

**Progress and preliminary results of a study to develop robust stock status indicators for broadbill swordfish and tropical tunas**

**Marinelle Basson and Natalie Dowling**  
CSIRO Marine and Atmospheric Research  
Hobart, Tasmania  
Australia

## **1. Introduction**

This information paper reports on progress and provides some preliminary results from a three-year study, funded by the Australian Fisheries Research Development Corporation (FRDC) and the CSIRO, to develop robust stock status indicators for billfish and tropical tunas. The work is focused on the Indian Ocean and the fishery off the West Coast of Australia, but has wider general relevance. Previous papers reporting on this study and submitted to IOTC meetings are: IOTC Working party on billfish (Dowling and Basson, 2004) and Working party on Tropical Tunas (Basson and Dowling, 2004). These papers reported on the standardisation of indicators.

This paper presents results on the second phase of the study which uses simulation models to investigate:

- the relative sensitivity of indicators to life-history parameters
- the relative responsiveness of indicators to exploitation
- the time-series behaviour of indicators and their likely effectiveness in feedback decision rules

Given the evidence of potentially strong sexual dimorphism in growth of broadbill swordfish (e.g. Sun et.al., 2002), we have explored the behaviour of size-based indicators in this context, and we compare the behaviour of indices based on pooled data or data for males and females separately. Only a subset of results have been extracted for presentation here, and we have focused on those aspects most likely to be relevant to broadbill swordfish.

The study is now in its last phase and final results will be made available to interested parties and the IOTC<sup>1</sup>. Work in the final phase is focused on:

- the development and evaluation of decision rules based on indicators
- the development and evaluation of a frameworks for using multiple indicators in decision rules
- the evaluation of the performance of decision rules under different hypotheses about spatial structure.

It is worth briefly noting that although the work is being conducted within the context of the Indian Ocean, we are using a relatively general spatial model which can represent alternative hypotheses, rather than a specific 'Indian Ocean' operating model for billfish (or the tropical tunas). The main reason for this is the lack of direct data on movement and spatial structure. This work should, however, lay the groundwork for further evaluation of indicators once the

---

<sup>1</sup> Please contact the authors for further information: [marinelle.basson@csiro.au](mailto:marinelle.basson@csiro.au) or [natalie.dowling@csiro.au](mailto:natalie.dowling@csiro.au)

development of more realistic and data-based operating models for swordfish (or tropical tunas) in the Indian Ocean becomes possible.

Methods and results are presented under four headings:

- Theoretical and equilibrium considerations
- Equilibrium considerations with sexual dimorphism in growth
- Behaviour of time-series of indicators
- Behaviour of a simple indicator-based decision rule

## 2. Methods

Descriptions of methods have only been summarised here. The final report of this study would clearly contain full details. A note with respect to  $F_{msy}$  is necessary, however. We have used  $F_{msy}$ , and proportions or multiples of  $F_{msy}$ , in many of our explorations. The main reason for this is that the response of an indicator to exploitation is directly and strongly affected by the harvest rate relative to the life-history parameters. As such,  $F_{msy}$  is a convenient quantity which retains the scaling between harvest rate and life-history parameters and hence make results comparable. Another option would have been to use proportions or multiples of natural mortality,  $M$ , but given the importance of steepness in the stock-recruit relationship to sustainable harvest rates, we prefer to use  $F_{msy}$ . The fact that we have used  $F_{msy}$  in these explorations should not be interpreted as a suggestion that  $F_{msy}$  is an appropriate target harvest rate. The appropriateness of alternative harvest rates are best explored in the context of feedback decision rules using a range of performance measures, including the risk of low spawning biomass ( $SSB^2$ ).

### 2.1 Theoretical and Equilibrium considerations

The reason for considering equilibrium values of indicators under a constant harvest rate is to explore the sensitivity of indicators to life-history parameters and to evaluate their responsiveness to harvesting. An indicator, such as mean length or mean weight, is only likely to be useful if it is expected to change sufficiently in response to harvesting at a sustainable level, and if it is responsive to changes in biomass. If the indicator is monitored from the start of a new fishery on a virgin stock, the expected change would obviously be larger than for a stock which is already exploited. In both cases, however, it is informative to look at the expected values for the indicator (or the indicator relative to its value under no exploitation) under a sustainable harvest rate. A harvest rate of  $F_{msy}$  should provide the maximum change (compared to unexploited) that one would want to observe in an indicator, because the harvest rate should ideally not be above  $F_{msy}$  for any substantial amount of time. If the expected change in an indicator, when harvesting at  $F_{msy}$ , is small (for example 5%) then the chance of detecting that change before the stock is overfished or being overfished is smaller than if the expected change is larger (for example 30%).

Under deterministic and equilibrium assumptions, the relative change in an indicator is essentially a pure signal. There is no ‘noise’ from, for example, recruitment variability or sampling error, and an indicator that does not change by much for a large change in biomass under these conditions, are unlikely to perform well when stochasticity is introduced into the underlying dynamics.

---

<sup>2</sup> We use  $SSB$  and  $B$  interchangeably to indicate spawning biomass unless otherwise stated

The operating model for this part of the study is a very simple, age-structured, life-history driven model. The model has quarterly time steps, with recruitment occurring in quarters 1 and 2. The plus-group age is 20 years. This model has no spatial structure; that is, it comprises a single stock with no movement dynamics, and it is harvested by a single fleet.

Key input parameters are the steepness<sup>3</sup> ( $h$ ) defining the Beverton-Holt stock-recruitment curve, the growth rate,  $k$ , in the von Bertalanffy growth curve, natural mortality ( $M$ ) and ages at maturity ( $t_m$ ) and first capture ( $t_c$ ). The latter was assumed to be related to age-at-maturity as:  $t_c = t_m/2$ ,  $t_c = t_m$ , or  $t_c = t_m+1$ . We explored knife-edge selectivity at different ages, assuming that this would mimic selectivity of longline gear (though there is sometimes evidence of dome-shaped selectivity for longlining, as seen in some assessment results for Southern Bluefin Tuna). We also examined a ‘triangular’ selectivity-at-age pattern to mimic purse-seine fishing, but this is not presented here because it is not relevant to swordfish.

Mean length-at-age was calculated from the von Bertalanffy growth curve, and mean weight-at-age was derived from a power relationship to mean length-at-age<sup>4</sup>:

$$w_L = aL^b \text{ where } a = 0.00002, b = 2.9$$

Note that scaling quantities such as recruitment  $R$ ,  $L_\infty$ , and  $a$  in the weight-length relationship, are of little relevance because we are primarily interested in the change in an indicator at some harvest rate relative to its value at  $F=0$ . Setting these quantities at arbitrary values should therefore not affect results of relative change in indicators.

Data were generated by running the model to stability (30 years) under a given harvest rate and ‘sampling’ the whole population (i.e. no sampling error) to obtain size-based indicators of stock status. These indicators were compared with their values under no exploitation.

The following indicators were calculated:

- the mean, median, 95<sup>th</sup> percentile of length and weight distributions in the catch
- the proportion of “small” fish in the catch (defined as the proportion of fish  $\leq t_c+1$ )
- the proportion of “large” fish in the catch (defined as the proportion of fish  $\geq 170$  cm TL<sup>5</sup>)
- the proportion of mature fish in the catch (defined as the proportion of fish older than the age-at-maturity,  $t_m$ ).

Results presented here do not include the ‘proportion small fish’ indicator because this indicator increases rather than decreases when biomass decreases and requires somewhat different treatment and interpretation.

Scenarios comprising combinations of three alternative levels of the five key life-history input parameters were run (Table 1). That is, 243 scenarios for each harvest rate were considered. These spanned the range of possible extremes for tuna and billfish-like species, although it should be noted that some combinations of life-history parameters may be less likely in reality (eg a species is unlikely to have very low natural mortality,  $M$ , and a very high growth rate,  $k$ , or

---

<sup>3</sup> Steepness is defined as the ratio between recruitment ( $R$ ) when spawning biomass is at 20% of unexploited spawning biomass ( $B_0$ ), and recruitment at the unexploited spawning biomass i.e.  $h = R(0.2B_0) / R(B_0)$ .

<sup>4</sup> Historic data from Australian observers on Japanese longliners off the East Coast of Australia were used to obtain these parameter estimates.

<sup>5</sup> This is an arbitrary choice and more comments are made regarding the choice of cut-off size in the results section.

vice versa). For simplicity the  $t_0$  parameter in the von Bertalanffy growth curve was set to -1 for all these scenarios.

**Table 1:** Grid showing the three values of the input parameters evaluated: steepness ( $h$ ), von-Bertalanffy growth rate ( $k$ ), natural mortality, age at maturity ( $t_m$ ), age at first capture ( $t_c = t_m/2, t_m, t_m+1$ ).

$h$	$k$	$M$	$t_m$	$t_c$
0.3	0.2	0.1	2	$t_m/2$
0.6	0.4	0.2	4	$t_m$
0.9	0.6	0.4	6	$t_m + 1$

The harvest rates considered were: i)  $0.5 F_{msy}$ , ii)  $F_{msy}$ , and iii)  $2 F_{msy}$ , but results here focus on  $F_{msy}$  only.

## 2.2 Equilibrium considerations for sexual dimorphism in growth

In this section we include the complication of sexual dimorphism in growth and variability in length-at-age into the population model. It is highly likely that sexual dimorphism exists for Indian Ocean swordfish, although the extent and exact form of dimorphism in growth for Indian Ocean swordfish is not well known, as far as we are aware. The reasons for considering these models are the same as for the simpler deterministic models, but they should provide a somewhat more realistic reflection of the likely dynamics of swordfish.

