covering a wide range of Catch at age analysis, for instance VPAs.

levels have been fluctuating without major trend they tend to conclude that selectivities and catchabilities have been permanently showing major changes for most gears (longline & purse seine) during the entire history of the fishery.

These alternate SPA or Bayesian production models are often providing strong & logical proofs:

- That changes in tuna fisheries (selectivity & catchability) are most often very important in the typical tuna stock assessment done on long periods of time (ideally, most often during more than a half century)

- that these major trends in the fisheries tend to introduce major bias in the results of the models (such as the present MFFC model), because many/most of them are assuming a constancy of these fishery factors (selectivity & catchability): for instance producing totally artificial and unrealistic trends in the recruitment levels (and of course also in biomass and fishing mortality levels).

**(B) Correspondingly, unexplained changes in biomass or fishing mortality:**

The 2009 MFCL on the yellowfin stock provides a good example of this type of pending question: total estimated recruitment were surprisingly stable during the 1975-2002 period, followed since 2004 by a major & sudden decline in the estimated recruitment, figure 8.

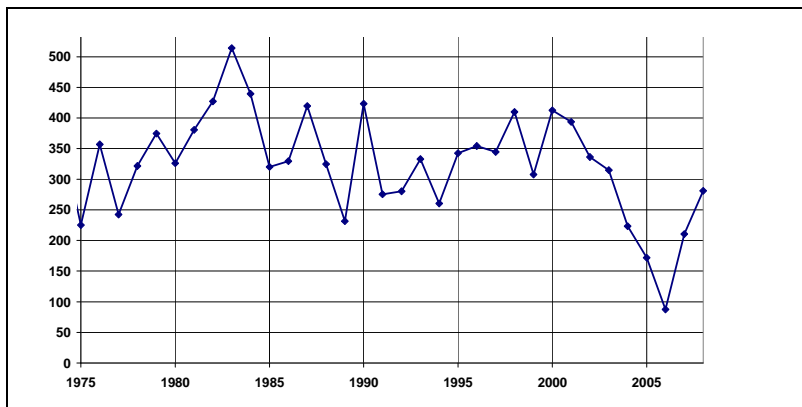
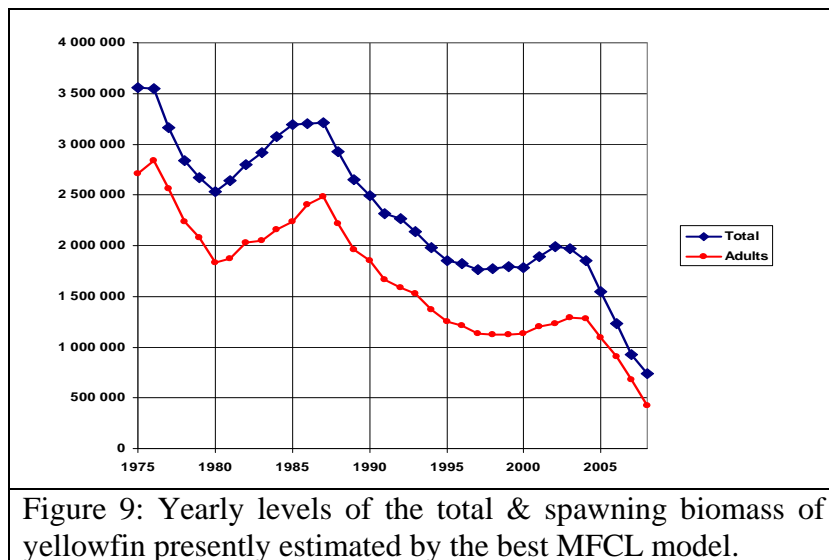


