ICES Journal of Marine Science

ICES International Council for the Exploration of the Sea
CIEM Conseil International pour l'Exploration de la Mer

ICES Journal of Marine Science (2015), 72(1), 7–18. doi:10.1093/icesjms/fsu015

Original Articles

Contemporary fisheries stock assessment: many issues still remain

Mark N. Maunder[1,2]* and Kevin R. Piner[3]

[1]Inter-American Tropical Tuna Commission, 8901 La Jolla Shores Drive, La Jolla, CA 92037-1508, USA
[2]Center for the Advancement of Population Assessment Methodology, Scripps Institution of Oceanography, La Jolla, USA
[3]NOAA Fisheries, Southwest Fisheries Science Center, 8901 La Jolla Shores Dr, La Jolla, CA 92037-1508, USA

*Corresponding author: tel: +1 858 546 7027; fax: +1 858 546 7133; e-mail: mmaunder@iattc.org

Maunder, M. N., and Piner, K. R. Contemporary fisheries stock assessment: many issues still remain. – ICES Journal of Marine Science, 72: 7–18.

Received 23 September 2013; accepted 17 January 2014; advance access publication 25 February 2014.

Interpretation of data used in fisheries assessment and management requires knowledge of population (e.g. growth, natural mortality, and recruitment), fisheries (e.g. selectivity), and sampling processes. Without this knowledge, assumptions need to be made, either implicitly or explicitly based on the methods used. Incorrect assumptions can have a substantial impact on stock assessment results and management advice. Unfortunately, there is a lack of understanding of these processes for most, if not all, stocks and even for processes that have traditionally been assumed to be well understood (e.g. growth and selectivity). We use information content of typical fisheries data that is informative about absolute abundance to illustrate some of the main issues in fisheries stock assessment. We concentrate on information about absolute abundance from indices of relative abundance combined with catch, and age and length-composition data and how the information depends on knowledge of population, fishing, and sampling processes. We also illustrate two recently developed diagnostic methods that can be used to evaluate the absolute abundance information content of the data. Finally, we discuss some of the reasons for the slowness of progress in fisheries stock assessment.

Keywords: abundance, catchability, diagnostics, fisheries stock assessment, growth, natural mortality, recruitment, selectivity.

Introduction

It is surprising that half a century after the publication of Beverton and Holt's (1957) manual on fisheries stock assessment, many of the same issues still remain unresolved. These unresolved issues include critical biological and fisheries processes such as growth, natural mortality, recruitment, and selectivity. Despite the lack of progress in some areas, there have been several major developments in our understanding of other areas of fishery modelling and analysis (Quinn, 2003). Age-structured models used for yield-per-recruit analysis and the notion of density-dependence in the stock–recruitment relationship caused assessment scientists to realize that the traditional Schaefer model (Schaefer, 1954) with maximum sustainable yield (MSY) occurring at 50% of the unexploited level was suboptimal. This understanding led Pella and Tomlinson (1969) to develop a generalized surplus production model for which MSY could occur at any level of unfished biomass. Age-structure models were also developed to reconstruct abundance from the time-series of age-composition data [e.g. cohort analysis (Pope 1972), virtual population analysis (VPA; Gulland, 1965), extended survivors analysis (Shepherd, 1999)]. A lack of a complete time-series of composition data for many species has led to the extensive use of integrated models that use all available data simultaneously in a statistical framework (Fournier and Archibald, 1982; Deriso *et al.*, 1985; Maunder and Punt, 2013) and the development of the software to implement them [e.g. AD Model Builder (ADMB); Fournier *et al.*, 2012]. Lack of age data for hard-to-age species has led to the development of integrated length-structured models (Punt and Kennedy, 1997; Punt *et al.*, 2013). Models have also been developed to include multiple species (e.g. Jurado-Molina *et al.*, 2005). Several general models are now available to implement integrated analysis [e.g. Stock Synthesis (Methot, 1990; Methot and Wetzel, 2013), Coleraine (Hilborn *et al.*, 2000), CASAL (Bull *et al.*, 2005), MULTIFAN-CL (Fournier *et al.*, 1998; Hampton and Fournier, 2001), Gadget (Begley and Howell, 2004)] and greatly facilitate the application of stock assessments.

Fisheries stock assessment has also been at the forefront of Bayesian applications (Punt and Hilborn, 1997). The early applications of Bayesian analysis for integrated stock assessment models were some of the most complicated applications in any field at the time (e.g. McAllister *et al.*, 1994; McAllister and Ianelli, 1997). Many of the early applications used the sampling/importance resampling algorithm (McAllister *et al.*, 1994). Bayesian analysis was greatly facilitated by the development of Markov chain Monte Carlo (MCMC; e.g. Punt and Kennedy, 1997) and inclusion of an MCMC algorithm in ADMB (Fournier *et al.*, 2012). One of the early ADMB-based general models, Coleraine (Hilborn *et al.*, 2000), was based on the Bayesian paradigm (e.g. priors were implemented for all estimated parameters). Bayesian analysis accommodates the use of prior information in integrated assessments, allowing sharing of information from other species or the inclusion of expert opinion with the associated uncertainty. This is particularly useful for low information species (Cope, 2013). It also allows for the representation of uncertainty in a probabilistic context, which is ideal for decision analysis. The use of Bayesian analysis promoted the development of methods to develop priors such as meta-analysis of data from many populations (e.g. Hilborn and Liermann, 1998; Myers *et al.*, 1999; Dorn, 2002). Finally, due to the realization that there is substantial uncertainty in most stock assessments, a whole field of research has evolved around developing and testing management procedures that are robust to the uncertainty (de la Mare, 1986; Butterworth *et al.*, 1997; De Oliveira *et al.*, 1998; Butterworth and Punt, 1999; Smith *et al.*, 1999). These management procedures are comprehensive and include the data collected, the analytical methods, and the resulting management actions.

Stock assessment models make use of multiple data types to inform modelled processes. The data typically available for use in the model include catch by fleet (e.g. gear), indices of relative abundance from either fisheries catch per unit effort (cpue) or fishery-independent surveys, and catch and/or survey composition (age, length, or weight). In addition to information on absolute abundance and abundance trends, these data must provide information on growth, recruitment including the stock–recruitment relationship and variation around the relationship, natural mortality, and selectivity, among other processes. Examples of directly informing model process include age–length data from otoliths that provide information on growth or composition data to inform selection. However, modelled processes receiving information indirectly from data sources are often poorly estimated, and therefore, the parameters representing these processes (e.g. the steepness of the stock–recruitment curve and natural mortality) are typically assumed and fixed in the model. Unfortunately, the lack of reliable direct data on these processes results in the need to make highly uncertain assumptions. Sensitivity analyses to these assumptions are often conducted to investigate the consequences of this uncertainty, but without probabilistic statements for different values of the parameters, which are often not available because data are not informative, sensitivity analyses are difficult to interpret in terms of management advice. In addition, there is a lack of diagnostic tools to determine when a parameter value is assumed incorrectly.

