CCAMLR Science, Vol. 18 (2011): 57–73

LENGTH-BASED ASSESSMENTS REVISITED – WHY STOCK STATUS AND FISHING MORTALITY OF LONG-LIVED SPECIES SUCH AS TOOTHFISH CANNOT BE INFERRED FROM LENGTH-FREQUENCY DATA ALONE

P.E. Ziegler⌧, D.C. Welsford and A.J. Constable
Australian Antarctic Division
Department of Sustainability, Environment,
Water, Population and Communities
203 Channel Highway
Kingston, Tasmania 7050
Australia
Email – philippe.ziegler@aad.gov.au

Abstract

In many of CCAMLR's new and exploratory fisheries, stock assessment techniques are still in development. The failure of the tag-recapture program to provide robust assessments in many areas and doubts about using unstandardised catch rate data have led to a desire to draw inferences about stock status and fishing mortality from length-frequency data alone. In this paper, a number of length-based stock assessment methods are revisited and some limitations in their ability to draw robust inferences about stock status and fishing mortality for long-lived fish species such as Patagonian toothfish (*Dissostichus eleginoides*) are illustrated. While length-frequency data can be useful in assessments when integrated with other informative estimators for stock abundance and productivity, the review and the simulations highlight that interpretation of raw length-frequency data and single length-based indicators for stock status and fishing mortality is problematic and should be avoided.

Keywords: fishery assessment, length-frequency data, Patagonian toothfish, CCAMLR

Introduction

CCAMLR aims to ensure that the development of a new and exploratory fishery does not occur faster than the Commission is able to achieve the objectives of Article II, notably maintaining stable recruitment of the fished populations and minimise the risk of ecosystem changes that are not reversible over two or three decades (Constable et al., 2000). At the core of this approach are rules for notification of intention to fish, measures requiring data collection, and decision rules for setting precautionary catch limits. The intent of the data collection is to provide sufficient data from research and fishery operations for an assessment of the status and productivity of the fished stocks to estimate the precautionary catch limits.

For new and exploratory fisheries of Patagonian toothfish (*Dissostichus eleginoides*) and Antarctic toothfish (*D. mawsoni*) in the Convention Area, a research plan, a data collection plan, and a tagging program specify the collection of four main data types: spatially referenced catch and effort data, catch length-frequency data, tag-recapture data, and otoliths or scales for age determination. Based on spatially concentrated tagging effort, tag-based assessments using ageing data have been developed for Subarea 48.4 and the Ross Sea (Subareas 88.1 and 88.2) (SC-CAMLR, 2010, paragraph 5.7). However, robust advice on catch limits according to CCAMLR decision rules has not been possible for most other new and exploratory toothfish fisheries.

Some reasons for this failure are the poor quality and quantity of data, and the strong reliance on tag-recapture data for abundance estimates. Tag-recapture data have been considered poor in many instances due to general problems with low tagging rates, insufficient overlap in the geographical location and the length-frequency of released tags and subsequently landed catch, poor tag detection rates, high post-tagging mortality of tagged fish, and the effects of IUU fishing on tag-based abundance estimation (SC-CAMLR, 2010, paragraph 5.5). Ageing of otoliths and a concerted effort to use the data for a stock assessment have so far occurred only in Subarea 48.4 and the Ross Sea in conjunction with tagging.

Catch rates have been used as indices of relative abundance in comparisons between areas and in depletion studies (McKinlay et al., 2008; SC-CAMLR, 2008, paragraph 5.82). But catch rates are often not directly comparable within or between areas due to inconsistencies in fishing operations with varying vessels, gear types and fishing locations (SC-CAMLR, 2010, paragraph 5.6). Longer time series of catch rates are required for a standardisation of catch rates and estimation of relative abundance. In addition, the effects of spatial variability in stock distribution and the non-random behaviour of fishing vessels are often underestimated. Trends in catch rates need to be analysed at a small spatial scale since fishers in exploratory fisheries rapidly gain experience and adapt their fishing tactics resulting in increased fishing efficiency (McKinlay et al., 2008). While catch rates can reflect local depletion, they can also be highly variable in areas of low abundance. Over larger areas, catch rates are notoriously hyper-stable, and as a consequence the interpretation of increasing or stable catch rates can be highly ambiguous.

With the imperative to provide an assessment of fished stocks in the CAMLR Convention Area, the failure of the tag-recapture program in many areas and doubts about using catch rate data have led to a desire to draw inferences about stock status and fishing mortality from length-frequency data. As a response to calls by the Scientific Committee for management advice for toothfish fisheries in Subarea 58.4 (e.g. SC-CAMLR, 2006, paragraph 4.150), the stability of length frequencies and unstandardised catch rates from some vessels has been used to argue that stock status is stable or increasing, if not above CCAMLR target levels, in areas where rapid localised declines in catch rates and large IUU catches suggested that stocks were likely to be depleted (SC-CAMLR, 2007, paragraphs 5.77 to 5.79; SC-CAMLR, 2008, paragraph 5.117).

However, several previous studies have indicated major problems with assessments that are based on length-frequency data alone. It is important that the lessons from these studies are not forgotten, since they are very likely to apply to toothfish fisheries as well. In this paper, the history of a number of length-based methods is revisited, and then some limitations in their ability to draw robust inferences about stock status and fishing mortality for Patagonian toothfish as an example of a long-lived fish species are illustrated.

Stock assessment methods based on length-frequency data

It makes some intuitive sense that the length-frequency composition of a fish population is affected by fishing: fishing is size-selective and truncates the length-frequency composition by generally removing larger fish. And because length-frequency data can be relatively easily collected, or may even be the only data type available, a range of methods have been developed to use length-frequency data for estimation of life-history and stock assessment parameters (see e.g. Pauly and Morgan, 1987 and Shepherd et al., 1987 therein; Gulland and Rosenberg, 1992; Gallucci et al., 1996). However, length-frequency data are generally more variable and their quality is considered to be inferior to that of age-frequency data, since a wider range of lengths corresponds to one age than vice versa. This variability severely reduces the usefulness of length-frequency data, and age-based assessment methods are generally preferred where individuals can be aged.

If age–length keys or growth functions are known, length-frequency compositions can be back-calculated into the corresponding age-frequency compositions. Without such data, a number of models can be used that follow modal progression in length-frequency distributions to identify cohorts and estimate various biological and fishery parameters (e.g. ELEFAN – Pauly, 1987; FiSAT II – Gayanilo et al., 2005; MULTIFAN-CL – Fournier et al., 1998; see also Gallucci et al. (1996) for a discussion). Some of these models have been found useful for assessing species that are difficult to age, such as tropical fish with little seasonal structure in fish otoliths or scales. But dissecting length-frequency data into several components representing cohorts is limited by the overlap of older cohorts, and Sparre and Venema (1998) argued that it is not feasible in practice to separate more than three or four cohorts within a dataset. Therefore, such methods are likely to be successful only for short-lived species with few and clearly distinguishable length modes.

For a direct estimation of fishing mortality F or total mortality Z, simpler length-based methods investigate the shape of the length-frequency composition, e.g. through mean length or an upper length percentile and length-corrected catch curves. These methods are based on the expectation that large fish will be abundant under low fishing mortality and rarer under high fishing mortality. Similar to their age-related counterparts, they share the assumptions of steady-state recruitment and mortality, exponential decline of the population with natural mortality, and knife-edge selectivity.

But departure from the steady-state assumption in regards to constant recruitment and mortality can introduce significant biases. A strong year-class can result initially in decreasing mean length, and as the cohort grows, mean length steadily increases until it drops again when the next strong year-class enters the fishery. Conversely, if recruitment fails, e.g. because of low spawning biomass, mean length may increase since small fish no longer enter the fishery and the remaining population continues to increase in size (and age). If there is no systematic trend in recruitment, smoothing of length-frequency samples can alleviate biases due to variable recruitment to some degree, i.e. population samples collected over several occasions can be pooled. Importantly though, fishing mortality will rarely be constant as this requires for the catch removal to be exactly balanced by new recruitment to the fishery. Applying constant mortality models to fish populations with changing fishing mortality can therefore introduce strong biases and effectively impede the reliable estimation of fishing or total mortality (see also simulations below).