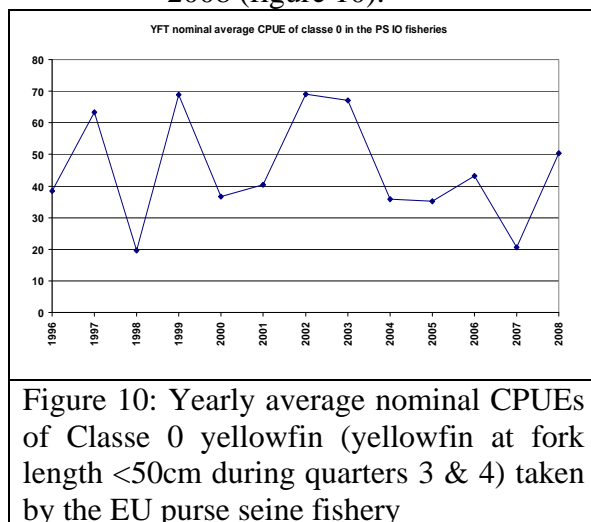
Figure 8: Yearly levels of recruitment estimated by the best MFCL run

This decline of estimated recruitment producing subsequently a major & fast decline in the biomass (total and spawning) (figure 9).



However, the fishery data shows that this major decline is not fully visible in the purse seine or in the longline CPUE data:

- ✓ Purse seine CPUEs of juvenile and adult yellowfin appear to be quite stable until 2008 (figure 10).



- ✓ Japanese & Taiwanese longline CPUEs do not show such major & sudden decline during the 2002-2006 period, but simply 2 very bad CPUEs in 2007 and 2008 (the worse CPUEs ever observed in the history of the longline yellowfin fisheries)(figure 5).

The paradox is that this « estimated collapse » of the estimated recruitment and adult biomass may simply be due to the reduced size sampling of Japanese longliners, see figure 2.

### **-C- Subsequently, very strange trends in recent fishing mortality?**

As a consequence of these estimated trends in stock recruitment and biomass, the present MFCL model estimates that there was a major increase of fishing mortality during



recent years, i.e, since 2006, even when the total yellowfin catches came back to their typical levels of the 1990-2002 period, see figure 11.