Reference points that are typically determined using stock assessment models and evaluated using the results of stock assessments are also sensitive to biological parameters (Maunder, 2012) and selectivity (Goodyear, 1996). For example, Maunder (2002, 2012) showed that the biomass corresponding to MSY as a ratio of the unexploited biomass, a standard reference point that is also often used in harvest control rules, depends on the steepness of the stock–recruitment relationship, natural mortality, and the growth rate. Maunder (2003) also showed that the same reference point is also sensitive to the selectivity of the gear. Uncertainty or bias in the biological or fishing processes can substantially impact the estimates of reference points and has led to the use of proxy reference points that are considered robust to uncertainty or are precautionary (Clark, 1991, 2002). The implications may be greater in management systems for which complex decision rules are used to set annual catch limits. The abundance estimates are used twice, first to determine what harvest rate is applied (e.g. the harvest rate is reduced from F_{MSY} if the biomass is less than B_{MSY}) then as a product with the harvest rate to determine the catch level.

The objective of this paper is to make evident the fact that a substantial amount of additional research remains that could improve stock assessments and ensure sustainable management of fish stocks. We use information content of typical fisheries data on absolute abundance to illustrate some of the main issues in fisheries stock assessment. We concentrate on information about absolute abundance from indices of relative abundance combined with catch and age- and length-composition data and how the information depends on knowledge of population, fishing, and sampling processes. Next, we summarize the uncertainty in our knowledge of these processes. We then illustrate two recently developed diagnostic methods that can be used to evaluate the absolute abundance information content of the data. Finally, we describe, from our perspective, some of the factors hindering the progress of improving fisheries stock assessment.

Absolute abundance information

Estimating absolute abundance is an essential component of most fisheries stock assessments because total allowable catch is often calculated as a harvest rate applied to the model estimate of abundance. In contemporary integrated stock assessments, a variety of data types can be included in the analysis and simultaneously provide information on all estimated parameters (Maunder and Punt, 2013). In these integrated models, the production processes (growth, natural mortality, and stock–recruit relationship) are often fixed (the parameters representing these processes are not estimated, but set at predetermined values), making estimation of absolute abundance the focus (intended or not) of the analysis. However, the complexity of integrated models makes it difficult to understand the role that data types have on the model estimates of scale. Data interact with structural assumptions producing estimates that are often unexpected and unintuitive. Here, we illustrate that the information on absolute abundance from each data type depends on population (recruitment, natural mortality, and growth), fishing (selectivity), and sampling processes.

Indices of relative abundance

Indices of relative abundance are a common data type used in integrated stock-assessment models. They can be derived from dedicated research surveys or developed from fishery cpue data. Typically, the catchability of the gear is not known even for statistically designed surveys, and the catchability coefficient (the proportionality constant that converts the relative index into absolute abundance) of the index of abundance is estimated as a parameter in integrated stock-assessment models. So, in practice, these data are generally not used as measures of absolute abundance (i.e. survey catchability is not set to 1).

The depletion in abundance, as measured by an index of relative abundance, caused by a known level of catch can be used to calculate the initial biomass, as illustrated in Figure 1 (initial biomass equals catch divided by depletion proportion). The actual depletion will be

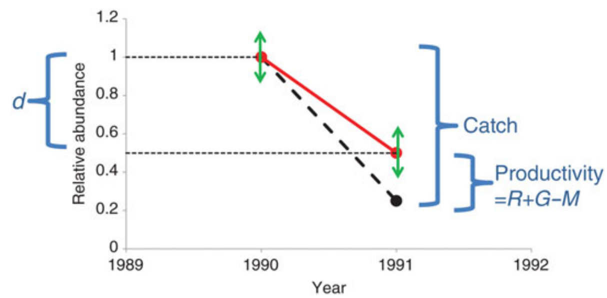

Figure 1. Illustration of the absolute abundance information contained in indices of relative abundance and its reliance on biological, fishing, and sampling processes. The solid and dashed lines represent depletion in the presence and the absence of biological processes, respectively. The arrows represent sampling error. *d*, depletion; *R*, recruitment; *G*, growth; *M*, natural mortality.

less because production (the difference between increase in abundance due to growth and recruitment and the decrease in abundance due to natural mortality) will generally increase the abundance between the two relative biomass estimates. There is also the complication of what component of the population is represented by the index of abundance, as measured by the gear selectivity, and what component of the population is represented by the catch, particularly if the catch and index come from different gears. Each relative index data point is a sample of the abundance, not a census, so it contains some random sampling error. In addition, other factors not accounted for in the index (e.g. temporal variability in the environment, expansion of the oxygen minimum layer; Stramma *et al.*, 2008) may cause catchability to vary over time introducing further error into the index. This clearly illustrates that information about absolute abundance from indices of relative abundance conditioned on catch depends on population dynamics (e.g. natural mortality, recruitment, and growth), fishing (e.g. selectivity), and sampling (e.g. random sampling variability and temporal variation in catchability), and these must be taken into consideration when interpreting the data.

Age- and length-composition data

Composition data, either from fisheries or from surveys, are used in most current assessments of major fish stocks. Length-composition data are the most common and can be converted to age composition using age–length keys either inside (e.g. by fitting to both age conditioned on length and length-composition data) or outside the assessment model. Age composition is the more direct measure and is therefore considered more informative when the ageing procedures are well established. VPA and cohort analysis, traditionally used for major fish stocks, rely on comprehensive age-composition data for all years, and therefore, data collecting programmes for these stocks have concentrated on collecting age-composition data or using methods (e.g. cohort slicing) to convert length-composition data into age composition. However, the collection of composition data is usually expensive, an expense that increases with the additional sample processing associated with production ageing, although the per-otolith costs may decrease with production ageing. Thus, there are limited age-composition data for many stocks, sometimes relying entirely on length-composition data. It is this lack of consistent age-composition data that has led to the extensive use of integrated analysis (Maunder and Punt, 2013; Punt *et al.*, 2013). Integrated analysis has the added advantage of including other types of data, such as weight composition, sex composition, and stage categories.