Mean length, maximum length or an upper length percentile are easy to calculate, but their underlying relationship with stock abundance or a meaningful reference point must be known if they are used as an index for stock abundance. Punt et al. (2001) investigated mean length and the upper 5th percentiles of length (and weight) in individual fish as a stock indicator for swordfish (*Xiphias gladius*) population biomass. As expected, both measures decrease with fish biomass, but the relationships were not linear and the indicators were quite imprecise. Reference points based on these relationships depended on assumptions about stock structure and were often triggered too early or too late.

The effects of total mortality Z depend on fish growth parameters (e.g. the von Bertalanffy growth parameters L_∞ and K), and results are often expressed as the ratio Z/K. Analogous to the age-based catch curve, the length-based catch curve can be used to estimate the ratio Z/K. Since in contrast to age, growth is not constant, it takes a fish longer to grow through the larger length classes. Thus, the slope of the log-converted length-frequency distribution is curved and a function of time, mortality and the changing value of fish length. For the von Bertalanffy growth model, the slope of the length-frequency distribution, i.e. the relationship between numbers of individuals by size class versus size class, can be estimated as:

$$\log(N_l) = \alpha + \left(\tfrac{Z}{K}\right)\log(L_\infty - l) \qquad (1)$$

where N_l is the numbers at length l, and K and L_∞ are parameters of the von Bertalanffy growth function (Jones, 1984; Gallucci et al., 1996). This relationship is a straight line with positive slope Z/K and intercept α when $\log(N_l)$ is plotted versus $\log(L_\infty – l)$.

Other estimates for the ratio Z/K from length-frequency data are based on probability distributions of lengths, and regression and maximum likelihood methods (for an overview see Wetherall et al., 1987; Gallucci et al., 1996). Beverton and Holt (1957) provided the classical estimator by:

$$\frac{Z}{K} = \frac{L_\infty - L_{mean}}{L_{mean} - L_c} \qquad (2)$$

where only the mean length L_{mean} in a sample of fish above the length at first capture L_c and an estimate of L_∞ are required. In addition to steady-state conditions, the model assumes that the length-frequency distribution is not truncated, i.e. the maximum age in the catch t_λ is the same as the theoretical maximum age t_{max}. This method is sensitive to the accuracy with which L_∞ is known, and to length-class bin width (Laurec and Mesnil, 1987). In addition, Ehrhardt and Ault (1992) showed that this equation overestimates Z when the maximum age in the catch is less than that in the population. This bias decreased with higher fishing mortalities due to the truncation of the length-frequency distribution. Simulations by Wetherall et al. (1987) and Gallucci et al. (1996) highlighted different levels of precision and biases of the various methods.

Based on a method by Powell (1979), Wetherall et al. (1987) also suggested an improved method to estimate Z/K and L_∞ simultaneously. This method was examined with simulations by Somerton and Kobayashi (1991) who found that biases may occur even when the population was in equilibrium. Further, disequilibrium biases in Z/K and L_∞ estimates from variable fishing effort and recruitment could be large. While smoothing of data by averaging the length-frequency over some years helped to reduce the estimation bias caused by recruitment variation, it had little effect for the bias caused by changing fishing effort. The authors therefore cautioned against using this method where fishing effort and hence fishing mortality is likely to vary.

In an alternative approach, Froese (2004) suggested three simple indicators based on proportions of the catch length-frequency composition to assess status and trends in fish stocks and hypothesised about ideal target levels. Three indicators reflect the percentage of mature fish in the catch (P_{mat}) with 100% as a target level, the percentage of individuals that provide the highest yield (P_{opt}) with 100% as a target level, and the percentage of ‘mega-spawners’ (P_{mega}) with 0% as a target level or 30–40% where no upper size limit exists. The proportional length-based indicators can be estimated as:

$$P_{mat} = \sum_{L_{mat}}^{L_{\max}} P_l \quad (3)$$

$$P_{opt} = \sum_{0.9L_{opt}}^{1.1L_{opt}} P_l \quad (4)$$

$$P_{mega} = \sum_{1.1L_{opt}}^{L_{\max}} P_l \quad (5)$$

where P_l is the proportion of fish in a particular length class l, L_{mat} is the length at 50% maturity, and L_{opt} is the length-at-age at which the product of weight-at-age and numbers-at-age under zero fishing mortality is maximised.

Cope and Punt (2009) investigated reference points based on these indicators and their sum P_{sum}:

$$P_{sum} = P_{mat} + P_{opt} + P_{mega} \quad (6)$$

across a range of different stock productivity and selectivity patterns. The three indicators could adopt a wide range of values and were often not very sensitive to stock status except when stock productivity was very high, i.e. when the steepness h in the Beverton-Holt stock-recruitment relationship approached 1.0. For example, major decreases in stock biomass occurred even when P_{opt} equalled 1.0, thought to be the ideal situation, and fish stocks with P_{mat} and P_{opt} values much less than 1.0 could theoretically be fished at a target spawning biomass of 40% relative to an unfished population. The values of the indicators also depended strongly on the pattern of fishing selectivity and could not be interpreted correctly without the knowledge of the underlying selectivity pattern.

In the following, the limitations of some of these methods are illustrated using a simulated Patagonian toothfish population under changing fishing mortality, different selectivity functions and growth parameters, and alternative assumptions about recruitment patterns. The simulations clearly show that the variable behaviour of these measures restricts their meaningful application for estimating stock status and fishing mortality.

Methods

The Generalised Yield Model (GYM) (Constable and de la Mare, 1996; Constable et al., 2003) was used to simulate a single age-structured population of Patagonian toothfish under fishing exploitation, and to calculate length-based indicators as they would be observed in catch data (for general settings see Table 1). The life-history parameters for the Patagonian toothfish population were taken from the assessment in Division 58.5.2 around Heard Island and the McDonald Islands (Table 2) (SC-CAMLR, 2009; Candy et al., 2007).

The model assumed a single population without movement or migration. It represented age classes from 0 to 50 years, with the last age class being a ‘plus class’ in which older fish accumulated. The removal of catch was modelled through three knife-edge fishing selectivity functions representing combined gear efficiency and fish availability (Figure 1). The first function was dome-shaped and selected immature fish between the age of 4 years and the age of 50% maturity (‘immature selectivity’). This fishing selectivity targeted young fish that would eventually grow out of the selected length

Table 1: Input parameters for population projections of Patagonian toothfish using the Generalised Yield Model (see Constable et al., 2003, for model specification).

Category	Parameter	Values
Age structure	Recruitment age	1 y
	Oldest age in initial structure (plus group)	50 y
Recruitment	Mean recruitment (lognormal)	1 000 000
	Standard deviations σ_R for constant and variable recruitment	0.0, 0.6
	Level of SSB_0 below which recruitment declines linearly	0.2
Fishery parameters	Season	1 Dec–30 Nov
	Catch levels for the constant catch scenarios (tonnes)	0, 250, 500, 750, 1 000, 2 500, 5 000, 10 000
Simulation specifications	Number of runs in simulation	1 (constant rec), 100 (variable rec)
Individual trial specifications	Reference year	2000
	First year of projection	2001
	Years to project stock in simulation	60 y
	Reference start date in year	01/12
	Increments within year	12
	Reasonable upper bound for annual F	5.0

Table 2: Parameters for Patagonian toothfish (SC-CAMLR, 2009; Candy et al., 2007) population dynamics in the scenario 1 (base case), and alternative values for K and M that have been assumed in scenarios 2–5 (all other parameters remained the same as in the base case).

Parameters of Patagonian toothfish population dynamics	Scenario				
	1 (Base case)	2	3	4	5
Von Bertalanffy growth parameters					
L_∞	2 871 mm				
K	0.02056	0.02056	0.02056	0.015	0.025
t_0	–4.2897 y				
Natural mortality M	0.13	0.10	0.16	0.13	0.13
Weight at length W (mm to kg)					
A	2.59E-09				
B	3.2064				
Maturity					
A_{mat50}	14.0 y				
Range 5–95% maturity	11.0–17.0 y				
L_{mat50}	900 mm				
Optimum age and length (estimated)					
A_{opt}	15.7 y	20.3 y	12.5 y	16.7 y	15.1 y
L_{opt}	966 mm	1141 mm	838 mm	775 mm	1017 mm

range, mimicking observed selectivity functions of trawl and longlines. The second function was a simplified logistic function and selected all mature fish above the age of 50% maturity ('mature selectivity'). The third selectivity function was similar to the mature function, but truncated at the upper end such that only the first 10 mature year classes were selected ('truncated mature selectivity').