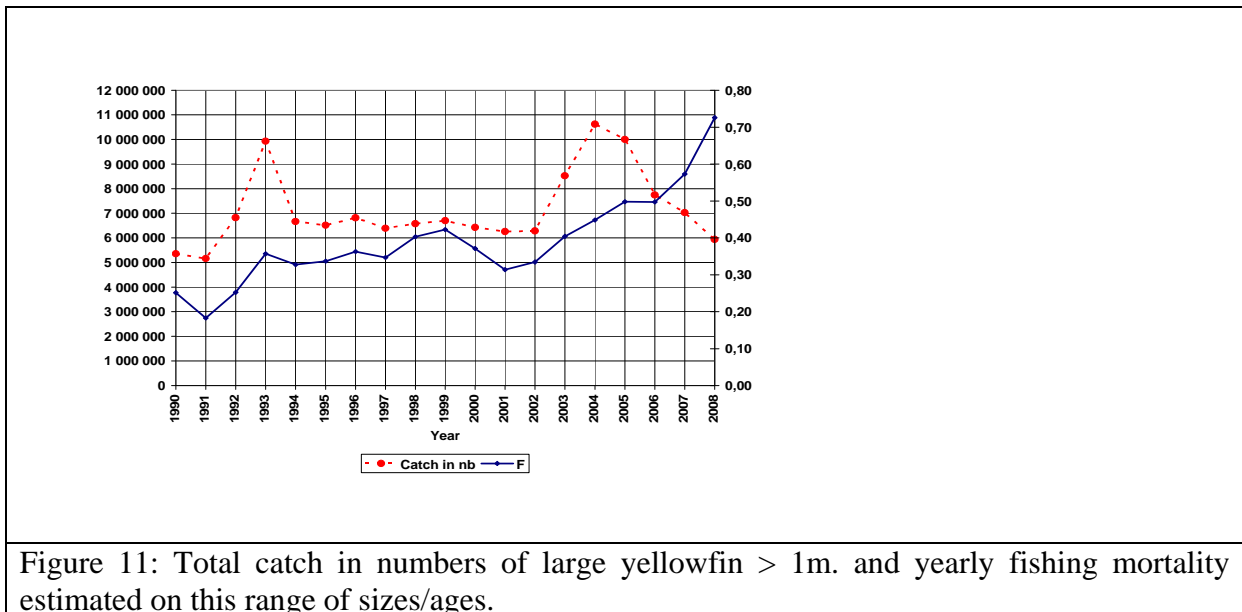
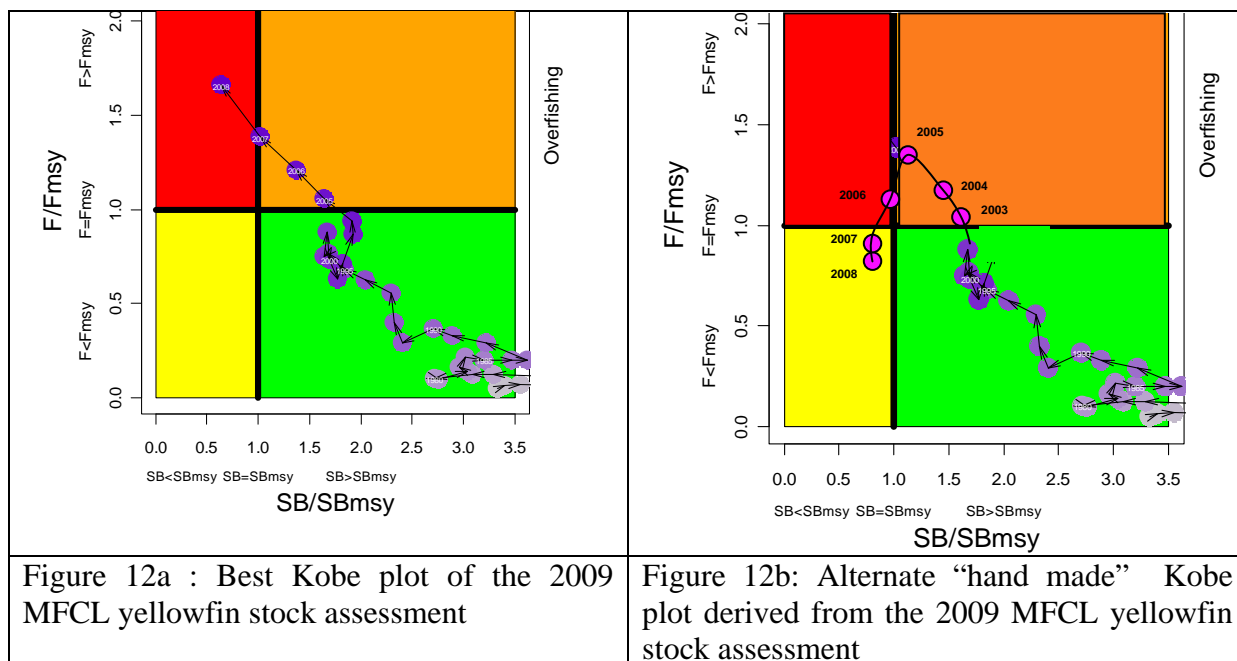


Figure 11: Total catch in numbers of large yellowfin > 1m. and yearly fishing mortality estimated on this range of sizes/ages.

The present MFCL model estimates (strangely?) that there was (1) a quite stable recruitment during the nineties, (2) a moderate increase of  $F$  during the 2003-2006 period of very high catches and (3) a very large increase of fishing mortality in 2007 and 2008. This variability estimated by the present MFCL is quite strange and possibly biased: most tuna biologist working in the Indian Ocean tends to consider that:

- a) during the 2003-2006 period of very high yellowfin catches, there was a combination of very high fishing efforts and very high catchability, then probably a peak in the adult fishing mortality (as there is no major recruitment visible in the catch at size data during the late nineties)
- b) During the 2007-2008 period of reduced nominal efforts and average catchability, scientists would tend to assume that fishing mortalities have been kept + or – stable, and not widely increasing as in the MFCL results (due to the “collapse» of the estimated adult stock).