Age-composition data have generally been considered to provide information on cohort strength (recruitment) and selectivity in integrated stock assessment models. However, as part of an integrated analysis, age- and length-composition data are also linked to the population governing processes, and through these provide information on fishing mortality. The higher the fishing mortality the less likely a fish will survive to an old age. In combination with known catch, fishing mortality provides information on absolute abundance by definition, since fishing mortality approximately equals catch divided by biomass ($F \approx C/B$). This relationship between composition data and fishing mortality has roots in the tradition of using a linear regression applied to the logarithm of the relative abundance of a cohort at age to estimate total mortality (catch curve analysis). In the simplest form, the relative abundance of a cohort over time is usually not available so the analysis is applied to catch-at-age, which is the relative abundance of multiple cohorts of different ages. However, this regression has several problems, as illustrated in Figure 2. First, it is an estimate of total mortality so there must be an estimate of natural mortality to separate out fishing mortality. The cohorts at different ages are not observed equally so the selectivity of the sampling gear must be taken into consideration. In addition, the fishing mortality and recruitment are assumed to be in equilibrium, but the fishery selectivity will mean that different ages will experience different fishing mortality and fishing mortality may change over time. The different cohorts may experience different recruitment levels. There is also sampling error, since the data are not a census. All these assumptions can be fully explored in an integrated analysis with sufficient data.

Age data are not available for many species, so length-composition data are used. This brings the additional complication of converting length into age in age-structured models, which requires knowledge of the mean length-at-age and the variation of length-at-age using standard approaches in integrated stock assessment. Typically, the variation at a given age is assumed to be normally distributed (Figure 3). The mean length-at-age and variation of length-at-age are generally assumed to follow functional forms, so similar ages have similar mean length and variation. This means that changing the expected mean length for the oldest age also changes the expected mean length for slightly younger ages similarly, compounding the effect of growth on interpreting fishing mortality information from length-composition data. In particular, the mean length and variation of length for the oldest ages are highly influential on the estimated fishing mortality and biomass levels. The mean length of old fish and estimated fishing mortality are negatively correlated. Increasing the mean length of the oldest age causes the estimated relative abundance of the oldest age to reduce to fit the length composition of the largest fish (Figure 4, upper panel). Similarly, increasing the variation of length for the oldest age causes the estimated abundance of the oldest age to reduce to fit the length composition of the largest fish (Figure 4, lower panel). The reduction in the abundance of the oldest fish is generally estimated by increasing fishing mortality. Dome-shape selectivity can also explain the reduction, but most assessments assume that at least one selectivity curve is asymptotic to avoid "cryptic biomass" (e.g. older fish that are no longer vulnerable to the gear). This clearly illustrates that information about absolute abundance from age- or size-composition data depends on population dynamics (e.g. natural mortality, recruitment, and growth), fishing (e.g. selectivity), and sampling (e.g. random sampling variability and temporal variation in catchability). Our

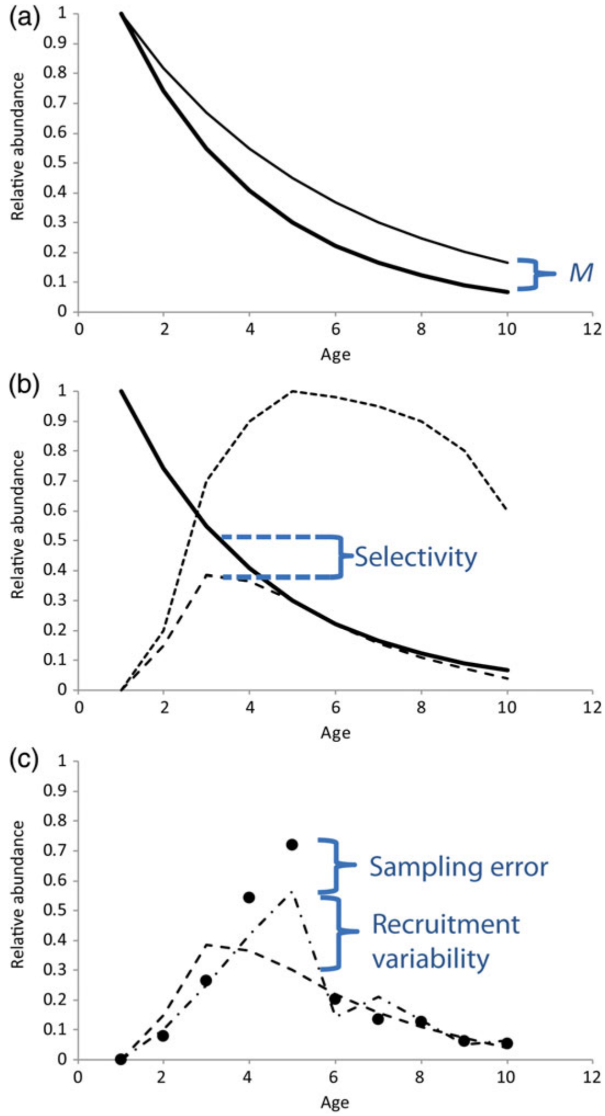

Figure 2. Influence of biological, fishing, and sampling processes on the information content of age-composition data. (a) The influence of natural mortality on the catch curve. The thick and thin lines represent with and without natural mortality, respectively. (b) The influence of selectivity. Dashed, selectivity; long dashed and solid lines represent the catch curve with and without selectivity, respectively. (c) The influence of recruitment variability and sampling error. The dash-dot and dashed lines represent the catch curve with and without recruitment variability, respectively. The dots include the sampling error.

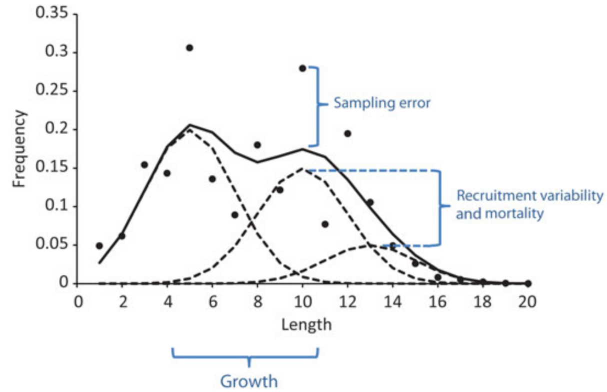

Figure 3. Influence of biological processes on information contained in length-composition data. The dashed lines represent the length-composition data from each of three cohorts. The solid line represents combined length-composition data. The dots include the sampling error.