The literature provides several estimates of growth for male and female swordfish from many oceans (see e.g. Radtke and Hurley, 1983; Tserpes and Tsimenides, 1995; Vanpouille et.al., 2001; Sun et.al., 2002). Young and Drake (2004) provide estimates for the west coast of Australia. We have used these ‘Young and Drake’ growth curves, which are in fact very similar for males and females, but we also constructed 3 ‘caricature’ or stylised hypothetical dimorphic growth curves to explore. We noticed from the curves in the literature, that male and female growth curves can be divergent, intersecting or parallel and we have constructed the ‘caricature’ curves to represent those three patterns (Figure 1, Table 2). Of main concern here is (a) whether sexual dimorphism implies that it is always better to consider male and female indicators separately, or whether pooled data can be used (noting the difficulties associated with collecting size data by gender) and (b) whether the type of dimorphism (divergent, intersecting, parallel) matters.

In the results presented here, selectivity for “longline-like” fishing was knife-edged at 120cm TL, as for the simple equilibrium models. Instead of using a single length-at-age, however, length-at-(quarterly)-age for each gender was assumed to follow a normal distribution with mean as calculated from the von-Bertalanffy growth curve, and standard deviation ( $\sigma_a$ ) a linear function of the square root of the (annual) age  $a$ :  $\sigma_a = 5.4 \cdot \sqrt{a/4} + 6.5$ . This relationship was loosely based on information in Young and Drake 2004. Selectivity at age for each gender was then calculated as the integral of this normal distribution above the knife-edge length of 120cm TL. As such, the age at first capture depends on growth curve and gender.

Reproductive studies by Young and Drake (2002) reported size at 50% maturity for females to be approximately 200cm. For the Young and Drake growth curve (2004) this corresponds to an approximately 9 - 10 year old female, but for the stylised growth curves, this length corresponds to lower ages (3-6 years). As such, an age at maturity of 5 was used across all growth curves, for

males and females, but for the Young and Drake growth curve, an additional model was run where the age at maturity was set to 9 years<sup>6</sup>. The model assumes knife-edged maturity at the gender-specific length corresponding to the age-at-maturity (5 years or 9 years). As with selectivity, the proportion mature-at-(quarterly)-age was then calculated for each gender as the integral of the normal distribution of length-at-age above the knife-edge length at maturity.

The sex ratio among recruits was assumed to be equal (50% per gender). Values for steepness in the stock-recruitment relationship were  $h=(0.3, 0.6, 0.9)$  and for natural mortality were  $M=(0.1, 0.25, 0.3)$ . Results presented here are only for the mid-points:  $h=0.6$  and  $M=0.25$ . These two parameters are likely to be the least well known of the life-history parameters and in the study as a whole, we therefore consider the ranges of values.  $F_{msy}$  was calculated for the combined catch of males and females.

The following indicators were derived from the “sampled” catch, in addition to those mentioned above:

- the mean length of “large” ( $\geq 170\text{cm TL}$ ) fish in the catch
- the mean weight of “large” ( $\geq 170\text{cm TL}$ ) fish in the catch

Indicators were calculated both separately for males and females, and for the catch sample pooled across genders.

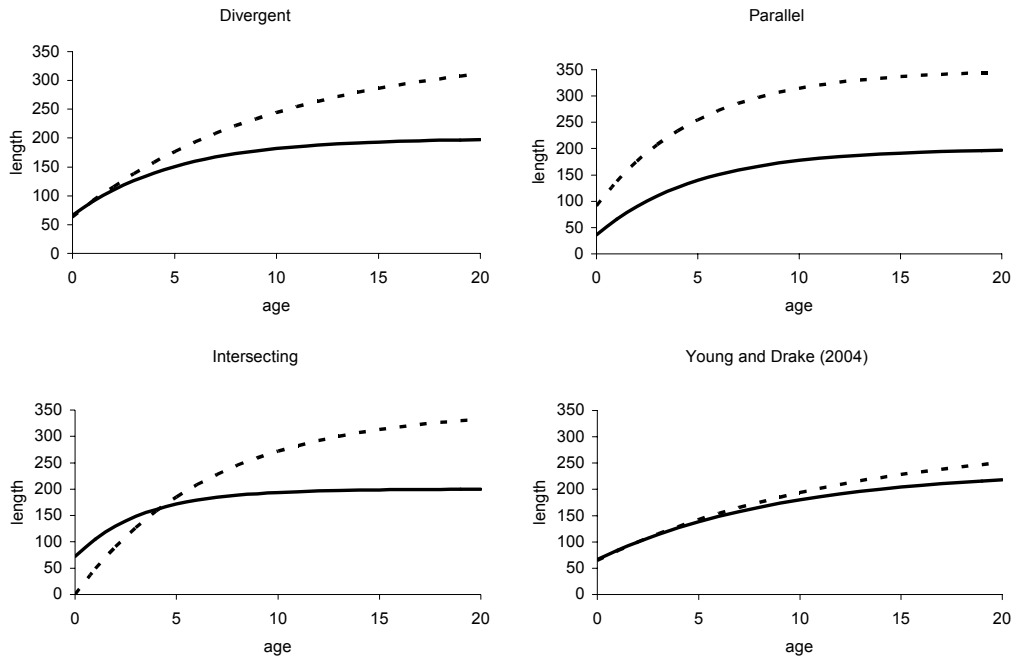
To summarise, the five scenarios considered in this section are:

1. Divergent growth,  $t_m=5$
2. Parallel growth,  $t_m=5$
3. Intersecting growth,  $t_m=5$
4. Young and Drake growth,  $t_m=5$
5. Young and Drake growth,  $t_m=9$

**Table 2:** Von Bertalanffy growth parameter values for the stylised sexually dimorphic growth curves, and the parameter values reported by Young and Drake (2004) for swordfish from the west coast of Australia

	Female			Male		
	$L_\infty$	k	t0	$L_\infty$	k	t0
Young and Drake (2004)	296.51	0.0815	-3.0148	236.9	0.1096	-3.0118
Divergent	350	0.1	-2	200	0.2	-2
Parallel	350	0.2	-1.5	200	0.2	-1
Intersecting	350	0.15	0	200	0.3	-1.5

<sup>6</sup> There is clearly still large uncertainty about life-span and age at maturity for swordfish. For an assumed natural mortality rate of 0.25, which gives average life-span  $1/m$  of 4 years, maturity at age 9 (when only about 10% of the year-class is still expected to be alive under no fishing) seemed rather high. This is why we included the scenario of the Young and Drake growth with age-at-maturity of 9.



**Figure 1. Stylised growth curves for sexually dimorphic male (solid line) and female (dashed line) growth, together with the growth curves reported by Young and Drake (2004) for swordfish from the west coast of Australia.**

### 2.3 Behaviour of time-series of indicators

The models with sexual dimorphism in growth were next run with stochasticity introduced into recruitment (CVs of 5%, 20% and 40%) and with three sampling levels: good (100% of the catch sampled), poor (2% of the catch sampled) and very poor (100 individual fish sampled from each gender). For each model and sampling level, 100 replicates of 30 years each were run. The initial stock status was always unexploited, and levels of harvest were again  $0.5F_{msy}$ ,  $F_{msy}$  and  $2F_{msy}$ , though results here focus on  $F_{msy}$ .

#### *'No-change error' summaries*

The 30-year time-series of each indicator was analysed by taking increasing periods of time (first 5 years, first 10, 15 and 20 years) and estimating the slope of the indicator over time by simple linear regression. The number of times the slope is NOT significantly different from 0 is then counted and expressed as a % of the total number of replicates (100). A high count means that an indicator is unlikely to signal a change; a low count means that it usually picks up some signal. In the rest of this document, the number of non-significant slopes is called the 'no-change error'. Clearly, this approach does not reflect the magnitude of change – for example the mean length may only drop by 10% for a drop of, say, 60% in the SSB – but it is meant to show the difference in behaviour for different levels of (a) recruitment variability and (b) sampling (good, poor and very poor). It is again within the feedback rules that the issue of magnitude of change in an indicator relative to that in SSB will be addressed more fully.

This approach also makes a simplistic assumption of a linear change in indicators. Although the response is generally slightly curvilinear (changing faster at the start of exploitation and then

slowing down), the deviation from linearity is relatively small compared to that for SSB, which is very obviously non-linear. This approach is simply being used as a tool for comparison and summary.

#### *'GAM MSE' summaries*

The time-series of indicators for 100 realisations were also analysed with respect to their relative ability to predict changes in biomass. As noted before, the biomass changes in a very non-linear way, particularly at the start of exploitation on a virgin stock. This analysis was therefore done by fitting a general additive model (GAM) to the indicator and simulated biomass, as follows:

$$dB_t \sim s(dI_t) + s(I_t)$$

where

$B_t$  is the simulated spawning biomass at time  $t$  and

$dB_t = (B_{t+1} - B_t)/B_t$  is the relative change in simulated SSB from time  $t$  to  $t+1$

$I_t$  is the value of the indicator at time  $t$  and

$dI_t = (I_{t+1} - I_t)/I_t$  is the relative change in the indicator from time  $t$  to  $t+1$

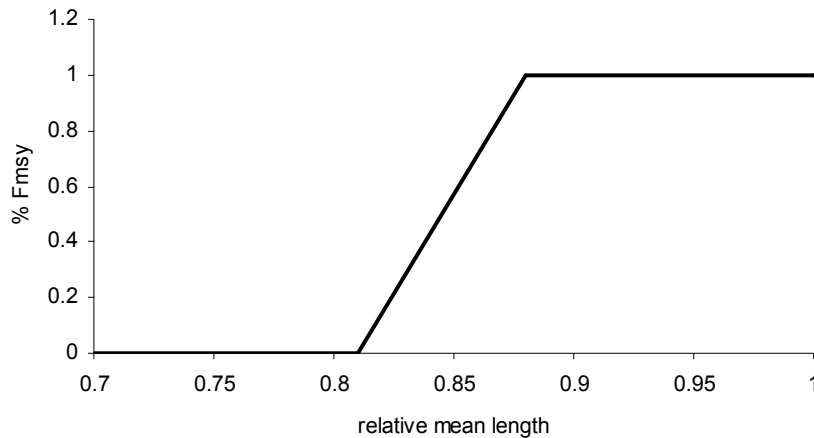
's' indicates a spline-based smooth in the GAM model

Both the change in the indicator and the level of the indicator is therefore used to predict the change in the SSB (this was found to be substantially better than just using the change in the indicator). Although this may not lead to the 'best' model fit in all cases, the approach is useful in a simulation context where it is impossible to fine-tune the fit for every realisation. This approach also leads to comparable results for all indicators and all realisations, and is therefore again a useful tool for comparison and summary. We use the mean squared error (MSE, i.e. the mean of the squared difference between the simulated SSB and model predicted SSB) as a measure of the performance of each indicator. Low values of the MSE indicate good performance.