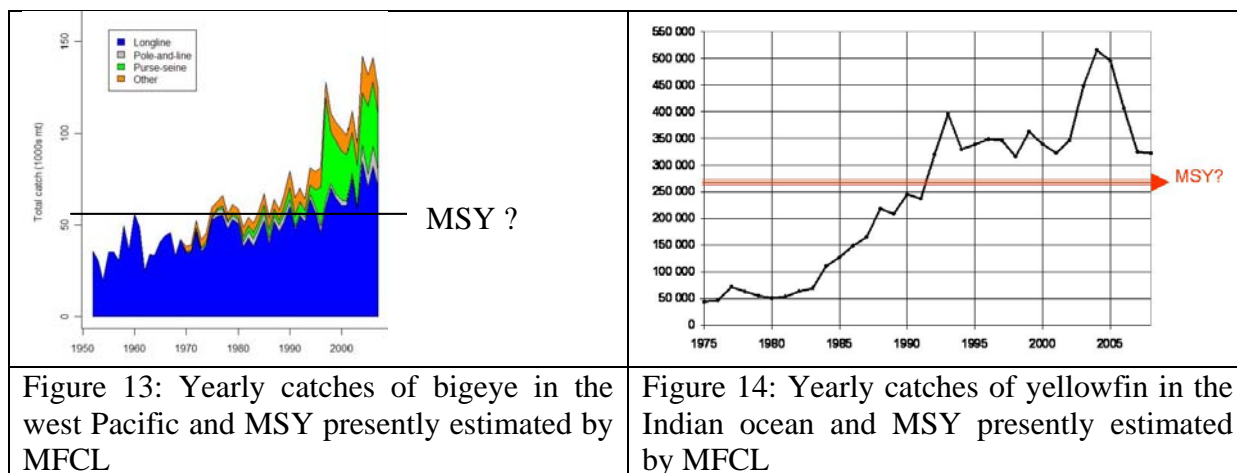
This major question/uncertainty in the present stock status should at least be discussed by scientists and alternate runs of the model should help to explore this basic serious uncertainty. Furthermore, scientists should also propose in alternate to the MFCL statistical Kobe plot, an alternate hand made Kobe plot constructed based on their scientific expertise. Such ‘hand made’ Kobe plot is for instance shown by figure 12b in comparison to the official plot (figure 12a).



There is probably a good probability that such Kobe plot based on the expertise of a group of scientists experts in fisheries and in stock assessment may be more realistic than a given statistical “best fit” of a questionable MFCL result.

#### -D- Too low or too high estimated MSY?

There is a frequent tendency of statistical models (WCPFC bigeye stock, Indian Ocean yellowfin) to estimate MSY that are quite unrealistic, being either too low or too high<sup>3</sup>. These MSY are often estimated to be at a very low level, for instance 30 or 50% lower than the sustained catches observed during 15 years (twice the duration of the exploited life), see the following figure 13 (Western Pacific bigeye) and figure 14 (Indian Ocean yellowfin).



Various other cases of surprisingly high MSY can also be found in the scientific literature of tuna RFOs.