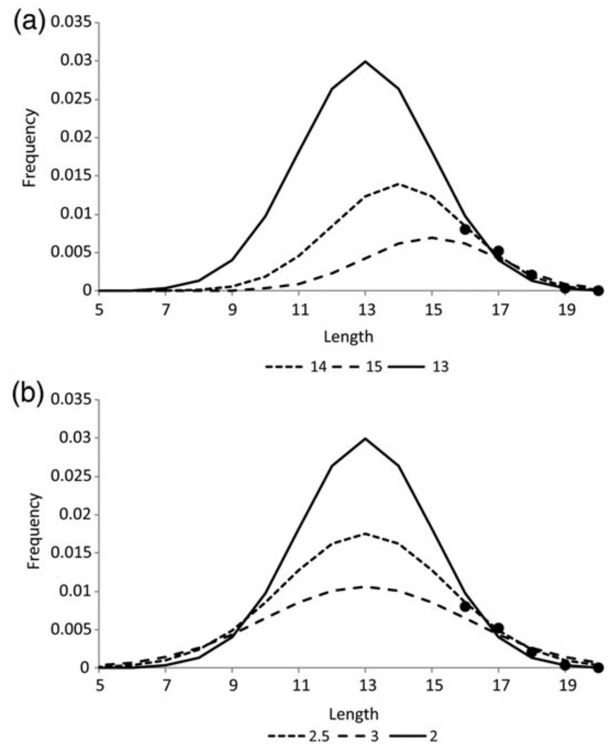

Figure 4. Illustration of (a) the mean length of the oldest fish and (b) the variation in length of the oldest fish on interpreting length-composition data. A lower peak of the length-composition curve relates to higher fishing mortality.

understanding of these model processes must be taken into consideration when interpreting the information on absolute abundance from age- or length-composition data.

Uncertainty in population, fishing, and sampling processes

The illustration of how absolute abundance information is extracted from relative indices of abundance and age- or size-composition data shows that it probably cannot be reliably done without knowledge of growth, recruitment, natural mortality, selectivity, and sampling processes. Similarly, these are important for determining fisheries management and reference points. Unfortunately, these are poorly known for many, if not most, species. Below, we summarize the uncertainties about these processes.

Growth

Growth is generally considered to be one of the most well-estimated processes due to the prevalence of ageing data. However, growth is often poorly estimated for many species due to lack of appropriate data, particularly for short-lived tropical species. There is often large uncertainty with variability from different ageing techniques and the ageing techniques have not been appropriately validated. For example, Chang and Maunder (2012) found a large variation in the estimates of growth parameters from different ageing methods (otolith, scale, and length-composition analysis) and among different stocks for dolphin fish (*Coryphaena hippurus*). Even for species for which ageing data have traditionally been thought to be good, ageing can turn out to be biased (Piner *et al.*, 2005; Stewart and Piner, 2007). For example, modes in the length-composition data from surveys for Pacific cod (*Gadus macrocephalus*) indicate that the otolith often puts down one additional ring, which produces underestimates of the growth rate if counted. Consequently, the current stock assessment estimates the ageing error internally (Thompson and Lauth, 2011). Ageing of yellowfin and bigeye tuna from daily rings in otoliths becomes difficult after about age 4 years (e.g. Wild, 1986), so assumptions must be made about the asymptotic length to extrapolate growth rates above age 4. However, the assessment model is fit to length-composition data so the evaluation of standard biomass and fishing mortality reference points are highly sensitive to the assumed asymptotic length (e.g. Aires-da-Silva *et al.*, submitted).

Use of length-composition data in integrated stock assessment models requires both the mean and the variation of length-at-age. Most growth studies do not report the variation of length-at-age. Therefore, the variation of length-at-age is often a missing component of stock assessment models, and a value has to be assumed. The estimates of management quantities can be sensitive to the assumed value.

Recruitment

The relationship between spawner stock size and the resulting recruitment has had a great deal of attention and research (Hilborn and Walters, 1992). However, for most stocks, little is known about the stock–recruitment relationship. The uncertainty results from several factors, including lack of contrast in spawning stock size, environmental variability, temporal autocorrelation in deviations about the stock–recruitment relationship, and error in estimating spawning stock size and recruitment (Hilborn and Walters, 1992; Quinn and Deriso, 1999). Typically, a stock–recruitment model (e.g. Ricker, 1954; Beverton and Holt, 1957) is fit to a time-series of stock and recruitment estimates. Stock size and recruitment are estimated in stock assessment models. It is intuitively better to estimate the stock–recruitment model inside the stock assessment model because it automatically takes the uncertainty of the stock size and recruitment into consideration (Maunder and Punt, 2013). However, several simulation studies have shown that the estimates of the stock–recruitment relationship are often highly uncertain or biased (Magnusson and Hilborn, 2007; Conn *et al.*, 2010; Lee *et al.*, 2012) and the benefits of estimating the stock recruitment model inside the stock assessment model are case-specific (Haltuch *et al.*, 2008, 2009). In particular, the estimate of steepness of the Beverton and Holt (1957) model is positively biased towards recruitment being independent of stock size, which produces greater productivity rates and often produces more optimistic stock status and management advice (e.g. greater catch limits). Regime shifts in recruitment caused by environmental conditions also frequently bias estimates of the stock–recruitment curve (Gilbert, 1997).

Given the inability to reliably estimate the stock–recruitment relationship, it is common to use default parameter values or borrow parameter values from similar stocks or species. Meta-analysis has been used to analyse stock and recruitment data from a collection of species and provide advice on parameter values that can be used for stocks for which there is little information (Myers *et al.*, 1999; Dorn, 2002; Punt and Dorn, 2013). Unfortunately, the reliability of advice from meta-analyses is questionable because knowledge of each time-series is needed to ensure that the results are not affected by other factors (e.g. regime shifts) and meta-analysis is vulnerable to the same biases as the analysis of individual stocks. Therefore, the capacity of meta-analysis to provide additional reliable information on uncertain parameters is unclear and its usefulness will probably differ among applications. Attempts to estimate parameter values for the stock–recruitment relationship based on the life-history theory (He *et al.*, 2006; Mangel *et al.*, 2010) do not, for any practical purposes, narrow the range of possible parameter values (Maunder, 2012).

Because of the inherent difficulties in reliably estimating the stock–recruitment relationship for most stocks, management strategies must be robust to uncertainty in the stock–recruitment relationship. Due to the flatness of the yield curve when steepness is high, less yield (in equilibrium) is lost by assuming a steepness lower than the actual steepness than assuming steepness is higher than it actually is, suggesting that assuming a lower steepness may be a more robust assumption in terms of maximizing equilibrium yield (Zhu *et al.*, 2012). Williams and Shertzer (2003) suggest that reference points should be developed for individual stocks by selecting appropriate proxy steepness values rather than using proxy reference points. Further work in this area is needed because of our inability to define the spawner–recruit relationship and its importance in both biological reference point calculations and future projections.