#### 2.4 Behaviour of single-indicator decision rules

Simple decision rules based on mean length were investigated to determine how well a single indicator might perform when used to manage a fishery. While various rules will ultimately be considered, that presented here is a variation on a simple "40/10" rule (usually defined in terms of biomass or CPUE), whereby fishing mortality is constant above a threshold indicator level, zero below a limit indicator level, and declines linearly between the two indicator values. The indicator, in this example mean length based on pooled data, is calculated at the end of each year and the coinciding harvest rate is determined from the decision rule. This harvest rate is then applied to the population to generate catch in the following year. Although the decision rule is formulated in terms of fishing mortality, a conversion to a TAC can easily be made in the model. We assume that management is perfectly implemented (i.e. the harvest rate from the decision rule is applied to the population without error), and that estimates of quantities such as biomass are available (and unbiased) if management is through TAC. We recognise that this represents ideal and somewhat unrealistic conditions. The results presented are, however, intended to be comparative rather than absolute. We emphasise that the work on decision rules is still underway, and issues such as implementation error and the availability of estimates for key quantities are being considered.

In our case the decision rule was defined as shown in Figure 2. Fishing mortality was set to  $F_{msy}$  above a threshold value expressed as a fraction of the expected mean length of the unexploited stock. This fraction was chosen to be 0.88, corresponding to the ratio of the expected mean length at  $F_{msy}$  to the ‘virgin’ mean length for a population growing according to the Young and Drake von Bertalanffy growth parameters. The limit indicator level was set at the (arbitrary) relative mean length of 0.81, corresponding to a length of 130cm (i.e. 10cm above the knife-edged selectivity length assumed in the simulations).



**Figure 2. Schematic illustration of the simple decision rule based on mean length (relative to virgin mean length)**

Additional scenarios were run where the threshold indicator level was increased to a relative mean length of 0.93. Under this decision rule, fishing mortality was lowered earlier and by a greater amount relative to the previous threshold value. The decision rule was invoked after 5 years of fishing at  $1.6 F_{msy}$ . This level of fishing mortality was chosen so that the impact on the stock after 5 years would be substantial.

Scenarios were based on the two sexually dimorphic growth curves showing the most extremes in  $F_{msy}$ . Of the four growth curves considered, populations with the Young and Drake growth curve had the highest  $F_{msy}$  (0.235) while populations with the stylized “parallel” sexually dimorphic growth had the lowest  $F_{msy}$  (0.145). Note, however, that the relative thresholds, i.e. inflection points, for the decision rule remained at the values described above.

In order to gain insight into how a decision rule might behave when the population grows in a different way to that assumed in the construction of the decision rule, scenarios with the “correct” decision rule (i.e.  $F_{msy}$  coinciding with the underlying growth curve), as well as with a “wrong” decision rule (i.e.  $F_{msy}$  NOT coinciding with the underlying growth curve) were run (see below).

As a reference case, and for comparison, one scenario had no feedback decision rule and the population was harvested at  $1.6F_{msy}$  for the entire 30 years:

Scenario 1: Young & Drake,  $1.6F_{msy}$  (based on Young and Drake)

Six scenarios with decision rules were run. The scenarios are in pairs, with one using the lower threshold (0.88, as in Figure 2), the other using the higher (0.93) threshold, and with the following combinations of growth curve in the stock dynamics and  $F_{msy}$  in the decision rule:



Scenarios 2 and 3: Young & Drake growth with its own  $F_{msy}$   
 Scenarios 4 and 5: Parallel growth curve with its own  $F_{msy}$   
 Scenarios 6 and 7: Parallel growth curve with the Young and Drake-based  $F_{msy}$

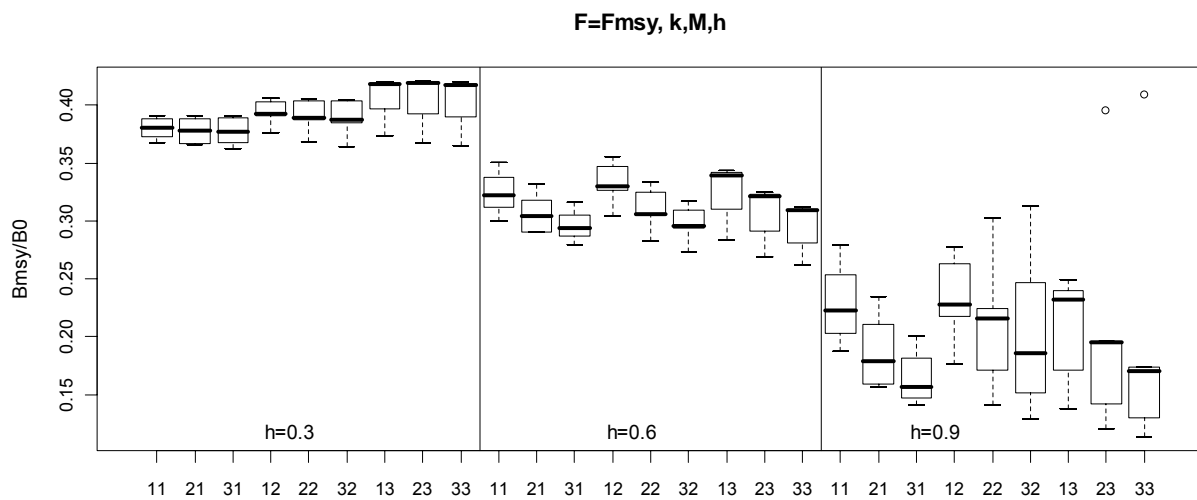
Results presented here are based on 100 replicates of 30 years with steepness of 0.6, natural mortality of 0.25, and the standard deviation for recruitment 0.4. Perfect (100%) catch sampling was assumed, and all outputs considered were those pooled across gender.

### 3. Results

#### 3.1 Theoretical and Equilibrium considerations

This part of the study is based on deterministic stock dynamics to explore the effect of life-history parameters (von Bertalanffy growth rate ( $k$ ), natural mortality ( $M$ ), steepness of the stock-recruit relationship ( $h$ ), age at maturity and the age/size at first capture) on the magnitude of expected change in indicators and spawning biomass when harvesting at  $F_{msy}$ .

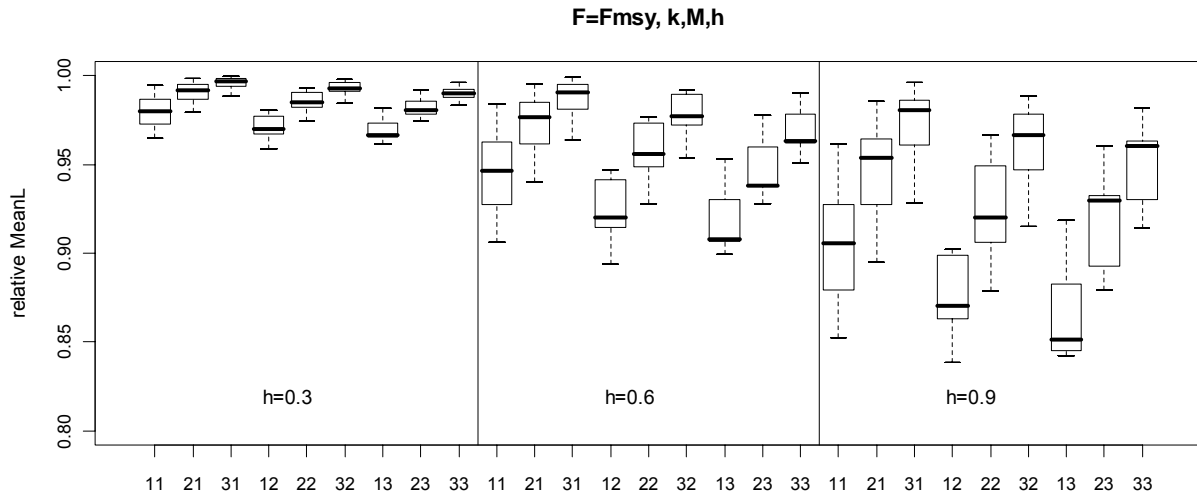
Steepness, which is likely to be the least well known quantity for a stock, has a strong effect on  $B_{msy}/B_0$  (Figure 3) because it has a strong effect on  $F_{msy}$  (higher steepness can support a higher  $F_{msy}$ ).  $B_{msy}/B_0$  is also generally lower for high values of  $k$ , but the effect of steepness still dominates. The ‘boxes’ in Figure 3 reflect different age at maturity and age at first capture, and some combinations of these quantities with  $M$  and  $k$  are omitted because they are unlikely – e.g. high  $M$  with high age at maturity or capture). The figure is only intended to provide a general impression of which factors have the most effect on  $B_{msy}/B_0$ .