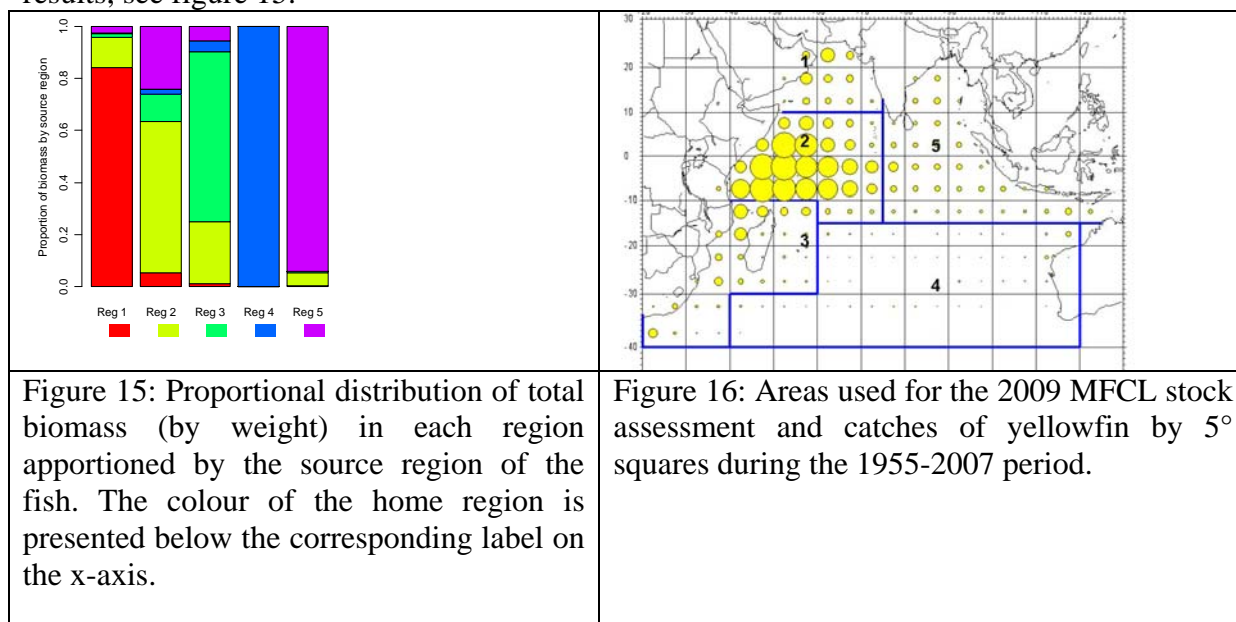
<sup>3</sup> These frequent errors in estimated MSY being mainly the result of the major uncertainties in all the tuna stock/recruitment relationship & their steepness.

These “strange” MSY should not be accepted by scientists, unless when there is a clear scientific reason explaining a change of MSY. It can be assumed that these anomalies in estimated MSYs are mainly due to the major uncertainties faced by the stock recruitment relationship, and by its steepness parameter.

As MSY is a key stone parameter in the functioning of all tuna RFOs, being most often their management target, these too low of too high MFCL MSY should not be accepted blindly by scientists, unless they can be explained by a well identified decline in yield per recruit or to a known change in the environment. In many cases, any alternate production model may be much more realistic to estimate the real MSY that are at least fully consistent with the history of sustainable catches taken on the stocks.

#### **-E- Very strange movement patterns and biomass per area**

Very strange trends and absolute biomass levels have been estimated by many MFCL models on various tuna stocks. Correspondingly, the estimated movement rates between areas are often estimated to be unrealistic. This problem was also noticed in the 2009 MFCL yellowfin results, see figure 15.



Most scientists would tend to conclude, based on fishery data and on the RTTP-IO results, that the mixing rates estimated by the present MFCL model, are much too low and by far. The same yellowfin tuna experts would also tend to assume that most yellowfin are born in the equatorial areas 2 and 5, and later moving at great scale in the other areas (towards the temperate areas 3 and 4 or towards the Arabian Sea, area 1), and probably showing a nearly full mixing in the Indian Ocean at adult ages,. Most/all scientists tend to consider that yellowfin from area 4 do not constitute a small isolated stock as in the MFCL results. In the same way, most/all scientists would disagree on the conclusion obtained by the present MFCL model, that the total biomass in the small south western area 3 was higher than the biomass in the much larger area 2, the core of the yellowfin stock (at least based on fishery data).

### **Conclusion**

The 2009 yellowfin MFCL model has been often producing surprising and probably false results, & *inter alia*:



Questionable movement patterns and biomass per area, that should be carefully analyzed by scientists, as they widely condition the exploitation of the stocks



Questionable recent major decline in recruitment estimated by the model & a major decline in biomass.



And subsequently very strange major increase in recent estimated fishing mortality

These strange results should not be accepted blindly as they may well be due to cascading errors, bias or statistical problems in the present best MFCL model.