Natural mortality

Despite its importance in determining the productivity of a population and the consequent impact on sustainable yields, there has been a surprising lack of attention paid to developing direct (e.g. based on tagging data for the species of interest) estimates of natural mortality. Uncertainty could be reduced with more reliable estimates of natural mortality. However, most stocks lack a direct estimate of natural mortality and rely on indirect methods (e.g. correlations with life-history parameters (Pauly, 1980; Jensen, 1996; Gunderson, 1997) or from maximum age (Hoenig, 1983). The indirect methods are potentially biased with low precision as they are generated based on relationships with other unreliable estimates of natural mortality (Pascual and Iribarne, 1993; Maunder and Wong, 2011). In a recent review of the indirect methods, Kenchington (2013) stated "None of the 30 [indirect methods] can provide accurate estimates for every species, and none appears sufficiently precise for use in analytical stock assessments, while several perform so poorly as to have no practical utility." Many of the direct estimates of natural mortality for specific stocks are based on data that are used in contemporary integrated stock assessments (e.g. age composition), and even mark-recapture data can be included in integrated stock assessments to provide information about natural mortality (Maunder, 1998, 2001; Hampton and Fournier, 2001; Goethel *et al.*, 2011). Recent studies have shown some promise in estimating natural mortality inside integrated assessments conditioned on including the right data, periods of low stock size, and the absence of significant model misspecification (Lee *et al.*, 2011). The issues addressed above do not even include the reality that natural mortality is likely to vary with age, sex, stock density, and time (Vetter, 1988). Given our current understanding of this key component of stock productivity, it seems likely that in the short term substantial uncertainty will remain, at least into the near future. Based on updating and testing indirect estimators of natural mortality using information on 201 fish species, Then *et al.* (2015) recommend the use of their updated maximum age-based estimator when possible and an updated von Bertalanffy K-based method otherwise. Despite remaining concerns over these methods, Then *et al.*'s advice is probably the best currently available, except in cases for which reliable direct estimates are available or simulation analysis indicates estimates within the stock assessment model are reliable.

Selectivity

Selectivity has traditionally been assumed to be well estimated in integrated stock assessment models when ample age- or size-composition data are available. However, recent research has shown that the functional forms currently used are often too inflexible and misspecification of the selectivity curve can have substantial influence on stock assessment results and management advice (Crone *et al.*, 2013; Lee *et al.*, in press). Spatial variability in fishery and population structure can result in the estimation of unexpected selectivity shapes, including dome shape, even when the underlying gear selectivity is asymptotic (Sampson and Scott, 2011). Dome-shapeness of the selectivity curve is very influential when using length-composition data and interacts with the estimate of the asymptotic length to determine fishing mortality and absolute abundance. In addition, selectivity is likely to change over time, and assuming constant selectivity will bias results. The issues with misspecified selectivity parameterization go far beyond a biased estimate of the selectivity pattern and resulting catch-at-age. Misspecified selectivity patterns can result in systematic misfit to age- or size-composition data that can result in degraded fit to other important data (Francis, 2011; Lee *et al.*, in press). A growing body of research suggests that more flexible selection parameterization (e.g. non-parametric, time-varying) should be considered (Crone *et al.*, 2013).

Catchability

Catchability is the process that scales the index of relative abundance into absolute abundance. In early applications, it has been customary to assume that catchability of surveys is known. The assumed value was often based on research (e.g. gear studies; Somerton *et al.*, 1999); and the assumption provided direct information on population scale thereby reducing uncertainty of stock assessment results. Setting the catchability parameter stabilizes parameter estimation by eliminating correlation with other parameters. Unfortunately, usually additional research, such as towing commercial gear side by side with the survey gear, eventually shows that catchability (after expansion of the survey index of abundance to the population level) rarely is 1 (commonly less, see Harley and Myers, 2001) and varies by species and even time. Indices of relative abundance based on fisheries cpue data tend to have even more temporal variation in catchability because standardization methods are unable to account for all the variation in fishing vessels, fisher behaviour, and the environment (Maunder and Punt, 2004). Temporal trends in catchability (e.g. Harley *et al.* 2001) in addition to uncertainty in mean catchability are particularly problematic, since they will bias estimates of depletion levels. Therefore, uncertainty in both the average level of catchability and the variation over time can contribute substantially to the uncertainty in stock assessment results and estimates of management quantities.

In some cases where catchability was misspecified, the selectivity curve estimated inside the assessment for the index is distorted to account for the misspecification. In practice, catchability is typically estimated, but is confounded with other model parameters (e.g. natural mortality) and can be difficult to estimate inside the stock assessment model. Information on catchability comes from the information on absolute abundance from the index of relative abundance conditioned on the catch and the composition data as explained earlier.

Sampling and process error/variability

Contemporary stock assessment models typically account for both sampling (observation) and process error (variability). Data collected from a stock is not a census, but a sample of the population and different (hypothetical) samples will provide somewhat different information (sampling error). The nature of the sampling error should be taken into consideration when using the data to estimate the model parameters and the associated uncertainty. Contemporary statistically integrated stock assessment models generally use the sampling distribution to define the likelihood function, which is used as the basis for the objective function used in parameter and uncertainty estimation (e.g. maximum likelihood and Bayesian analysis). The distributional form and the parameters of the sampling distribution must be defined. However, it is the variance-related parameters (e.g. sample size of a multinomial used for composition data or the standard deviation of a lognormal distribution used for the relative index of abundance data) that tend to be most influential because they determine the weight each dataset gets in the parameter estimation and the overall uncertainty in parameter estimates. Commonly, the variance parameters are fixed, but methods are available to estimate the variance parameters (Deriso *et al.*, 2007; McAllister and Ianelli, 1997; Maunder, 2011). The estimated variance parameters are statistically determined, based on how consistent the data is with themselves (e.g. low variance is often associated with a relative index of abundance that does not change much from one year to the next despite how reliable it is), the assumed population dynamics, the other data, and how many and what type of parameters are estimated to allow the model to fit the data closely (e.g. the more parameters estimated, the closer the fit and the less sampling variance estimated).

Process error is additional variability in the population (e.g. recruitment), fishing (e.g. selectivity), or sampling processes (e.g. survey catchability) that are not represented by the main structure of the model. The most commonly modelled process error is temporal variation in recruitment. In this context, the term process error is somewhat confusing since the recruitment strength each year is an important component of the model and intuitively should be part of the main model parameters. However, without good age-composition data for each year, the annual recruitment may be uncertain for some years. To improve estimation, the assumed distribution of recruitment variation (e.g. lognormal around the stock–recruitment relationship) can be taken advantage of (Maunder and Deriso, 2003). To estimate the variance of the recruitment distribution, the process error should be treated as a random effect (or, equivalently, a state-space model used) and integrated out. This is a computationally intensive multidimensional integral in stock assessment models and has had limited application outside a Bayesian context [see Nielsen and Berg (in press) for an exception].