**Figure 3.  $B_{msy}/B_0$  (relative spawning biomass) for different values of  $k$ ,  $M$  and steepness ( $h$ ). Numbers on the x-axis refer to  $k-M$  pairs with  $k=(1=0.2, 2=0.4, 3=0.6)$  and  $M=(1=0.1, 2=0.2, 3=0.4)$ . Factors reflected in the ‘boxes’ are age at maturity and age at first capture.**

A similar plot for relative mean length (Figure 4) again shows the strong effect of steepness, but also shows that the effect of  $k$  on the change in mean length is strong. The expected change in mean length is largest at high steepness and when  $k$  is small (i.e. slow growth), particularly in combination with relatively high  $M$  (e.g.  $k=0.2, M=0.4$ ; we again note that, from a biological

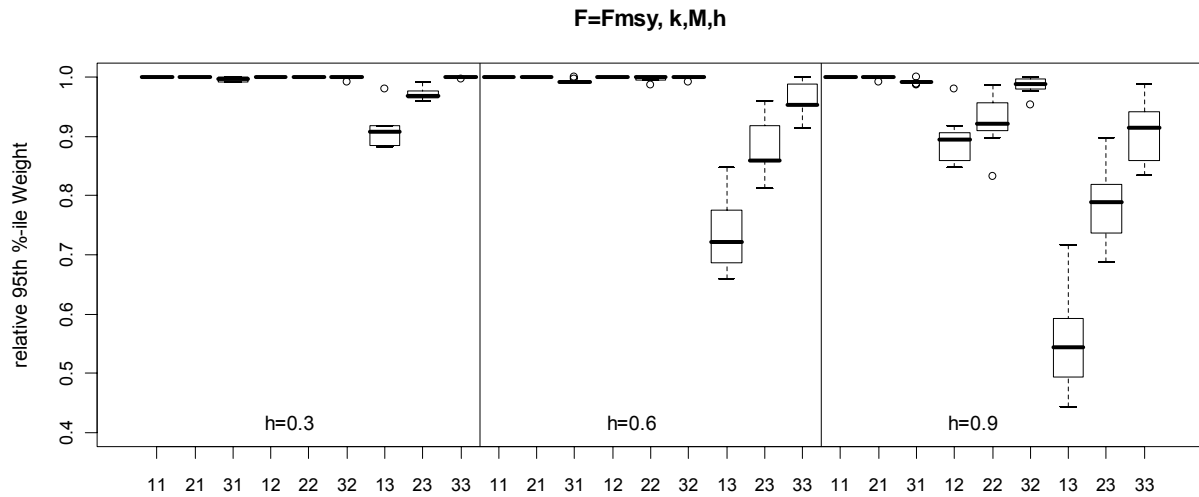
point of view, some of these combinations of very low  $M$  with very high  $k$ , or vice versa, may be less likely than combinations where  $M$  and  $k$  are more similar). The large expected change for LOW values of  $k$  is in contrast to the pattern shown in Figure 1 for relative spawning biomass. Recall that high  $k$  would tend to imply many older age classes with similar mean length. A change in the overall mean length of the catch would therefore only be evident once the mean age has been reduced quite substantially. In comparison, a low  $k$  tends to imply increasing size for many older age classes, so that a relatively small change in mean age could reflect a change in overall mean length of the catch.



**Figure 4** Relative mean length ( $\text{meanL}(F_{msy})/\text{meanL}(F=0)$ ) for harvest at  $F_{msy}$  given 3 steepness ( $h$ ) levels. Numbers on the x-axis refer to  $k$ - $M$  pairs with  $k=(1=0.2, 2=0.4, 3=0.6)$  and  $M=(1=0.1, 2=0.2, 3=0.4)$ . Factors reflected in the ‘boxes’ are age at maturity and age at first capture.

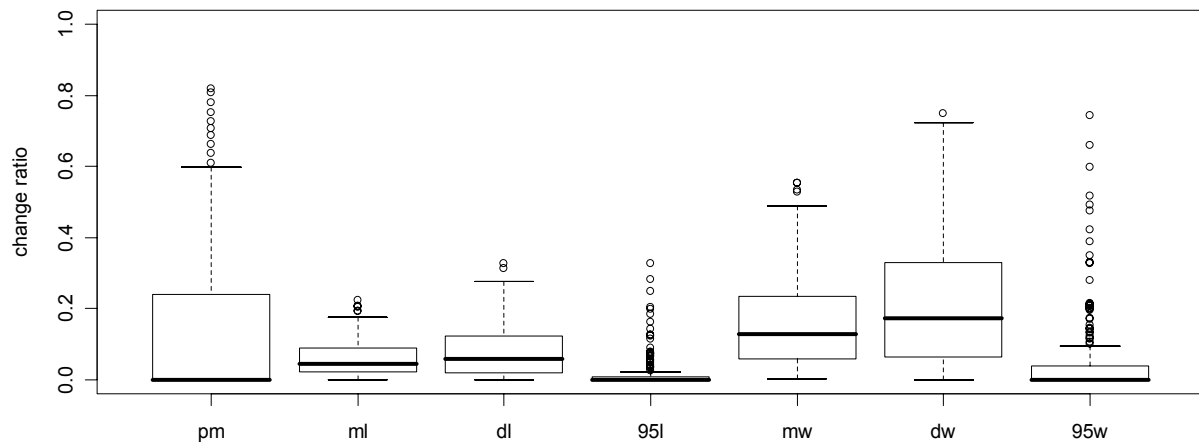
Figure 4 also implies that, if a large change in mean length is observed for a stock with high growth rate, particularly in combination with low mortality rate, then this should be cause for concern, because it is likely that the harvest rate is already too high.

Figures for other indicators are similar to that for relative mean length, though it is worth showing the figure for the 95<sup>th</sup> percentile of weight (Figure 5). This indicator shows particularly poor response (lack of response) at low steepness, low  $k$  and low  $M$ . It is only really at high steepness and high mortality, particularly with slow growth, that this indicator shows some change.



**Figure 5** Relative 95<sup>th</sup> percentile of length ( $L95(F_{msy})/L95(F=0)$ ) for harvest at  $F_{msy}$  given 3 steepness ( $h$ ) levels. Numbers on the x-axis refer to  $k$ - $M$  pairs with  $k=(1=0.2, 2=0.4, 3=0.6)$  and  $M=(1=0.1, 2=0.2, 3=0.4)$ . Factors reflected in the ‘boxes’ are age at maturity and age at first capture.

The scenarios we looked at lead to ratios of spawning biomass  $B_{msy}/B_0$  of between 0.15 and 0.40. The relative mean length at  $F_{msy}$  is, however, between 0.85 and values close to 1 (hardly any change; see Figure 4), showing that mean length is not a particularly responsive indicator. This is also illustrated in Figure 6 where the change in indicator is expressed relative to that in SSB:  $(1 - \text{Indicator}(F_{msy})/\text{Indicator}(F=0)) / (1 - B_{msy}/B_0)$ . An ideal indicator would have large values, close to or above 1 for this performance measure, i.e. the indicator changing almost as much or more than SSB. An indicator with low values, i.e. changing much less than SSB, is likely to be poor in terms of detecting the signal when stochasticity in underlying dynamics and sampling error come into play.

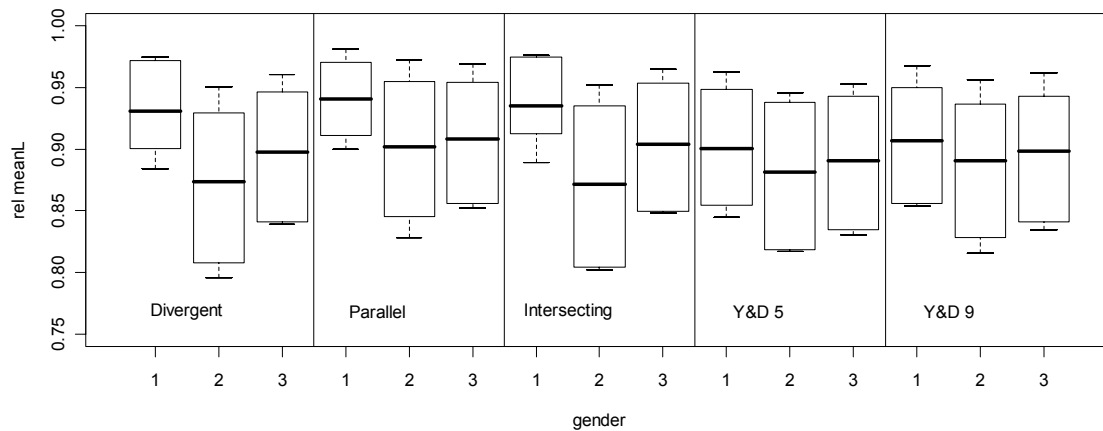


**Figure 6** Change in indicator relative to change in biomass,  $[1 - \text{indicator}(F_{msy})/\text{indicator}(F=0)]/[1 - B_{msy}/B_0]$ , for 7 size-based indicators: proportion mature, mean length, median length, 95<sup>th</sup> percentile of length, mean weight, median weight and 95<sup>th</sup> percentile of weight.

Indicators based on weight are generally more responsive than those based on length. This result is not new; see for example Punt et.al. (2001). The theoretical considerations also suggest that median weight is slightly better than mean weight, but the 95<sup>th</sup> percentile is only adequate in some cases as already shown (Figure 5). The proportion mature is variable in its performance, depending on the size/age at first capture relative to the size/age at maturity. Obviously, if age at first capture is greater than age at maturity, then the proportion mature is not informative, and those cases are included in Figure 6.