For each selectivity function, a number of simulations were conducted to evaluate the influence of changing fishing mortality, exploitation levels, alternative life-history characteristics, and recruitment variability on the length-based indicators. For different simulations of exploitation history, constant catches between 0 and 10 000 tonnes were removed over the projection period of 60 years starting arbitrarily in 2000. These catch levels were chosen to provide contrasting scenarios over the projection period. When removing a constant catch with constant recruitment, the spawning stock reaches a new equilibrium level or is depleted after a number of years (Figure 1).

Five scenarios with alternative life-history characteristics for natural maturity M and the von Bertalanffy growth parameter K were investigated (Table 2). In scenario 1 (base case), current best estimates for $M = 0.13$ and $K = 0.0256$ were assumed. For scenarios 2 and 3, natural mortality was assumed to be lower ($M = 0.10$) and higher ($M = 0.16$) respectively, and for scenarios 4 and 5, K was assumed to be lower ($K = 0.015$) and higher ($K = 0.025$) respectively.

Initial sets of simulations were conducted with constant recruitment to investigate indicator trends without recruitment variability as a major source of uncertainty. In subsequent simulations, recruitment R was modelled as a lognormal distribution with standard deviation σ_R. The standard deviation was set to 0.6 in line with current assessments.

In addition to examining raw length-frequency data directly, a number of the length-based indicators to infer information on stock status or fishing mortality were derived, namely mean length, the upper 75th percentile, the slope of the length-based catch curve, and Froese's (2004) proportional values. Instead of maximum length or a higher percentile, the 75th percentile that tends to be less prone to noise in the length-frequency distribution was used (Trenkel et al., 2007).

To demonstrate the direct behaviour of the indicators, all simulations assumed perfect knowledge of selectivity and life-history parameters and no sampling (observation) error when calculating indicator values from the catch length frequencies. Mean length (L_{mean}), the upper 75th percentile (L_{75}) and the slope of the length-based catch curve were calculated from individuals of all selected age classes. The only exception was the slope estimate for mature selectivity, for which the data were limited to 20 age classes. The proportional values were calculated as per equations (3) to (6).

Results

Examples of length-frequency distributions

As expected, the exponential decline of the length-frequency distributions under constant recruitment often steepened with decreasing biomass levels. This is shown in Figure 2(a) for immature selectivity and a constant catch of 1 000 tonnes at three different years during the projections 15 years apart, where stock levels fall from 0.72 to 0.26 and 0.04 and larger fish become increasingly rare. When stock levels drop below around 0.20, the catch is made up mainly of fish that have newly recruited to the fishery in that year, while larger fish are rapidly removed by the fishery each year. However, some length-frequency distributions remain remarkably stable. For a constant catch of 750 tonnes, spawning biomass stock status relative to virgin levels falls from 0.78 to 0.44 and then 0.30, while the overall length-frequency distribution does not change noticeably.

Variable recruitment events naturally modify these general patterns in the length-frequency distribution resulting in substantial variation between individual model runs. Figure 2(b) shows two example runs with a constant catch of 750 tonnes. The first run follows the development of the length-frequency distributions corresponding to an expectation that larger fish become rare with fishing over time. However, the second run shows an opposing pattern with a lower proportion of larger fish for 2012 than the bottom graph for 2042, whereas the true stock levels are at 0.92 in 2012 compared to only 0.28 in 2042.

Performance of length-based indicators under constant recruitment

The importance of the selectivity function on determining the absolute indicator values for mean (L_{mean}) and upper 75th percentile (L_{75}) of fish populations is shown in Figure 3. Under constant recruitment, L_{mean} and L_{75} in the catch are larger for mature fishing selectivity than for immature selectivity. Similarly, truncating a selectivity function decreases indicator values for L_{mean} and L_{75}.

The simulations generally support the hypothesis that L_{mean} and L_{75} decrease when stock size decreases (Figure 3). However, they also clearly highlight that the relationship between stock status and length-based indicators can be quite variable. For the mature selectivity function, there is a reasonably close and linear relationship between stock status and length-based indicators for a given set of life-history characteristics. However, for the dome-shaped immature and truncated mature selectivities most realised values of L_{mean} or L_{75} can represent a large range of stock status levels. For example, in the base-case scenario with the immature selectivity (black diamonds) an L_{mean} value of 547 mm can represent a spawning stock status relative to virgin levels between 0.0 and 0.92, while the smallest L_{mean} value of 449 mm can represent stock status values from 0.0 to 0.88. The deterministic growth function would have exacerbated this result to some degree, but growth variability would not change the basic result that dramatic changes in stock status can occur with relatively small changes in indicator values.

Both indicators can also vary for a given stock status depending on the catch history. This is particularly the case for immature selectivity where high catches immediately lower L_{mean} and L_{75} but stock status declines much slower due to the biomass of larger fish which is unfished and is only under the influence of natural mortality.

While the slope estimate of the Z/K ratio is little influenced by the selectivity function, the slope generally underestimates the true Z/K ratio (Figure 3). The slope is highly biased, particularly at low mortality levels, but generally improves over time for most scenarios when steady-state conditions are reached.

The significant effects of alternative life-history characteristics, such as natural mortality M and the growth parameter K, for the relationship between the length indicators and stock status or estimates of Z/K can be seen in Figure 4. For example, the maximum realised value for L_{75} with the mature selectivity varies from 1 055 to 1 129 mm between higher and lower M, and from 846 to 1 229 mm between lower and higher K. The negative bias of the Z/K ratio estimated by the slope is apparent in all tested scenarios.

The values of proportion indicators P_{mat}, P_{opt}, P_{mega} and P_{sum} depend strongly on the shape of the selectivity function, which is highlighted by the contrast in indicator values between immature and mature selectivity. Selecting immature fish results in all indicators being at, or close to, zero (not shown), while indicator values for mature selectivity are shown in Figure 5.

For mature selectivity, the relationship between P_{mat} and stock status is reasonably independent of the evaluated alternative life-history characteristics. Because the applied selectivity function is knife-edged, while the maturity function is not, P_{mat} includes some immature fish and is thus always smaller than 1.0. P_{mat} even approaches zero under severe stock depletion. For spawning stock status between 0.3 and 1.0 of virgin levels, P_{mat} is largely insensitive to changes. Both P_{opt} and P_{mega} are reasonably correlated with stock status for any given life-history scenario, however, values vary substantially across different scenarios. Worryingly, P_{opt} is even dome-shaped in some scenarios, rather than describing a monotonic relationship with stock status. Because the sum of stock proportions in P_{opt} and P_{mega} is very similar to P_{mat} for the mature selectivity, P_{sum} follows an overall similar pattern to P_{mat}. For truncated mature selectivity (not shown), P_{sum} has a less consistent pattern between scenarios with different life-history characteristics.

Performance of length-based indicators under variable recruitment

Besides the variable fishing mortality and alternative life-history characteristics, variable recruitment is probably the most important source for uncertainty in the relationship between length-based indicators and spawning stock status or estimates of Z/K. Under variable recruitment with $\sigma_R = 0.6$, realised values of all indicators can represent a large range of stock status levels or Z/K ratios even when only considering the base-case scenario (Figures 6

Table 3: Percentages of years that indicator values and spawning stock status increase or decrease for immature, mature and truncated mature selectivity (base-case scenario only).