The variance parameters used in sampling distributions generally require the process model, including the model that relates the sampled data to the population dynamics, to be correctly specified. However, the process model is typically not modelled correctly. For example, population dynamics (e.g. natural mortality), fishery (e.g. selectivity), and sampling (e.g. survey catchability) processes likely vary over time, but are assumed to be time invariant in many assessment models. Ignoring this variability will result in the model not fitting the observed data as well as suggested by the variance parameters of the sampling distributions. Estimation of the variance parameters of the sampling distributions inside the stock assessment model will result in the process error being assumed by the observation error variances. One example is the inflation of standard deviations for survey data because of temporal variability in catchability due to factors such as the environmental conditions. Another is the reduction in the effective sample size of composition data due to unmodelled correlation in the sampling process (i.e. many species school by size and repeated samples from a purse-seine set on a single school will be correlated). The observation and process errors are typically assumed to be random, but often there are correlations particularly in process errors, which show temporal autocorrelation. It is not clear if accommodating the unmodelled process error (e.g. temporal variation in selectivity) in the observation error is appropriate. It may be better to estimate the observation error outside the model (e.g. bootstrapping the composition data sampling design) and explicitly modelling the process error inside the model (e.g. time varying selectivity). This would require using computationally intensive methods [e.g. random effects (Fournier *et al.*, 2012; Nielsen and Berg, in press) or cross-validation (e.g. Maunder and Harley, 2011)] to estimate the variance parameters of the process error, which may be impractical for some applications. However, promising less computationally intensive approximations are available (Thompson and Lauth, 2012).

Diagnostics

Diagnostics are important tools to determine if the model fits the data adequately and that the model is well specified. Unfortunately, there are few standard diagnostics tools available for integrated stock assessment models that can provide the analyst with all the information needed to determine model performance. Here, we describe two recently developed diagnostic tools that can be used to evaluate the information content of data about absolute abundance and help determine if the model is correctly specified. However, more thorough testing of these methods is needed to determine their usefulness (e.g. Wang *et al.*, in press).

R0 likelihood component profile diagnostic

The R0 likelihood component profile diagnostic estimates all model parameters while fixing the population scaling parameter (often the virgin recruitment) at different values and plotting the resultant likelihood (or commonly, the negative log-likelihood) value for each data component against this parameter (Francis, 2011; Lee *et al.*, in press). The likelihood profiles of each data component usually follow a smooth parabolic curve, indicating the value with most support and the amount of uncertainty in that support. Different maxima (minima if the negative log-likelihood is used) among data components indicate possible conflict in the data sources about absolute abundance. The higher the gradient in the likelihood profile the more influential that data source on the model's estimate of scale. We recommend also applying the R0 profile to data simulated without error from a model structured the same as the application. This R0 profile (simulated R0 profile) will illustrate the information content of the data expected if the model is correctly specified and any differences from the actual R0 profile indicates conflict in the data or model misspecification.

The R0 profile has been used to diagnose selectivity misspecification (Lee *et al.*, in press; Wang *et al.*, in press). Despite age- and size-composition data possibly providing substantial information on absolute abundance as shown by R0 component profiles using correctly specified models (Wang *et al.*, in press), there has been a trend to deemphasize the abundance content of age- and size-composition data (Francis, 2011). This is because relatively minor model misspecification (e.g. a too inflexible selectivity curve) can have a large impact on the information about absolute abundance contained in the composition data (Lee *et al.*, in press). The weighting of the age- and size-composition data (e.g. the sample size) can be decreased until the R0 profile shows that the age- and size-composition data have relatively low influence compared with the other data. The idea is to remove catch out of the population at about the right size without providing additional information on abundance. Information focusing estimation that estimates selectivity and recruitment deviates only when fitting to the age- and size-composition data is a new approach in fisheries stock assessment that might achieve this (T. Kitakado, pers. comm.).

Age-structured production model diagnostic

A comparison of the results of the age-structured production model (ASPM) to those from a model estimating the full dynamics and fitting to all the data (e.g. an integrated analysis) can be used as a diagnostic tool. Similar to the R0 profile, this diagnostic can be used to evaluate model misspecification and the contribution of age- or size-composition data to estimates of absolute abundance. Unlike the profile diagnostic, comparison of the ASPM also can be used to evaluate if the catch alone (taken out of approximately the correct ages) can explain trends in the index of abundance. When catch does explain indices with good contrast (e.g. declining and increasing trends), it suggests that a production function is apparent in the data, therefore providing evidence that the index is a reasonable proxy of stock trend. If catch cannot explain the index, then either the stock is recruitment driven (or did not decline far enough to show changes in recruitment due to the stock–recruitment relationship), the model is incorrect, or the index of relative abundance is not proportional to abundance. These kinds of results are especially important in stock assessments based on fishery-dependent cpue whose true relationship with trends in abundance is always in doubt.

To perform the diagnostic, the parameters of the selectivity curve in the ASPM are fixed at those estimated from the fully dynamic model, the annual recruitment deviates are not estimated (fixed at zero so that recruitment follows the stock–recruitment relationship), and the age- and size-composition data are not used. If the age- and size-composition data are not informing the absolute abundance or the abundance trend and there is no pattern in recruitment, the results from the age-structured model should be similar to the full dynamics analysis. To determine if a recruitment trend is causing any differences, recruitment deviates can be added to the ASPM.

Two hypothetical examples based on experience using the diagnostic are used to illustrate the method. The first hypothetical example illustrates how the age-composition data are influencing the absolute abundance and reducing the weight on the composition data causes the integrated model to estimate similar abundance levels as the ASPM (Figure 5, upper panel). The second hypothetical example illustrates how a pattern in recruitment deviates in the

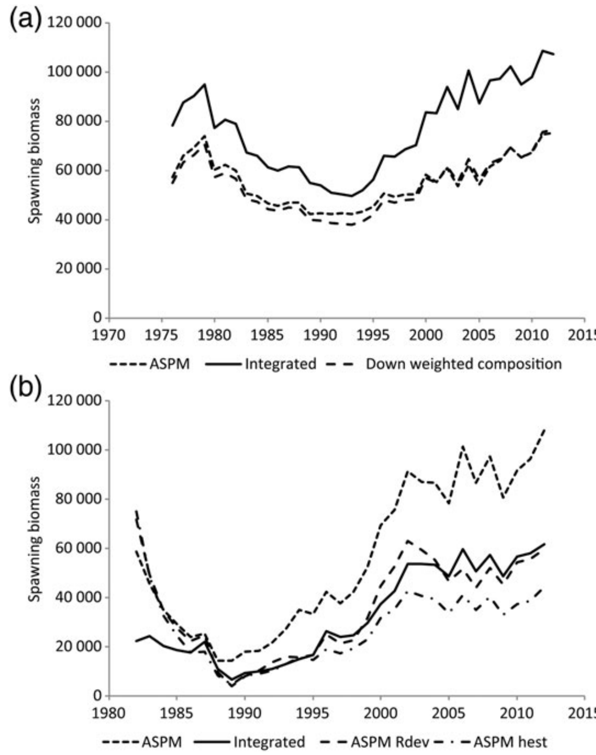

Figure 5. Hypothetical results of applying the production model diagnostic. (a) Composition data are influencing the estimates of absolute abundance. (b) Misspecified steepness of the stock–recruitment relationship.