### 3.2 Equilibrium considerations for sexual dimorphism in growth

In this section we consider the added complexity of sexual dimorphism in growth, using the growth curves illustrated in Figure 1. It is not surprising that the same life-history parameters that were important when no sexual dimorphism in growth was assumed, are important here. The suite of growth curves considered generally have lower growth rate but higher  $L_{inf}$  for females than males, and that leads to more response in indicators based on data for females than indicators based on data for males. The reverse would be true if both  $L_{inf}$  and  $k$  for females were higher than for males. The indicator values based on pooled data (males and females) are intermediate. They are not, however, necessarily halfway between values for the male and female indicators.



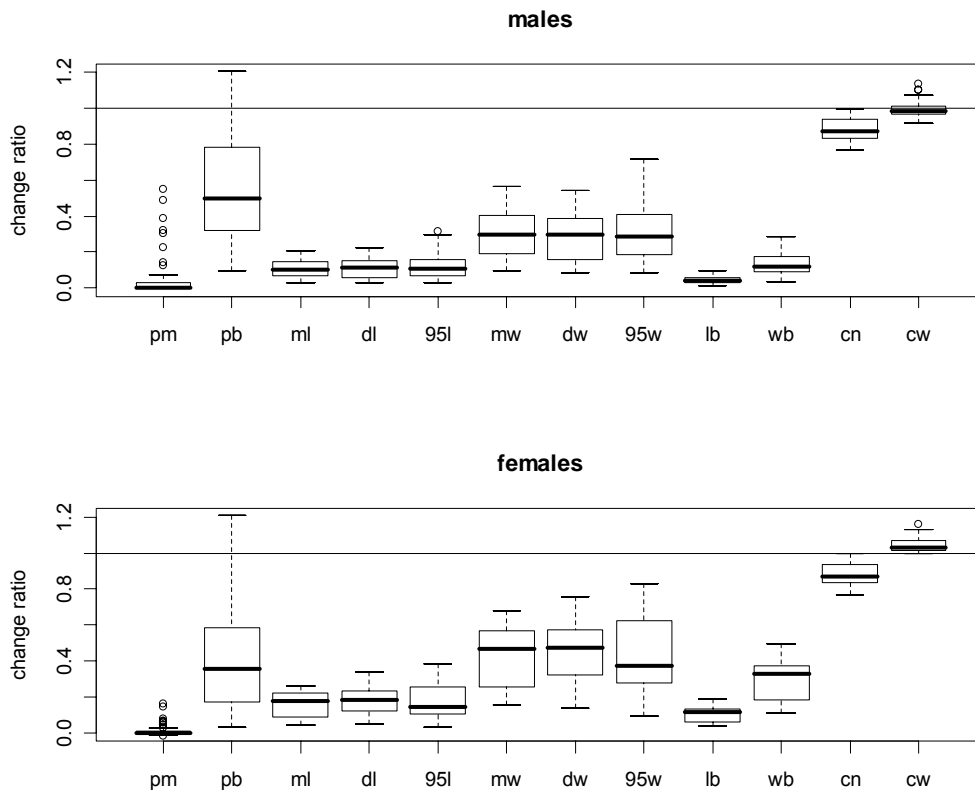
**Figure 7. Expected values for relative mean length:  $\text{meanL}(F_{msy})/\text{meanL}(F=0)$  by growth curve and gender. Other factors are steepness,  $m$ , and implicitly age at first capture noting that size at first capture is identical in all cases. Gender is indicated by: 1=Male, 2= Female, 3=Pooled. Y&D refers to the Young and Drake growth curves with age-at-maturity of 5 years (Y&D 5), or 9 years (Y&D 9).**

Absolute mean length of females would be larger than that for males, but the relative change in this indicator is bigger (hence lower values in Figure 7) for females than for males. The parallel growth curves have the same  $k$  for males and females (Table 2), but different  $L_{inf}$  and  $t_0$  parameters. Results in Figure 7 show that here these two quantities also play a role in the relative response of size-based indicators through their interaction with a single SIZE at first capture (and hence a different AGE at first capture for males and females).

It is also interesting that for the parallel growth curve, the pooled indicator behaves quite similarly to the female indicator. This is probably mainly due to the much lower age at first

capture for females than for males which presumably affects the sex ratio in the fishable population. Results suggest that under some growth patterns (e.g. divergent and intersecting) it could be beneficial to consider indicators based on data for females only, though in other cases (e.g. parallel, or Young and Drake which has a low degree of dimorphism) there is apparently not much to be gained, in terms of responsiveness of the indicator, by considering gender-specific indicators. This analysis does not, however, take into account issues such as spatial structure by size or gender, and in practice there may be other important reasons why size data should be collected, and indicators constructed, by gender, if at all possible.

Other size-based indicators show similar patterns to those for relative mean length (Figure 7), though the 95<sup>th</sup> percentile can sometimes be more responsive for pooled rather than female data (not illustrated). This result should, however, be treated with caution, and the performance of the 95<sup>th</sup> percentile as indicator should rather be evaluated within the context of stochasticity, sampling and feedback decision rules.

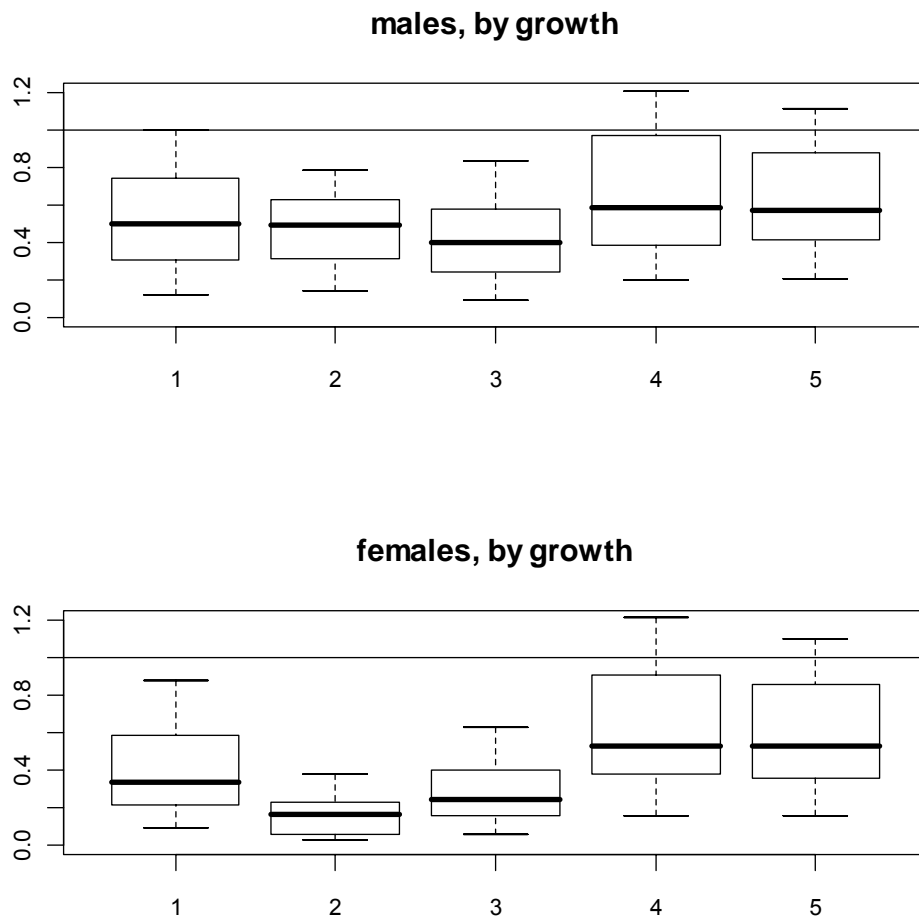


**Figure 8** Change in indicator relative to change in biomass,  $[1 - \text{indicator}(F_{msy})/\text{indicator}(F=0)]/[1 - \text{Bmsy}/B_0]$ , by gender, for 11 size-based indicators: proportion mature, proportion big, mean length, median length, 95<sup>th</sup> percentile of length, mean weight, median weight and 95<sup>th</sup> percentile of weight, mean length of big fish, mean weight of big fish; and two CPUE indicators: cpue in numbers and cpue in weight.

A comparison of the expected change in indicators relative to the change in biomass (as in Figure 6) is shown for the 5 scenarios combined in Figure 8. Although the responsiveness of a given indicator can differ between males and females, the relative responsiveness of indicators remains essentially the same (e.g. mean length is less responsive than mean weight for both males and

females). Recall that quantities near 1 reflect a responsive indicator (values >1 imply that the indicator changed more than the biomass).

It is worth commenting on the ‘proportion big’ indicator. In all cases this was based on the proportion bigger than 170cm. The choice of cut-off will have quite different implications for the different growth curves, and the results here, which suggest this indicator is only sometimes informative, should therefore be interpreted with caution. Figure 9 (below) shows the change in the ‘proportion big’ relative to change in SSB separately for each growth curve. This highlights that the wide range of performance is primarily driven by growth curve (and what the choice of size implies for that growth curve). It is likely that the cut-off used in constructing this indicator would have to be chosen with care.



**Figure 9** Change in the ‘proportion big’ indicator relative to change in biomass,  $[1 - \text{indicator}(F_{msy})/\text{indicator}(F=0)]/[1 - B_{msy}/B_0]$ , by gender and growth curve: 1=divergent, 2=parallel, 3=intersecting, 4=Young and Drake with age-at-maturity=5, and 5=Young and Drake with age-at-maturity=9.