		Immature		Truncated mature		Mature	
Stock status or true Z/K:		Decreasing (%)	Increasing or stable[1] (%)	Decreasing (%)	Increasing or stable[1] (%)	Decreasing (%)	Increasing or stable[1] (%)
Indicator							
L_{mean}:	Decreasing	9	7	7	10	8	11
	Increasing or stable	53	31	38	45	29	52
L_{75}:	Decreasing	10	7	8	10	9	12
	Increasing or stable	52	31	37	45	28	51
Slope:	Decreasing	1	47	1	45	0	47
	Increasing or stable	1	51	4	51	1	52
P_{sum}:	Decreasing	-	-	25	27	20	32
	Increasing or stable	-	-	20	28	16	32

[1] Within 1% of virgin stock biomass or 0.01 of fishing mortality that contributes to Z/K from previous year.

and 7). The slope estimate of Z/K tends to follow the true Z/K, but the spread of values for a given Z/K ratio is enormous. For mature selectivity, the slope estimate again almost always overestimates the true Z/K.

Changes in stock status and length-based indicators

Changes in values of most indicators are also a poor predictor for changes in spawning stock status (Table 3). For both selectivity functions, changes in L_{mean} and L_{75} from one year to the next correctly predict decreases or increases in stock levels for the same period in only about 40–60% of all years (base-case scenario). The proportional indicator P_{sum} performs correctly also only in around 53% of all years, while the slope of the size spectrum scores with around 52% to correctly predict changes in the true Z/K for all selectivity functions.

Discussion

Indicators based on length-frequency data have been typically used when stock dynamics and status were poorly known and alternative stock assessment methods could not be conducted. However, due to the inherent uncertainty in this type of data, it is exactly these stock dynamics and fishing parameters that must be well known for long-lived species such as toothfish to enable the correct interpretation of length-frequency data and related length-based indicators.

Raw length-frequency data and length-based indicators such as L_{mean}, L_{75}, the slope estimate of Z/K and Froese's (2004) proportional values contain little inherent conclusive information about stock status or the true Z/K ratio and cannot be generalised. Rather, as shown previously by a number of studies, length-frequency data of fish stocks are strongly influenced by parameters such as the variation in fishing mortality and recruitment, the form of the fishing selectivity function and life-history characteristics (see e.g. Ehrhardt and Ault, 1992; Punt et al., 2001; Somerton and Kobayashi, 1991; Cope and Punt, 2009).

The simulation results presented here should therefore come as no surprise, but simply highlight the limitations of using length-frequency data for the stock assessment of long-lived fish such as Patagonian toothfish. The violation of steady-state conditions through variable fishing mortality and recruitment is likely to be prevalent in many fisheries. In the case of Patagonian toothfish, recruitment variability may be substantial (Candy and Constable, 2008), and fishing mortality is likely to vary even with constant catches. As a result, catch length-frequency compositions, as well as changes within these compositions, can vary strongly independent of stock size and give a misleading indication of stock levels. Consequently, all length-based indicators tested here are unreliable estimators of stock status or the *Z/K* ratio on their own, even if the selectivity function and life-history characteristics were known exactly.

The form of the selectivity functions determines directly the range of fish lengths caught by the fishing gear and the interpretation of length-based data and indicators in respect to stock status. Yet, selectivity is notoriously difficult to estimate even with extensive datasets and complex stock assessment models. Fishing selectivity is a composite of the efficiency of the specific gear type to catch and retain fish of a given length, and the availability of such fish on the fishing grounds. Particularly the latter is rarely known. For example, a catch length-frequency with missing larger fish can be obtained with two different selectivity functions: one that catches large fish but these fish have been previously removed by fishing (equivalent to the mature selectivity function in the simulations), and one that does not catch large fish because emigration of these fish away from the fishing grounds simply makes them unavailable to the fishing gear (equivalent to the truncated mature selectivity function). Thus, the same values for L_{mean} or L_{75} could indicate low stock levels when assuming a mature selectivity where larger fish have been removed, or high stock status levels when assuming a truncated mature selectivity where larger fish move away from the fishing grounds.

Life-history parameters, such as natural mortality *M* and the growth coefficient *K*, that determine the productivity of fish stocks and have an effect on the catch length-frequency distributions, tend to be difficult to estimate exactly. Natural mortality *M* may even fluctuate between fish sizes and fishing years (Vetter, 1988). The *M* values examined in this study resulted in substantially different estimates for all length-based indicators, but all values were well within the confidence intervals of *M* estimated in a simulation study by Candy et al. (2011) using data from Heard Island and the McDonald Islands.

Joining length-based indicators and indicators based on other data types in a decision tree or multi-criterion indicators could provide alternatives to the simple approaches discussed here. A decision tree based mainly on P_{sum} for defining indicator values of stock status above certain spawning biomass reference points was proposed by Cope and Punt (2009). The tree requires further evaluations to examine the generality and reliability in practice, since it was unsuccessful in providing guidance about stock status in a quick examination of the toothfish example. In addition, the simulations used by Cope and Punt (2009) were based on a deterministic population dynamics model and it would be important to know whether a decision tree under stochastic recruitment variability was able to satisfy CCAMLR's Article II in the long term. Maybe more promising are decision trees that combine indicators from different data sources. In order to distinguish between the effects of fishing and recruitment, Davies et al. (2008) interpreted overall catch rates in the light of catch rates of small recruits and large (old) fish and changes in the proportion of large (old) fish. Likewise, Rochet et al. (2005) assessed ongoing changes in exploited fish populations and communities based on combined indices of abundance from trawl surveys with trends in average length of individuals (and other indicators for communities).

Joining indicators from different data sources may provide a viable approach also for the assessment of CCAMLR's new and exploratory fisheries. Indices of abundance derived from catch rate data or trawl surveys, similar to the research trawl survey by Australia around Heard Island and the McDonald Islands in Division 58.5.2 (Welsford et al., 2006), could be combined with age-frequency data, although issues with regards to ageing errors of otoliths and the representativeness of samples remain to be solved. Any such scheme will also need to be evaluated by simulations for appropriateness and robustness against a wide range of assumptions about fishery and fish stock characteristics, and embedded into the existing CCAMLR management framework with reference points and decision rules. Such an approach is much more likely to yield a successful assessment than reverting to methods that have proven to be, at best, difficult to interpret, or, at worst, misleading, at times when corrective management action is actually urgently required.

Conclusions

While length-frequency data can be useful in assessments when integrated with other informative estimators for stock abundance and productivity, any interpretation of raw length-frequency data and single length-based indicators for stock status and fishing mortality is problematic and should be avoided.

Acknowledgements

We are grateful to Steve Candy and John McKinlay for their constructive discussions and input to the simulations.

References

Beverton, R.J.H. and S.J. Holt. 1957. On the dynamics of exploited fish populations, UK. *Fish. Invest., London*, Ser. II, 19: 533 pp.

Candy, S.G. and A.J. Constable. 2008. An integrated stock assessment for the Patagonian toothfish (*Dissostichus eleginoides*) for the Heard and McDonald Islands using CASAL. *CCAMLR Science*, 15: 1–34.

Candy, S.G., A.J. Constable, T. Lamb and R. Williams. 2007. A von Bertalanffy growth model for toothfish at Heard Island fitted to length-at-age data and compared to observed growth from mark-recapture studies. *CCAMLR Science*, 14: 43–66.

Candy, S.G., D. Welsford, T. Lamb, J. Verdouw, J. Hutchins. 2011. Estimation of natural mortality for the Patagonian toothfish at Heard and McDonald Islands using catch-at-age and aged mark-recapture data from the main trawl ground. *CCAMLR Science*, 18: 29–45.

Constable, A.J. and W.K. de la Mare. 1996. A generalised model for evaluating yield and the long-term status of fish stocks under conditions of uncertainty. *CCAMLR Science*, 3: 31–54.

Constable, A.J., W.K. de la Mare, D.J. Agnew, I. Everson and D. Miller. 2000. Managing fisheries to conserve the Antarctic marine ecosystem: practical implementation of the Convention on the Conservation of Antarctic Marine Living Resources (CCAMLR). *ICES J. Mar. Sci.*, 57 (3): 778–791.

Constable, A.J., A.T. Williamson and W.K. de la Mare. 2003. *Generalised Yield Model: User's Manual and Specifications. Version 5.01b.* Australian Antarctic Division, Kingston, Australia: 165 pp.

Cope, J.M. and A.E. Punt. 2009. Length-based reference points for data-limited situations: applications and restrictions. *Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science*, 1: 169–186.