integrated model causes a difference with the ASPM (Figure 5, lower panel). Adding recruitment deviates to the ASPM gives similar results to the integrated model. Exploration of the recruitment deviates (positive deviates for low stock size and negative deviates for high stock size) suggests that the stock–recruitment relationship was misspecified at too low of a value. Estimating the steepness of the Beverton and Holt (1957) stock–recruitment relationship in the ASPM (without recruitment variability), which was originally fixed, produced results similar to the integrated model, supporting the misspecified stock–recruitment relationship hypothesis. We have only begun to use the ASPM diagnostic and without more testing, its utility remains speculative.

Impediments to progress

Despite the importance of commercial fisheries worldwide, there remains tremendous uncertainty in even our basic understanding of stock productivity and condition. Perhaps, the most basic impediment to improving our understanding is a lack of relevant and reliable data needed to assess the status of stocks. For many species, the lack of quality data may result from a reasonable decision because the cost of collecting good data cannot be justified based on the value of the fishery. However, for other more valuable or important stocks, the funding levels of science is either insufficient to collect the necessary data or has not been allocated in ways that would have improved the applied understanding of fisheries science.

The lack of sufficient funding cannot be blamed entirely on a lack of interest in fisheries or fisheries management. Both private and public groups have shown a willingness to spend, often large sums of money, in high-profile cases. There are many examples for which industry is funding science and data collection (e.g. the International Seafood Sustainable Foundation) or is offsetting the costs of public spending. Private interest groups are also spending sizeable amounts of money in fisheries; often willing to contract highly paid consultants when they perceive that the assessment and management is inappropriate and impacting the levels of catch or endangering the sustainability of the resource. However, in most situations, it is government funding of data collection and research that provides the information used to manage stocks.

If adequate funding is being spent on fisheries, but unacceptable uncertainties remain, we may need to reassess the ways we are spending those finite resources. In situations where basic data are missing, are we spending applied science resources on less immediate needs associated with long-term goals of ecosystem-based management or on what may be more theoretical research on climate change, marine protected areas, and environmental covariates, which could be arguably less useful for routine management of fish stocks. We need to carefully assess if the level of funding and those projects funded for research and management is aligned with the economic value of the resources and with benefits of the resources to society in general.

As large as the data gap is, the lack of highly trained stock assessment scientists and the finite time available for the limited number of qualified scientists that exist currently (Berkson *et al.*, 2009a, b) may turn out to be an even larger obstacle to improving fisheries management. To be useful for fisheries management, the multiple sources of data available have to be processed using ever more complex quantitative methods to extract and combine the information they contain. As discussed in this paper, research on improving our understanding of technical aspects of fisheries stock assessment modelling (particularly since stock assessment is done with limited data and simplifying assumptions need to be made) still need to be conducted. However, all these pursuits need considerable attention from an already overburdened pool of scientists. As more stocks are added to the list of species requiring quantitative management advice, even more requirements will be placed on the existing stock assessment scientists, further compounding the issue.

Efforts are underway to rectify the deficiency in the number of qualified stock assessment scientists; however, they may not be sufficient. Recent hiring in the United States of stock assessment academics to train graduate students is an important step, but may be too slow and produce too few scientists to make up for those lost to retirement. Part of the problem is a lack of critical mass in academic institutions due to the specialized nature of stock assessment, and improvements in this situation may require collaboration among academic institutions. Too much time is spent on attempting to train life science graduates using short courses in the fundamentals of stock assessment. Short courses are insufficient to train qualified scientists and keep their interest. These courses usually lack the basic mathematical content to underpin participants understanding of stock assessment concepts. Applied mathematics graduates are needed in stock assessment, but there are not enough role models at universities. A similar situation is apparent worldwide and is particularly of concern in small or developing nations for which researchers with expertise in stock assessment tend to move to the larger developed countries, where there is a higher concentration of talent, to improve their financial situations, or to advance their careers. Complicating the situation is the fact that graduate students trained overseas often stay overseas. A 1-year post-PhD diploma programme to retrain quantitative ecologists, statisticians, and other quantitatively trained scientists may be the quickest way to train the stock assessment scientists needed. The programme may require collaboration between multiple universities and management agencies to contain the courses needed and provide mentoring and practical experience. However, such a programme has not been implemented and therefore is unproven.

A greater commitment to improving the applied aspects of fisheries science is needed from all scientists including those who do not have direct management-related responsibilities. Solutions to the basic problems and not just influential publications are needed. For example, it may be possible to publish a paper on the first archival tag applied to a species where new behaviour is often discovered, but it is the next 100 recoveries that are needed to infer population level processes useful for stock assessment. Demonstrated improvement of fisheries management should be rewarded at least as much as influential publications in both career advancement and future funding.

Finally, our interactions with the larger public needs to be improved. Too often the public's perception of resource health is taken from controversial information published in high-profile journals. There have been several influential articles published in high-profile journals in the past decade that have focused on the dire state of the world's fish stocks (e.g. Casey and Myers, 1998; Pauly *et al*., 1998; Roberts *et al*., 2001; Conover and Munch, 2002; Baum *et al*., 2003; Myers and Worm, 2003; Worm *et al*., 2006). These articles gained considerable media attention and may have been more publishable for that reason (Hilborn, 2006). However, some high-profile articles have been shown to be seriously flawed, but only after being read by the scientific community, policy-makers, and the public. The costs of misinformation extend beyond the public perception as they may further reduce research time for practicing stock-assessment scientists who, must spend time responding to those controversial results (Hilborn, 2006) and attract graduate students away from basic stock assessment research. It is also concerning that rebuttals to these articles (e.g. Hilborn, 2002; Walters, 2003; Hampton *et al*., 2005; Essington *et al*., 2006; Maunder *et al*., 2006; Polacheck, 2006; Sibert *et al*., 2007; Branch, 2008) generally garner far less public attention as they are often published in the technical literature with less mainstream media fanfare. A better peer review and reward system is needed to ensure research is focused and constructive.