### 3.3 Behaviour of time-series of Indicators

In this section we explore the behaviour of the time-series of an indicator as a stock is exploited at a constant harvest rate, starting from an unexploited state. The equilibrium considerations in the previous section were done without stochasticity in the dynamics, but here variability in recruitment is introduced to the scenarios with sexually dimorphic growth. We explored different harvest rates ( $0.5F_{msy}$ ,  $F_{msy}$  and  $2F_{msy}$ ) and different degrees of variability in recruitment (CVs of 5%, 20% and 40%) in the study, but only highlight a few results here. We note that the assumption of starting in an unexploited state may be rather unrealistic. Nonetheless, this provides insight into the extent of response that could be expected under such an ‘extreme’ scenario. More importantly, it provides a reasonable comparison between indicators. Further explorations of the performance of indicators in feedback decision rules are being conducted under alternative assumptions of starting state of the stock and of build-up of effort at the start of a fishery.

Some results are not surprising. For a given set of life-history parameters, the ‘no-change error’ (see Methods section for definition):

- decreases as  $F$  increases,
- decreases when sampling level improves or increases,
- decreases as recruitment variability decreases

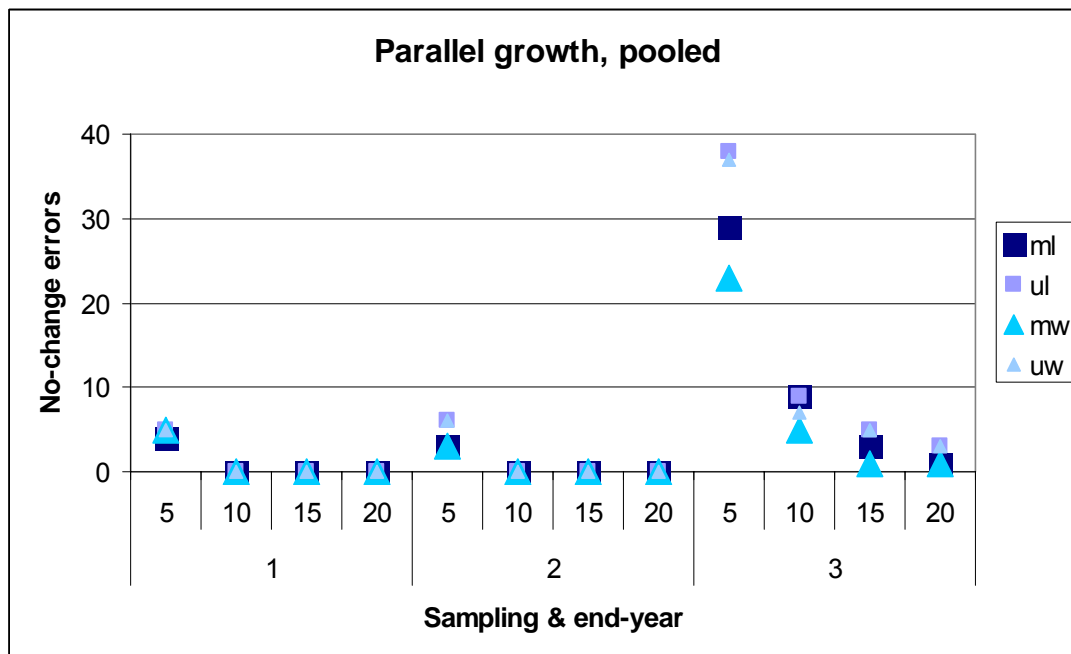


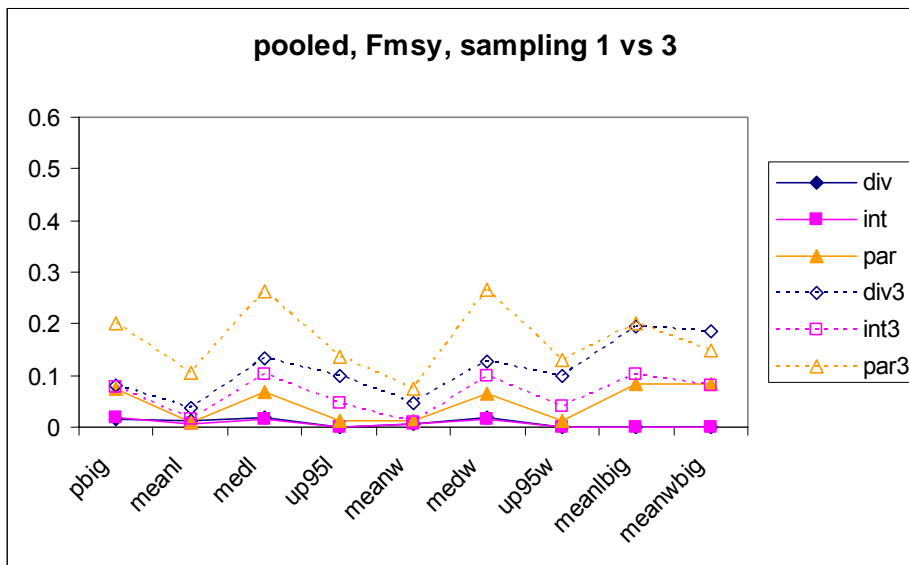
Figure 10. Percentage ‘no-change errors’ (from 100 runs) for ml=mean length, ul=95<sup>th</sup> percentile of length, mw=mean weight and uw=95<sup>th</sup> percentile of weight for 3 levels of sampling (1=good, 2=poor, 3= very poor) and 4 lengths of the time-series, all starting in year 1. The indicators are based on pooled data, the parallel growth curve was used with a recruitment CV of 40% and harvesting at  $F_{msy}$ .

It is also not surprising that some indicators are more sensitive to the above factors than others. For example, the 95<sup>th</sup> percentile of length or weight is more sensitive to sampling level than mean length or weight, particularly over a short initial time-period. Figure 10 shows that when

sampling is very poor (level 3), a significant change in the indicator is not detected in almost 40% of cases when the first 5 years' data are considered. This improves to about 10% when 10 years' data are used. One would expect indicators based on the 'tails' of distributions (e.g. the 95<sup>th</sup> percentile, or the proportion of individuals above some large size) to be more sensitive to sampling level than quantities such as the means.

Results with respect to gender-specific indicators are, however, less obvious. For the example based on pooled data and 'parallel' growth shown in Figure 10, the 95<sup>th</sup> percentile performs even more poorly when based on size data for males only. When based on data for females only, it appears to perform better than the mean length. It is not yet clear why this is the case – it could be spurious and due to the relatively small number of replicates (100 time series), it could be spuriously poor fits of the linear model in the case of the data for males only, or it could be a real effect. We note that this only occurs for the parallel growth curve, and we are exploring this further.

To consider the relative sensitivity of different indicators to sampling level, we have calculated the total 'no-change' errors over all the time-periods (5,10,15, and 20 year periods) and expressed them as a proportion. When indicators are based on pooled data, the proportion big, median length, median weight and the mean length and weight of big fish are particularly sensitive to sampling level, and more so for the parallel growth curves than the divergent or intersecting (Figure 11).



**Figure 11. Overall proportion of 'no change' errors for pooled data, over periods of 5,10,15 and 20 years, for harvesting at Fmsy and for the 3 growth curves with good sampling (1, indicated by 'div', 'int', 'par') and very poor sampling (3, indicated by 'div3', 'int3', 'par3'). See text for more detail on sampling level.**

The above figure also shows that the median length and weight appears to perform more poorly overall than the mean and the 95<sup>th</sup> percentile (NOTE here results have been combined over all time-periods, whereas Figure 10 shows performance separately for different time-periods, highlighting that the 95<sup>th</sup> percentiles perform better over longer time-periods).



The ‘no-change errors’ for gender-specific indicators show a somewhat different picture. Figure 12 shows that the ‘males only’ indicators (top panel) are much more sensitive to poor sampling than the ‘females only’ indicators (bottom panel).

This is likely to be a combination of the higher growth rate, lower  $L_{inf}$  for males than females, and the expected sensitivity of indicators based on the ‘tails’ of the distribution to poor sampling. It is not surprising that the means (length and weight) are the least sensitive of the suite of indicators shown.