Davies, C., R. Campbell, J. Prince, N. Dowling, D. Kolody, M. Basson, K. McLoughlin, P. Ward, I. Freeman and A. Bodsworth. 2008. Development and preliminary testing of the harvest strategy framework for the Western Tuna and Billfish Fishery. Report to the Australian Fisheries Management Authority, Canberra, Report No. R2006/0684: 60 pp.

Ehrhardt, N.M. and J.S. Ault. 1992. Analysis of two length-based mortality models applied to bounded catch length frequencies. *Trans. Am. Fish. Soc.*, 121: 115–122.

Froese, R. 2004. Keep it simple: three indicators to deal with overfishing. *Fish and Fisheries*, 5 (1): 86–91.

Fournier, D.A., J. Hampton and J.R. Sibert. 1998. MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. *Can. J. Fish. Aquat. Sci*., 55 (9): 2105–2116.

Gallucci, V.F., S.B. Saila, D.J. Gustafson and B.J. Rothschild. 1996. *Stock Assessment: Quantitative Methods and Applications for Small-Scale Fisheries.* CRC Press, Lewis Publishers, Boca Raton, Florida: 527 pp.

Gayanilo, F.C. Jr., P. Sparre and D. Pauly. 2005. FAO-ICLARM Stock Assessment Tools II (FiSAT II). Revised Version. User's Guide. *FAO Computerized Information Series (Fisheries),* 8. FAO, Rome: 168 pp.

Gulland, J.A. and A.A. Rosenberg. 1992. A review of length-based approaches to assessing fish stocks. *FAO Fish. Tech. Pap.*, 323. FAO, Rome: 100 pp.

Jones, R. 1984. Assessing the effects of changes in exploitation pattern using length composition data (with notes on VPA and cohort analysis). *FAO Fish. Tech. Pap.*, 256. FAO, Rome: 117 pp.

Laurec, A. and B. Mesnil. 1987. Analytical investigations of errors in mortality rates estimated from length distributions of catches. In: Pauly, D. and G.R. Morgan (Eds.). *Length-based Methods in Fisheries Research*. ICLARM Conference Proceedings, 13. ICLARM, Manila, Philippines, and Kuwait Institute for Scientific Research Safat, Kuwait: 239–281.

McKinlay, J.P., D.C. Welsford, A.J. Constable and G.B. Nowara. 2008. An assessment of the exploratory fishery for *Dissostichus* spp. on BANZARE Bank (CCAMLR Division 58.4.3b) based on fine-scale catch and effort data. *CCAMLR Science*, 15: 55–78.

Pauly, D. 1987. A review of the ELEFAN system for analysis of length-frequency data in fish and aquatic invertebrates. In: Pauly, D. and G.R. Morgan (Eds.). *Length-based Methods in Fisheries Research*. ICLARM Conference Proceedings, 13. ICLARM, Manila, Philippines, and Kuwait Institute for Scientific Research Safat, Kuwait: 7–34.

Pauly, D. and G.R. Morgan. 1987. *Length-based Methods in Fisheries Research*. ICLARM Conference Proceedings, 13. ICLARM, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait: 468 pp.

Powell, D.G. 1979. Estimation of mortality and growth parameters from the length-frequency in the catch. *Rapp. P-V. Réun. Cons. Int. Explor. Mer*, 175: 16–169.

Punt, A.E., R.A. Campbell and A.D.M. Smith. 2001. Evaluating empirical indicators and reference points for fisheries management: application to the broadbill swordfish fishery off eastern Australia. *Mar. Freshw. Res.*, 52 (6): 819–832.

Rochet, M.-J., V. Trenkel, R. Bellail, F. Coppin, O. Le Pape, J.-C. Mahé, J. Morin, J.-C. Poulard, I. Schlaich, A. Souplet, Y. Vérin and J. Bertrand. 2005. Combining indicator trends to assess ongoing changes in exploited fish communities: diagnostic of communities off the coasts of France. *ICES J. Mar. Sci.*, 62 (8): 1647–1664.

SC-CAMLR. 2006. *Report of the Twenty-fifth Meeting of the Scientific Committee (SC-CAMLR-XXV)*. CCAMLR, Hobart, Australia: 523 pp.

SC-CAMLR. 2007. Report of the Working Group on Fish Stock Assessment. In: *Report of the Twenty-sixth Meeting of the Scientific Committee (SC-CAMLR-XXVI)*, Annex 5. CCAMLR, Hobart, Australia: 251–406.

SC-CAMLR. 2008. Report of the Working Group on Fish Stock Assessment. In: *Report of the Twenty-seventh Meeting of the Scientific Committee (SC-CAMLR-XXVII)*, Part I, Annex 5. CCAMLR, Hobart, Australia: 305–457.

SC-CAMLR. 2009. Report of the Working Group on Fish Stock Assessment. Fishery report: *Dissostichus eleginoides* Heard Island (Division 58.5.2). In: *Report of the Twenty-eighth Meeting of the Scientific Committee (SC-CAMLR-XXVIII),* Annex 5, Appendix O. CCAMLR, Hobart, Australia.

SC-CAMLR. 2010. Report of the Working Group on Fish Stock Assessment. In. *Report of the Twenty-ninth Meeting of the Scientific Committee (SC-CAMLR-XXIX)*, Annex 8. CCAMLR, Hobart, Australia: 275–395.

Shepherd, J.G., G.R. Morgan, J.A. Gulland and C.P. Mathews. 1987. Methods of analysis and assessment: report of working group II. In: Pauly, D. and G.R. Morgan (Eds). *Length-based Methods in Fisheries Research*. ICLARM Conference Proceedings, 13. ICLARM, Manila, Philippines, and Kuwait Institute for Scientific Research Safat, Kuwait: 353–362.

Somerton, D.A. and D.R. Kobayashi. 1991. Robustness of the Wetherall length-based method to population disequilibria. *Fish. Bull.*, 89 (2): 307–314.

Sparre, P. and S.C. Venema. 1998. Introduction to tropical fish stock assessment. Part 1. Manual. *FAO Fish. Tech. Pap.*, 306.1, Rev. 2: 407 pp.

Trenkel, V.M., M.-J. Rochet and B. Mesnil. 2007. From model-based prescriptive advice to indicator-based interactive advice. *ICES J. Mar. Sci.*, 64 (4): 768–774.

Vetter, E.F. 1988. Estimation of natural mortality in fish stocks: a review. *Fish. Bull.*, 86 (1): 25–43.

Welsford, D.C., A.J. Constable and G.B. Nowara. 2006. Review of the use of survey data and length-at-age models in the assessment of *Dissostichus eleginoides* in the vicinity of Heard Island and McDonald Islands (Division 58.5.2). Document *WG-FSA-06/44*. CCAMLR, Hobart, Australia.

Wetherall, J.A., J.J. Polovina and S. Ralston. 1987. Estimating growth and mortality in steady-state fish stocks from length-frequency data. In: Pauly, D. and G.R. Morgan (Eds.). *Length-based Methods in Fisheries Research*. ICLARM Conference Proceedings, 13. ICLARM, Manila, Philippines, and Kuwait Institute for Scientific Research Safat, Kuwait: 53–74.