The general conclusion is that MFCL are powerful but unstable and dangerous tools and their results should be carefully analyzed by scientists before accepting these conclusions: remembering that MFCL results may be simply due to misspecification of the model or to errors & bias in its input data or in its working hypothesis. This complex work will need an active research on data and methods, see the following recommendations, especially trying to conciliate the best results of the best statistical models and the personal expertise of tuna scientists.

As this work will need quite a lot of additional research, in the present context, the IOTC should manage the yellowfin stock in a precautionary way and following the recommendations done by the Mombasa WG.

## **Research recommendations allowing to improve tuna stock assessment in the Indian Ocean**

**Justifications for the 14 subsequent personal recommendations on tuna stock assessment, mainly targeting yellowfin stock assessment & the use of statistical models to do this stock assessment:**

New statistical models (such as MFCL or SS3) are potentially the ideal stock assessment tools, but they are also very dangerous ones: their “best statistical fit” results often produce very misleading or even totally false results, due to a combination of multiple factors: uncontrolled changes in fisheries and weakness of biological and fishery data (especially size data of Japanese longline ). The complexity of these models never offers a warranty of the validity of their results: great care and meticulous scrutiny of the detailed results should be developed before approving any of these results, especially concerning their estimated MSYs and projections of stocks & fisheries that are widely dependent of the mysterious Stock recruitment parameters and their *infamous* « steepness ».

**In order to facilitate their discussion by the scientific committee, these recommendations have been stratified in 3 categories:**

\* Technical point and recommendations that should be easily and urgently studied and solved by IOTC scientists

\*\* More complex but urgent research recommendations: targeting a problem that should be solved before and during the next IOTC stock assessments

\*\*\* Longer term and more global research recommendations: targeting a basic problem in the tuna stock assessment

### **\* Recommendations**

- 1) **The major disruption in the GLM CPUEs** as the 1993-95 disruption in the Japanese longline yellowfin standardized CPUE series, should be fully explained and justified by the standardisation method (as disruption plays a major role in the recent very bad situation of the stock).
- 2) **Sampled sizes on Japanese longliners:** the potentially vicious effects of the recent **declining sample size of Japanese longliners** should be urgently analyzed, as the decrease of Japanese size samples may have accelerated the recent collapse of the modelled stock!
- 3) **Steepness mysteries?** Scientists should never be confident in the **steepness** based on the dubious stock and recruitment levels estimated by the model, and they should better explore the potential structure & uncertainty in the steepness parameter of the yellowfin Stock Recruitment relationship based on the biological characteristic of the species, for instance using the statistical method proposed by Mangel 2009.

## \* \* recommendations

- 4) **Trend in recruitments and in biomass** as estimated by the model, total and in each area, should be examined with great care and accepted by scientists only if they are estimated by the experts as being reasonable ones.
- 5) **Movements & migration:** Great scrutiny should be given to the **movement patterns and mixing rates that are estimated** between areas: these movements should be “accepted” by the experts as being reasonable ones. In the Indian Ocean yellowfin case, they should be compatible with tagging results and the global knowledge of movements by the experts, based on fishery and environmental data. Today, this is not yet the case, and by far!
- 6) **Population sizes by area:** when the stock assessment **parameters estimated by area** are estimated by the experts to be fully inconsistent, they should not be accepted by scientists: if population size and fishing mortality in each area are clearly wrong, it is highly questionable to assume that their total is OK. In this case, it would be much better to run the assessment models at the scale of the whole Indian Ocean stock.
- 7) **Unexplained trends in stocks & fisheries?:** When the MFCL Stock Assessment model is estimating that **major changes** have been recently occurring in the **stock biomass, recruitment levels or fishing mortalities**, these changes should be fully **explained and understood** before being accepted by scientists. For instance a major & sudden increase of fishing mortality should correspond to visible changes in fisheries or stock catchability.
- 8) **Too low or too high estimated MSY estimated by MFCL** should be accepted by scientists with great precaution: there is a frequent tendency of statistical models (WCPFC bigeye stock, Indian Ocean yellowfin) to estimate MSY that are quite unrealistic, being too low or too high<sup>4</sup>. When the MSY is estimated to be at a very low level, for instance 30 or 50% lower than the sustained catches observed during 15 years (twice the duration of the exploited life), this low MFCL MSY should not be accepted blindly by scientists, unless it can be explained by a well identified decline in yield per recruit or to a known change in the environment.
- 9) **Alternate simple stock assessment models:** Waiting the finalization of these “ideal” stock assessment models allowing a full use of statistical models, the alternate solution would be that **a wider range of stock assessment alternate models should also be used**<sup>5</sup>. Two types of models should be encouraged: PROCEAN type production models in their Bayesian flexible version, and ad hoc SPA models (ASPM type) analyzing the catch at age matrix & where the Input and output parameters are more « under control » than in the complex MCFL models. This positive recommendation has been already kept by the IOTC TT WG.
- 10) **Toward hand made Kobe plot by experts:** When the estimated trends and levels of the Kobe diagram estimated from the best MFCL model appears to be unrealistic during the critical recent years, tuna scientists should be encouraged to rebuild and to propose hand based alternate plots, and probably more realistic Kobe plots, based on their knowledge of experts and on the observed data.
- 11) **Projections: Scientists and WG should be** highly precautionary when doing **projections** of stocks & fisheries. The logical tendency to do these projections should