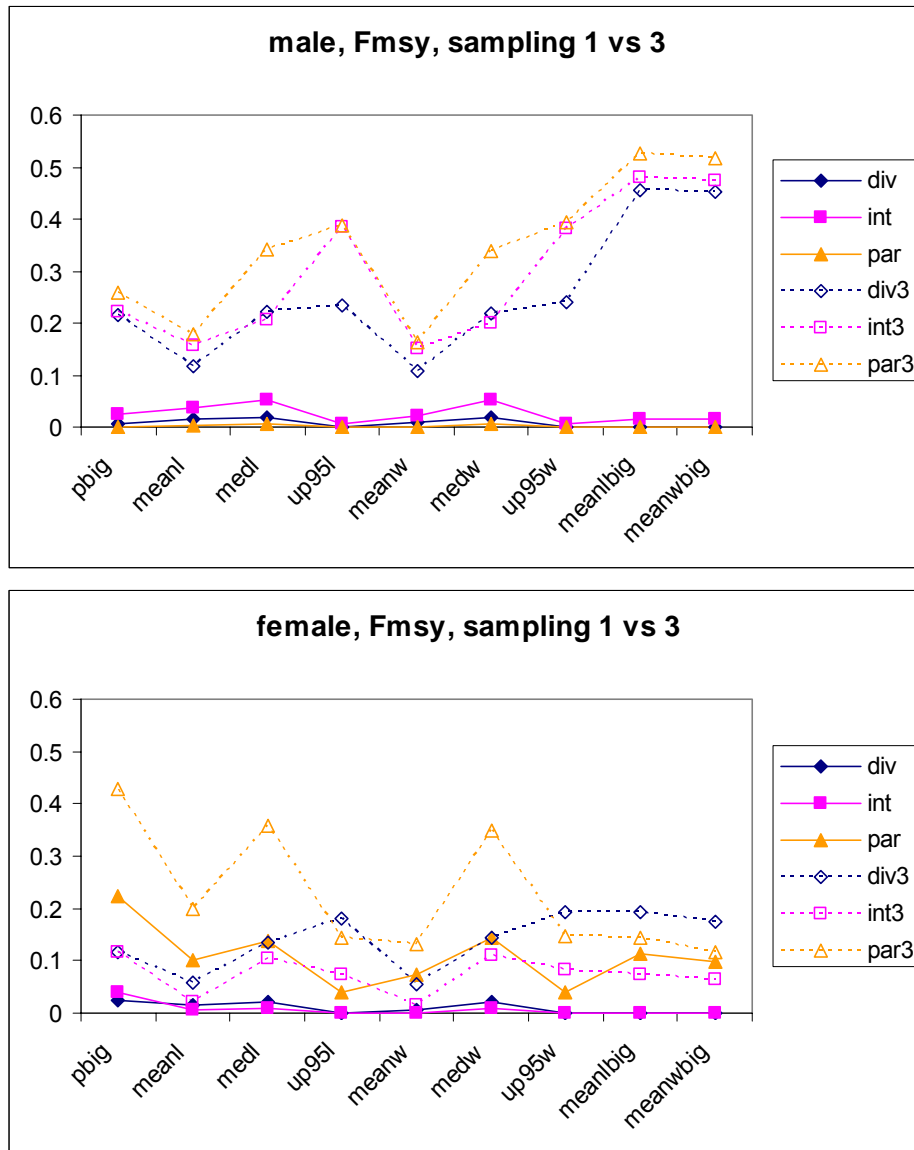


Figure 12. As Figure 9, but separately for males and females.

It is also interesting to note that the ‘female’ indicators under the parallel growth curve perform more poorly than under the other growth curves even when sampling is good. As mentioned above, the main reason for this is not obvious, and it is being investigated further.

We also considered how well the time-series of indicators can predict changes in spawning biomass, using the mean squared errors from GAM models (see methods section). For each indicator, 100 replicates of 30-year time-series were used, but the analysis could also be done on

shorter subsets of the time-series. The mean squared error (MSE) over all replicates is used to evaluate performance; low values of the MSE suggest a potentially good indicator.

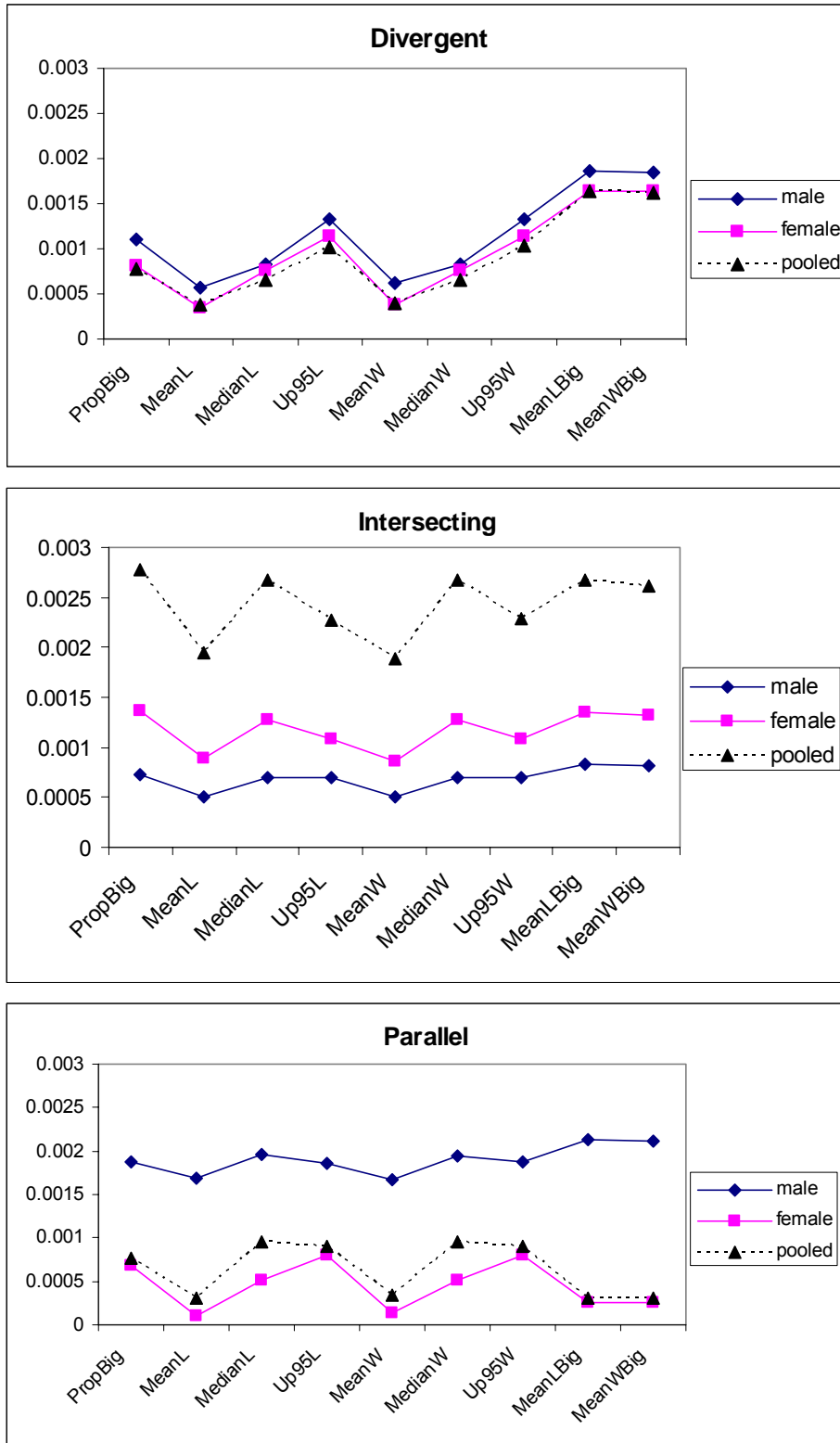


Figure 13. Mean squared error from GAM models for indicators by gender and pooled data. Other assumptions are: recruitment variability CV of 40%,  $F=F_{msy}$  and poor sampling (level 2).

Results (Figure 13) show the differences between the growth curves and choice of gender-based or pooled-data indices. There is little difference for the divergent growth pattern, but for the intersecting curves, the indicators based on pooled data perform more poorly than either of the gender-based indices. For the parallel growth pattern, however, the pooled indices perform much like the female-based indicators, and both perform better than indicators based on the data for males only. Whether the differences in MSE shown above are sufficient to have a strong effect on the performance of indicators in a feedback decision-rule context still has to be explored. The results confirm the good performance of mean length and mean weight relative to the other indicators. In some cases the mean length and/or weight of large fish also perform well, but as noted before, the performance of this indicator is sensitive to the choice of cut-off relative to growth parameters.

### 3.4 Behaviour of single-indicator decision rules

Summary results of the decision-rule scenarios, in terms of the change in the indicator and performance of catch and spawning biomass, are shown below. All quantities are relative to their initial values, except for the number of years SSB is below  $0.4 \cdot \text{SSB}_0$ , and catch which is expressed relative to the catch at  $\text{MSY}^7$  for the growth curve used in the population dynamics. Although it is customary to consider the proportion of time that spawning biomass is below 20% of unexploited SSB as an indication of risk, these proportions were almost always zero for scenarios other than scenario 1. Instead, we have shown the number of years (out of 30) that spawning biomass is below 40% of unexploited spawning biomass. Given that  $B_{\text{msy}}/B_0$  is around 0.35, this statistic should not be over-interpreted since one would expect the stock to be below 40% $B_0$  for some years without necessarily being over-exploited.

Figure 14 shows the strong effect a decision rule based on an incorrect assumption about growth can have on the performance of the stock. For scenario 6 where growth is ‘parallel’ but the decision rule is based on  $F_{\text{msy}}$  associated with the Young and Drake growth curves, the relative spawning biomass is almost as low as for the unmanaged and over-exploited scenario 1 (where  $F=1.6F_{\text{msy}}$ ). When the threshold or inflection point of the decision rule is increased, there is a slight improvement (scenario 7) in relative biomass, but it is not as good as for the case where the Parallel growth curve is used with the decision rule based on its associated  $F_{\text{msy}}$  (Scenarios 3 and 4).

An alternative way of looking at results is in terms of the trade-off between relative spawning biomass and catch relative to  $\text{MSY}$  (Figure 15). This figure more clearly shows that, in these examples, relative catch changes more slowly than relative SSB (e.g. there is less difference between relative catch for the 7 scenarios than for the relative SSB). It is therefore not necessarily the case that a large amount of catch needs to be forgone for an increase in spawning biomass which could imply a much lower risk of low SSB and possible recruitment failure.

---

<sup>7</sup> Since the actual level of recruitment is entirely arbitrary, the absolute levels of biomass and catch are not meaningful, and it is therefore necessary to rescale quantities here. Under circumstances where stock-recruit information was available from an assessment, for example, one would consider unscaled catch.