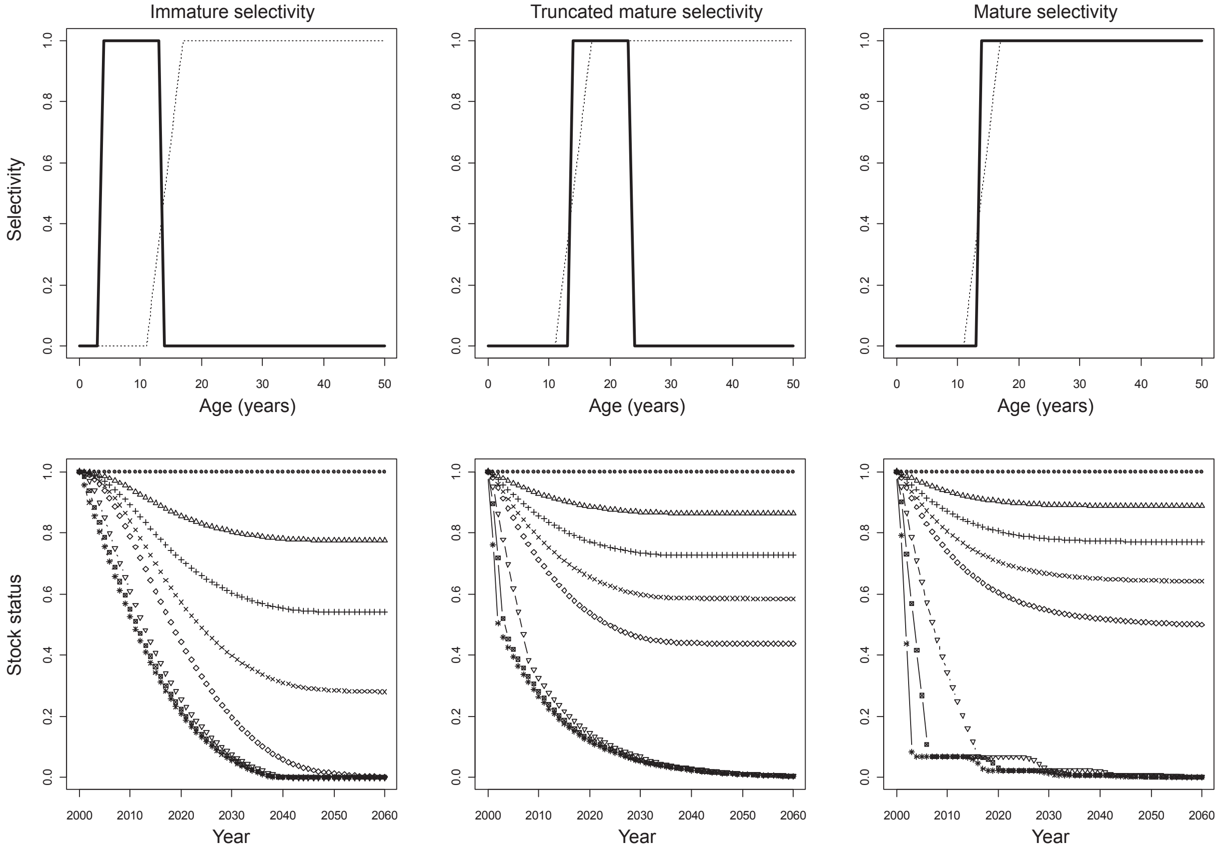

Figure 1: Immature, truncated mature and mature fishing selectivity functions (with dotted line for the maturity function), and spawning stock status relative to virgin levels of Patagonian toothfish over the projection period of 60 years for constant recruitment under different exploitation histories (0, 250, 500, 750, 1 000, 2 500, 5 000, and 10 000 tonnes distinguished by different markers) for the evaluated selectivity functions.

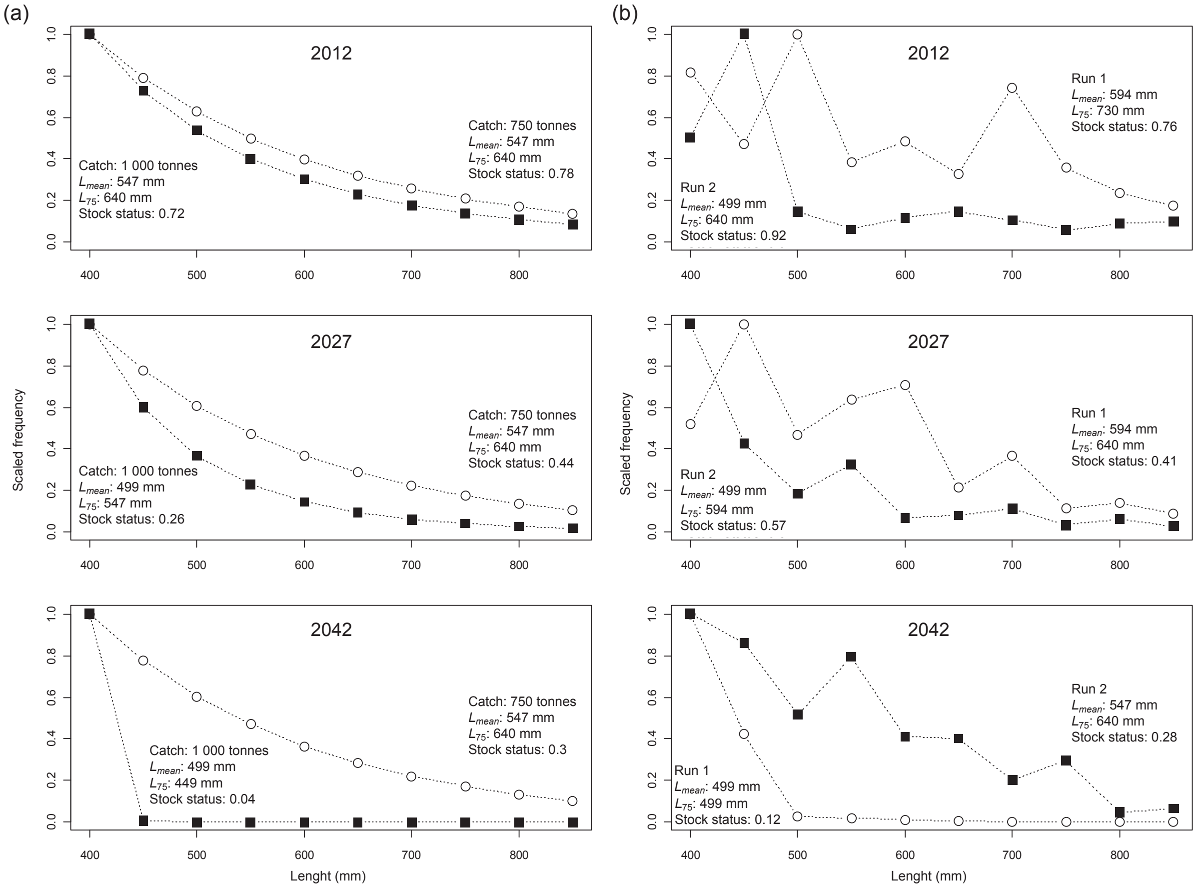

Figure 2: Examples for length-frequency patterns at three different years during the projections for immature selectivity and (a) constant recruitment with catches of 750 and 1 000 tonnes; and (b) for variable recruitment with catches of 750 tonnes.

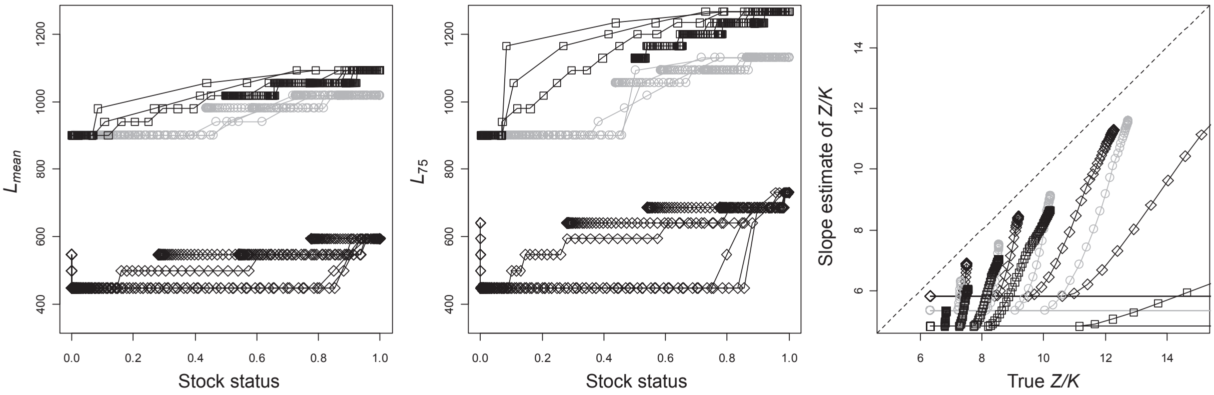

Figure 3: Length-based indicators mean length (L_{mean}) and length at 75 percentile (L_{75}) versus relative spawning biomass (stock status), and slope estimate of the Z/K ratio versus the true Z/K ratio (with dotted parity line) with constant recruitment and immature (black diamonds), truncated mature (grey circles) and mature selectivity (black squares) for the base-case scenario. Different catch histories are connected by lines. The distinct values of L_{mean} and L_{75} are due to the deterministic growth function and may slightly exaggerate differences between different catch histories.