---

<sup>4</sup> These frequent errors in estimated MSY being mainly the results of the major uncertainties in the stock recruitment relationship & its steepness.

<sup>5</sup> MFCL model : keeping in mind the French rule... never put all your eggs in the same bag....

be encouraged, but most of these projections have been proven to be totally wrong because they are hampered by cascading & cumulative sources of variability, most often producing exponential increases of errors. These major errors in the projections should be better recognized & identified.

- 12) **CPUE & Biomass:** further in depth analysis are strongly required in order to fully recognize, to estimate & solve this major problem: allowing to incorporate in the future models (a) the early excessive declines of longline CPUEs<sup>6</sup> & (b) the excessive stability of the recent longline catch rates (probably widely due its increased efficiency), (c) to eliminate the vicious artificial effects of monofilament lines in the GLM results and (d) to fully explain the major divergence between Taiwanese and Japanese CPUEs (e) to incorporate the effect of total local efforts on CPUEs. All future assessment models should at least also try to estimate the potential effects of increases in the longline efficiency.

**Tagging symposium:** the organization of a large final scientific symposium, as it was recommended by the 2008 Scientific Committee, would be necessary to obtain quickly a fully useful analysis of the multiple IOTTP and positive results obtained by the RTTP-IO & by the various pilot and small-scale tagging operations recently conducted under the IOTC framework.

### \*\*\* Recommendations

- 13) All the **results of the tagging programs** (movements, growth, Mi, Exploitation rates) should be: (1) fully **analyzed by specialized team of experts** and (2) their results better introduced in the MFCL model. All future assessment analytical models should be based primarily on the best parameters estimated from external in depth analysis & they should never have a statistical freedom to select “statistically, blindly & often stupidly” their best parameters (growth, age specific movements, natural mortality).
- 14) **Better use of tagging results in stock assessment models:** Increased scientific methodological efforts should be conducted in order to allow to develop stock assessment **models allowing to make a fully efficient use** of the fantastic **tagging data** that are now available for the 3 species of Indian Ocean tropical tunas: fully combining the results of scientific analysis and the tagging data set itself! Ideally, this next generation “ideal” stock assessment should be conducted at a multispecies scale for the 3 tropical species tagged.
- 15) The use of **Ecosystem fishery models, such as the SEAPODYM model (Lehodey)**, a model efficiently used by WCPFC, should be developed by the IOTC in conjunction of statistical models such as MFCL, in order to provide estimates of densities and movement rates compatible with the structure and productivity of the various Indian Ocean ecosystems.

---

<sup>6</sup> Excessive early CPUE declines: these early declines should at least be explained: if it can be demonstrated that these declines were due to catchability & not to biomass declines, future assessment models could easily assume a given range of these + or - fixed recruitments (and subsequently measure decline the corresponding decline in longline catchability).