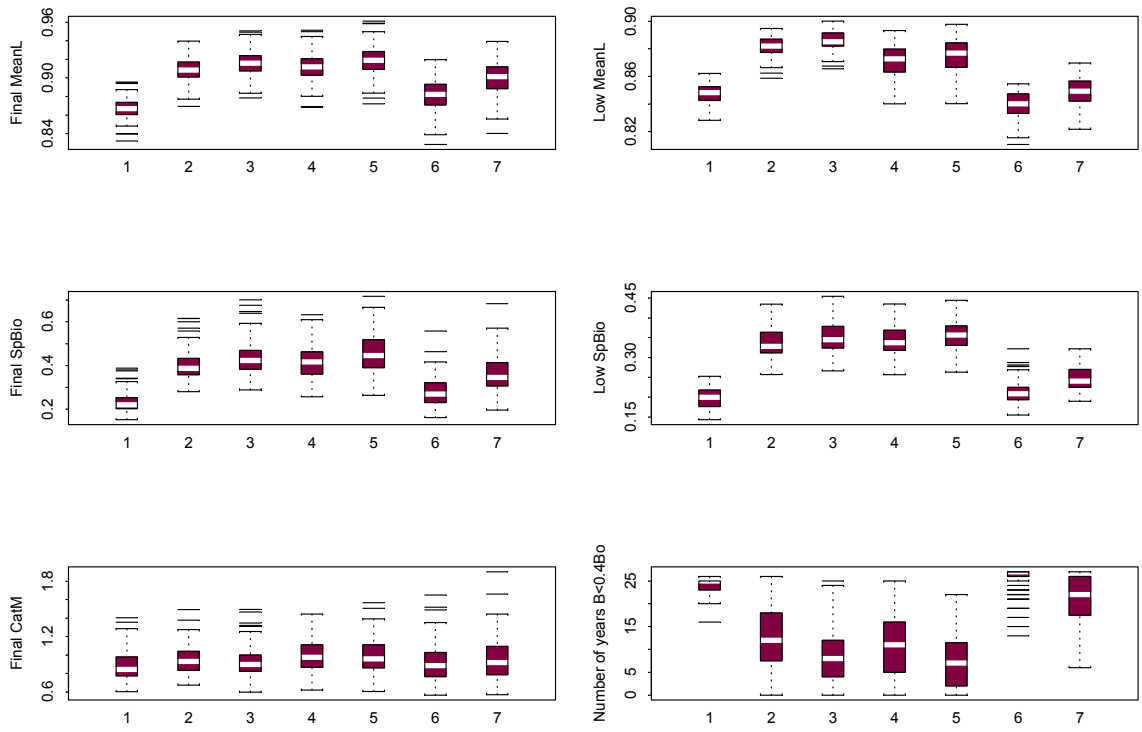


Figure 14. Summary results for 100 replicates of the 7 scenarios (numbered on the x-axis; see text for definitions). All quantities are relative to their initial values (unexploited stock), except for final catch which is relative to the MSY of the underlying growth curve, and last panel which is the number of years out of 30 that SSB is below 40% of unexploited SSB.

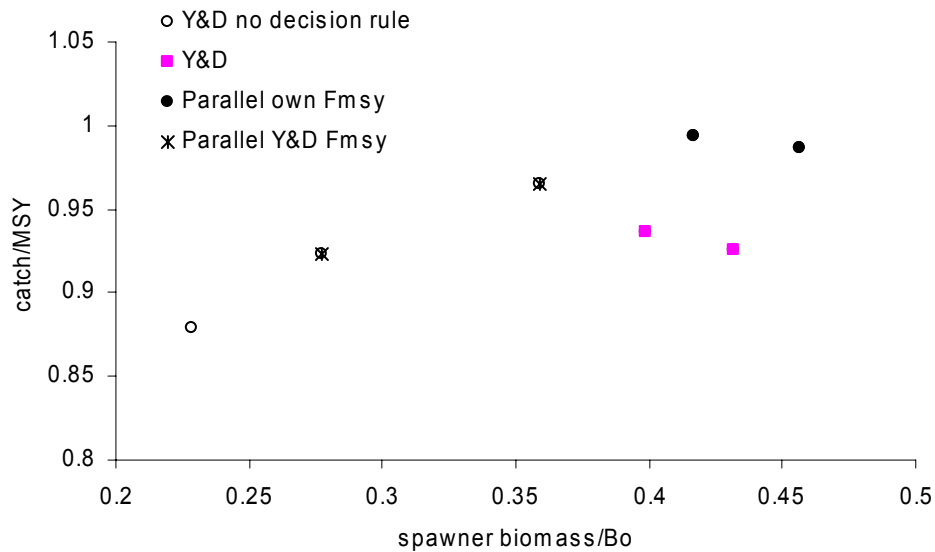


Figure 15. Trade-off plot in terms of spawning biomass ( $B/B_0$ ) and catch ( $Catch/MSY$ ), averaged over the 100 replicates, for the 7 scenarios (see text). The right-hand points in each pair are for the decision rule with higher threshold (0.93), those on the left, for the decision rule with lower threshold (0.88).

## Discussion

It is not surprising that the extent to which size-based indicators are likely to reflect changes in biomass, is strongly driven by the harvest rate. Sustainable harvest rates are, in turn, strongly driven by the steepness of the stock-recruit relationship, with higher harvest rates possible for a stock with high steepness. Given harvesting at  $F_{msy}$  (for comparability), results based on equilibrium considerations indicate that size-based indicators are likely to be more sensitive to changes in biomass for stocks with relatively low growth rate ( $k$ ) and relatively high mortality ( $M$ ). Even at low mortality, indicators are more responsive when  $k$  is low. The most important factor is, however, steepness. Although knowledge about steepness is admittedly unlikely, this result suggests that, if there is evidence or other reasons (e.g. specifics of reproductive biology) to believe that a stock has low steepness, then a size-based indicator, particularly those based on length, may not be responsive enough to detect changes in biomass in time. This statement is based on the small expected change in the indicator under such a scenario, even without sampling error or recruitment variability to add “noise” to the “signal”.

More usefully though, results show that, even at high steepness, the expected change in size-based indicators for a stock where growth rate is high (0.6 in these examples), is relatively small. This again suggests that size-based indicators for such a stock may not be responsive enough to detect changes in biomass in time. Conversely, it implies that if a large change in a size-based indicator is observed for a stock with high growth rate, this should be taken very seriously because the harvest rate may already be too high. The results therefore do not imply that size-based information should not be collected, or that such indicators cannot be used at all, for stocks where the response is likely to be small. It does mean that even small or uncertain trends should be seriously considered in the light of other information.

It is also not surprising that sexual dimorphism in growth implies a stronger response in an indicator based on the data from the gender with lower growth rate. In our swordfish examples, this is always females (unless growth rates are very similar). Pooled data can lead to an indicator behaving very much like the female-based indicator or something in between the male- and female-based indicators. It is the detail of the form of the male and female growth curves (as well as other factors, including size at first capture relative to  $L_{inf}$ , and  $t_0$ ) that determines whether the gender-specific or pooled indicators are preferable.

When “noise” is added to the indicators via recruitment variability and sampling level, it is not surprising that it is more difficult to detect a change in an indicator as the noise-to-signal ratio increases. What is interesting, however, is that the gender-specific indicators respond differently to an increase in “noise”, particularly sampling level. The type of sexually dimorphic growth (divergent, intersecting or parallel) also affects the response.

In general, the mean length and mean weight indicators are more robust than the others and the weight-based indicators are more sensitive to changes in biomass than the length-based indicators. Given the practical difficulties with obtaining individual weight information, we are currently investigating the behaviour of indicators based on a proxy for weight, such as length cubed.

The brief section on decision rules illustrates the way in which a size-based indicator could be used in a decision rule. The differences in performance of the stock under the 7 scenarios underline the importance of designing decision rules that are robust to uncertainty in the

underlying dynamics. Although we have only illustrated this for uncertainty in growth, taking uncertainty in steepness and mortality into account will also be important.

As indicated in the introduction, the final phase of the study involves further exploration of decision rules, particularly in the context of a spatial model with different hypotheses about movement and re-distribution of stock as harvesting occurs, and in terms of using more than one indicator in the design of decision rules.

## References

- Basson, M. and Dowling, N. (2004). Standardisation of size-based indicators for bigeye and yellowfin tuna in the Indian Ocean. IOTC-2004-WPTT-15 presented at the 6<sup>th</sup> Session of the IOTC WPTT
- Dowling, N. and Basson, M. (2004). Standardisation of size-based indicators for broadbill swordfish in the Indian Ocean. IOTC-2004-WPB-07 presented at the 4<sup>th</sup> Session of the IOTC WPB.
- Punt, A.E., G. Cui and A.D.M. Smith. (2001). Defining robust harvest strategies, performance indicators and monitoring strategies or the SEF. Final Report to the Fisheries Research Development Corporation, Project No. 98/102
- Radtke, R.L. and Hurley, P.C.F. (1983). Age estimation and growth of broadbill swordfish, *Xiphias gladius*, from the northwest Atlantic based on external features of otoliths. U.S. Dep. Commer., NOAA Tech. Rep. NMFS 8:145-150
- Sun, C-L, Wang, S-P. and Yeh, S-Z. (2002) Age and growth of the swordfish (*Xiphias gladius* L.) in the waters around Taiwan determined from anal-fin rays. Fishery Bulletin 100(4): 822-835.
- Tserpes, G. and Tsimenides, N. (1995). Determination of age and growth of swordfish, *gladius* L., 1758, in the eastern Mediterranean using anal-fin spines. Fishery Bulletin 93: 594-602.
- Vanpouille, K., Poisson, R., Taquet, M., Ogor, A. and Troadec, H. (2001). Etude de la croissance de l'espadon. In 'L'espadon: de la recherché a l'exploitation durable'. (eds. Poisson, F. and Taquet, M.) pp 170-211. Programme Palangre Reunionnais, Rapport finale.
- Young, J.W. and Drake, A. (2002). Reproductive dynamics of broadbill swordfish (*Xiphias gladius*) in the domestic longline fishery off eastern Australia. Final report to the Fisheries Research and Development Corporation Project No. 1999/108.
- Young, J.W. and Drake, A. (2004). Age and growth of broadbill swordfish (*Xiphias gladius*) from Australian waters. Final report to the Fisheries Research and Development Corporation Project No. 2001/014.