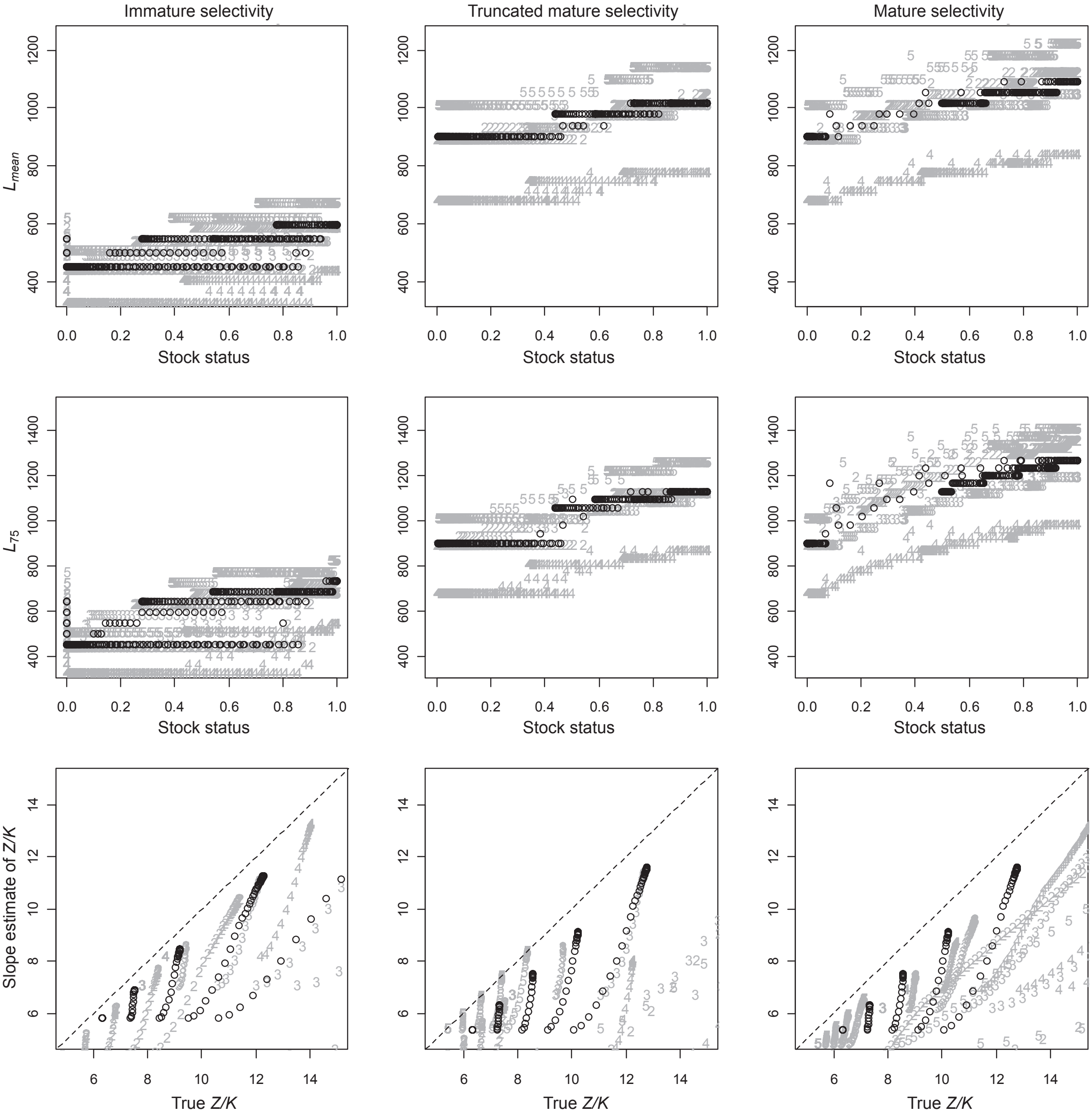

Figure 4: Length-based indicators mean length (L_{mean}) and length at 75 percentile (L_{75}) versus relative spawning biomass (stock status), and slope estimate of the *Z/K* ratio versus the true *Z/K* ratio (with dotted parity line) with constant recruitment and immature, truncated mature and mature selectivity, for the base-case scenario (black circles) and alternative life-history scenarios for *M* or *K* (grey numbers for scenarios, lower *M* (2), higher *M* (3), lower *K* (4), higher *K* (5)). The distinct values of L_{mean} and L_{75} are due to the deterministic growth function and may slightly exaggerate differences between different scenarios.

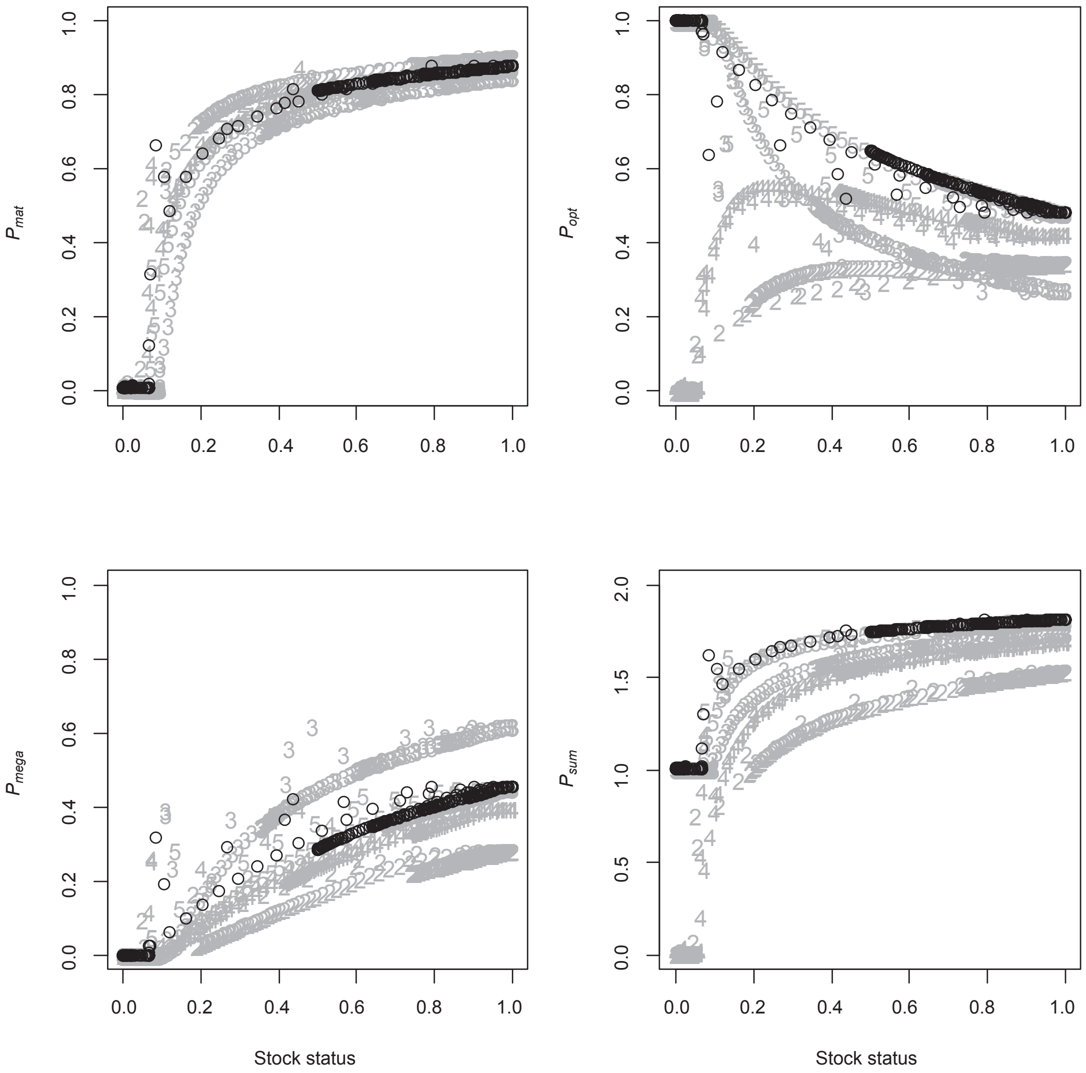

Figure 5: Proportion of mature (P_{mat}), optimally sized (P_{opt}), and 'mega-sized' fish (P_{mega}), and sum of these proportions (P_{sum}) versus relative spawning biomass (stock status) for constant recruitment and mature selectivity, for the base-case scenario (black circles) and alternative life-history scenarios for *M* or *K* (grey numbers for scenarios, lower *M* (2), higher *M* (3), lower *K* (4), higher *K* (5)).

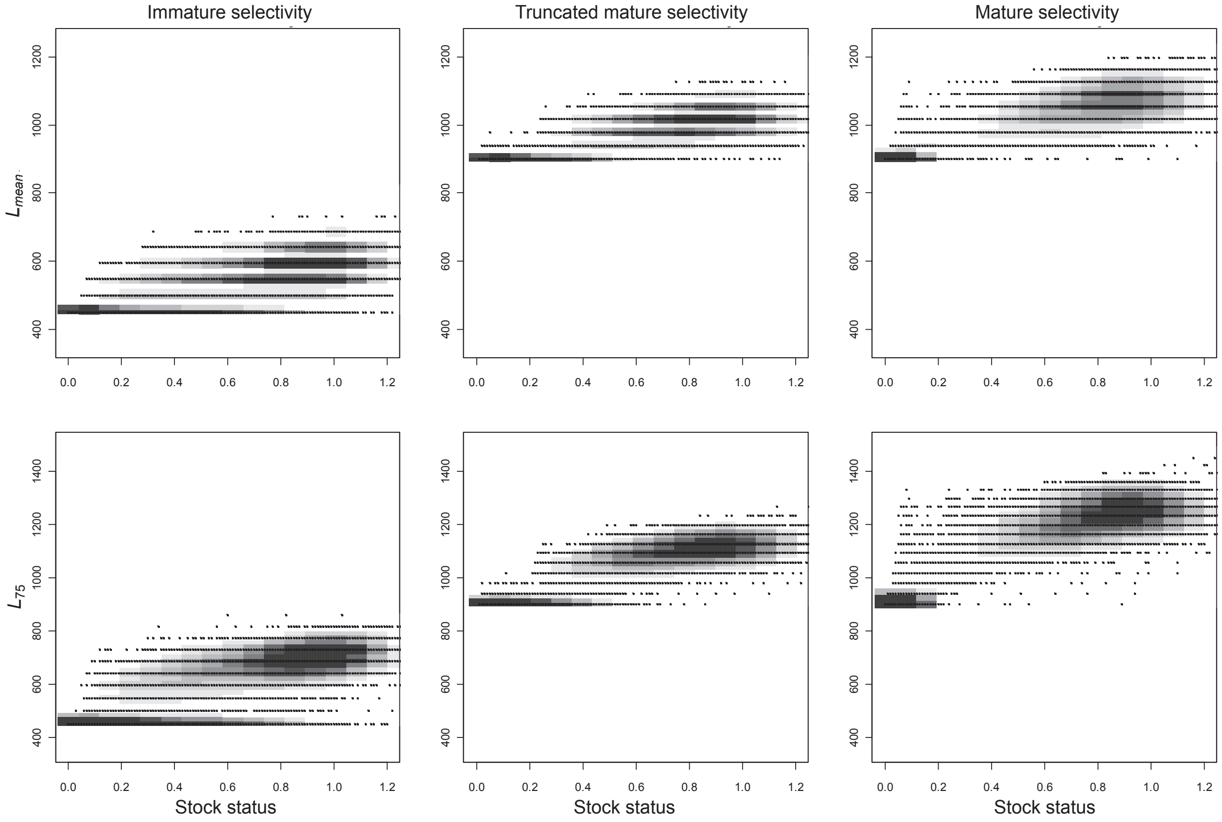

Figure 6: Length-based indicators mean length (L_{mean}) and length at 75 percentile (L_{75}) versus relative spawning biomass (stock status) for the base-case scenario and variable recruitment (100 simulations) with immature, truncated mature and mature selectivity. Individual points (with rounded values) indicate the spread of the data, the grey shading represents point densities as estimated by a bivariate normal kernel density function.

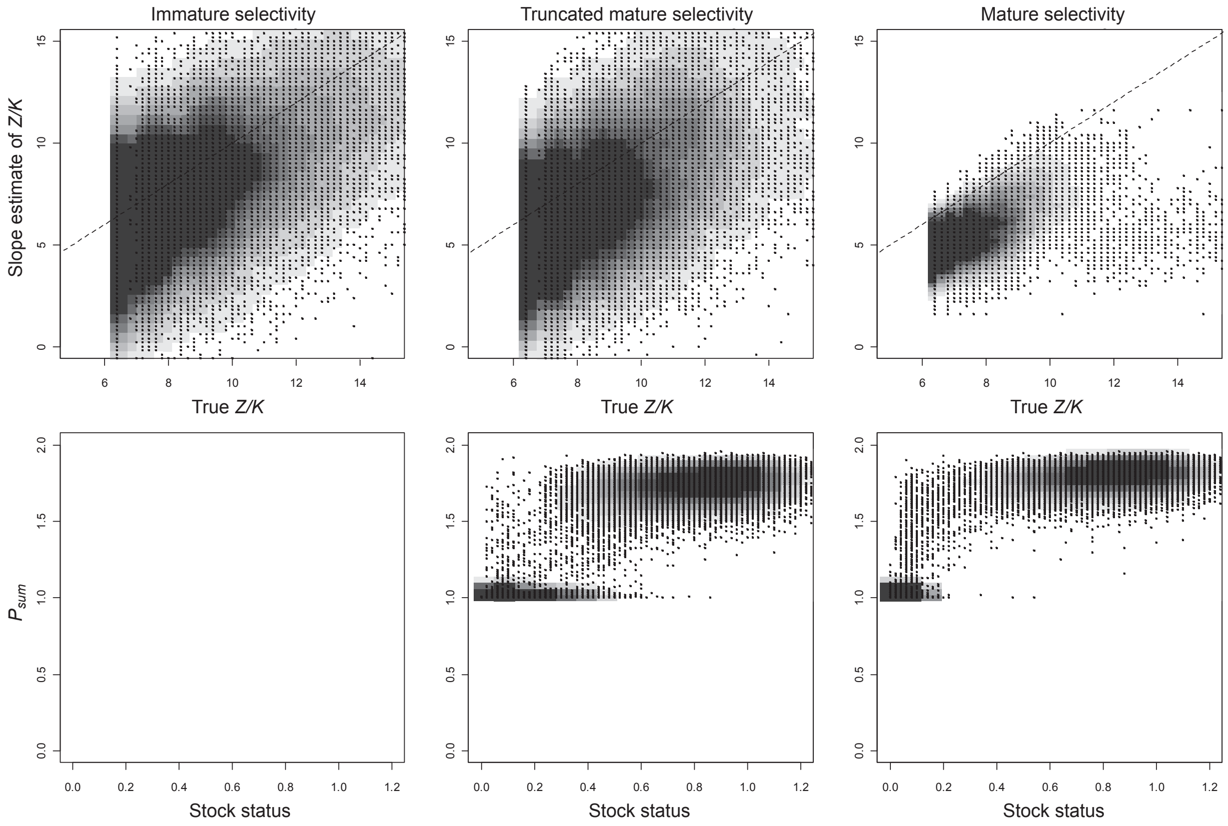

Figure 7: Slope estimate of *Z/K* versus true *Z/K* (with dotted parity line) and P_{sum} versus relative spawning biomass (stock status) for the base-case scenario and variable recruitment (100 simulations) with immature, mature and truncated mature selectivity. Note that when $F = 0$ the true *Z/K* ratio is equivalent to *M/K* (6.32 in the base case). Individual points (with rounded values) indicate the spread of the data, the grey shading represents point densities as estimated by a bivariate normal kernel density function.