Conclusion

Interpretation of data used in fisheries assessment and management requires the knowledge of population (e.g. growth, natural mortality, and recruitment), fishing (e.g. selectivity), and sampling processes. Without this knowledge, assumptions must be made, either implicitly or explicitly, based on the methods used. Incorrect assumptions can have a substantial impact on stock assessment results and management advice. Unfortunately, there is a lack of understanding of these processes for most, if not all, stocks and even for those processes that have traditionally been assumed to be well understood (e.g. growth and selectivity). There are various reasons for the lack of understanding, but fisheries science is at a point at which decisions must be made—whether comprehensive research should be focused on understanding these processes or whether management should be designed to be robust, not just conservative, to the uncertainty. If understanding is the path to be taken, then substantially more investment is needed in training stock-assessment scientists, stock assessment research, biological studies, and data collection in a coordinated and focused approach. Some of these problems are related the current socio-economic situation and, with any luck, may improve in the future. However, it is likely that many of the issues will still remain without a concerted effort to rectify them. Several efforts, but of limited scope, are being made in this respect, such as research towards developing a guide to good stock assessment practices [e.g. the Center for the Advancement of Population Assessment Methodology (CAPAM)].

Acknowledgements

This research addresses the good practices guide to stock assessment programme of the Center for the Advancement of Population Assessment Methodology (CAPAM). Andre Punt, Doug Butterworth, Laurie Kell, William Bayliff, and an anonymous reviewer provided comments that improved the manuscript. The opinions expressed in this manuscript reflect those of the authors and should not be considered to represent their agency or institution.

References

Aires-da-Silva, A. M., Maunder, M. N., Schaefer, K. M., and Fuller, D. W. Submitted. Improved growth estimates from integrated analysis of direct aging and tag-recapture data: an illustration with bigeye tuna (*Thunnus obesus*) of the eastern Pacific Ocean with implications for management. Fisheries Research.

Baum, J. K., Kehler, D., Myers, R. A., Worm, B., Harley, S. J., and Doherty, P. A. 2003. Collapse and conservation of shark populations in the Northwest Atlantic. Science, 299: 389–392.

Begley, J., and Howell, D. 2004. An overview of Gadget, the Globally applicable Area-Disaggregated General Ecosystem Toolbox. ICES Conference and Meeting Documents 2004/FF: 13. http://www.hafro.is/gadget/.

Berkson, J., Hudy, S. F., McMullin, S. L., Murphy, B. R., Prager, M. H., Kelly, M. J., and Thompson, N. B. 2009b. Addressing the shortage of stock assessment scientists through undergraduate workshops. Fisheries, 34: 220–227.

Berkson, J., Hunt, K. M., Whitehead, J. C., Murie, D. J., Kwak, T. J., and Boreman, J. 2009a. Is there a shortage of fisheries stock assessment scientists? Fisheries, 34: 217–219.

Beverton, R. J. H., and Holt, S. J. 1957. On the dynamics of exploited fish populations. UK Ministry of Agriculture, Fish and Fisheries Investigations (Series 2), 19.

Branch, T. A. 2008. Short communication: not all fisheries will be collapsed in 2048. Marine Policy, 32: 38–39.

Bull, B., Francis, R. I. C. C., Dunn, A., McKenzie, A., Gilbert, D. J., and Smith, M. H. 2005. CASAL (C++ algorithmic stock assessment laboratory): CASAL user manual v2.07–2005/08/21. NIWA Technical Report, 127. http://www.niwascience.co.nz/ncfa/tools/casal.

Butterworth, D. S., Cochrane, K. L., and De Oliveira, J. A. A. 1997. Management procedures: a better way to management fisheries? The South African experience. *In* Global Trends: Fisheries Management, pp. 83–90. Ed. by E. L. Pikitch, D. D. Huppert, and M. P. Sissenwine. American Fisheries Society Symposium 20, Bethesda.

Butterworth, D. S., and Punt, A. E. 1999. Experiences in the evaluation and implementation of management procedures. ICES Journal of Marine Science, 56: 985–998.

Casey, J. M., and Myers, R. A. 1998. Near extinction of a large widely distributed fish. Science, 281: 690–692.

Chang, S-K., and Maunder, M. N. 2012. Aging material matters in the estimation of von Bertalanffy growth parameters for dolphinfish (*Coryphaena hippurus*). Fisheries Research, 119–120: 147–153.

Clark, W. G. 1991. Groundfish exploitation rates based on life history parameters. Canadian Journal of Fisheries and Aquatic Sciences, 48: 734–750.

Clark, W. G. 2002. F35% revisited ten years later. North American Journal of Fisheries Management, 22: 251–257.

Conn, P. B., Williams, E. H., and Shertzer, K. W. 2010. When can we reliably estimate the productivity of fish stocks? Canadian Journal of Fisheries and Aquatic Sciences, 67: 1–13.

Conover, D. O., and Munch, S. B. 2002. Sustaining fisheries yields over evolutionary time scales. Science, 297: 94–96.

Cope, J. M. 2013. Implementing a statistical catch-at-age model (Stock Synthesis) as a tool for deriving overfishing limits in data-limited situations. Fisheries Research, 142: 3–14.

Crone, P., Maunder, M., Valero, J., McDaniel, J., and Semmens, B. 2013. Selectivity: theory, estimation, and application in fishery stock assessment models. Center for the Advancement of Population Assessment Methodology (CAPAM) Workshop Series Report, 1. 46 pp.

de la Mare, W. K. 1986. Simulation studies on management procedures. Reports of the International Whaling Commission, 36: 429–450.

De Oliveira, J. A. A., Butterworth, D. S., and Johnston, S. J. 1998. Progress and problems in the application of management procedures to South Africa's major fisheries. *In* Fishery Stock Assessment Models, pp. 513–530. Ed. By F. Funk, T. J. Quinn, II, J. Heifetz, J. N. Ianelli, J. E. Powers, J. J. Schweigert, P. J. Sullivan, *et al.* Alaska Sea Grant College Program Report, AK-SG-98-01, University of Alaska Fairbanks.

Deriso, R. B., Maunder, M. N., and Skalski, J. R. 2007. Variance estimation in integrated assessment models and its importance for hypothesis testing. Canadian Journal of Fisheries and Aquatic Sciences, 64: 187–197.

Deriso, R. B., Quinn, T. J., II, and Neal, P. R. 1985. Catch-age analysis with auxiliary information. Canadian Journal of Fisheries and Aquatic Sciences, 42: 815–824.

Dorn, M. W. 2002. Advice on West Coast rockfish harvest rates from Bayesian meta-analysis of stock–recruit relationships. North American Journal of Fisheries Management, 22: 280–300.

Essington, T., Beaudreau, A., and Wiedenmann, J. 2006. Fishing through marine food webs. Proceedings of the National Academy of Sciences of the USA, 103: 3171–3175.

Fournier, D. A., and Archibald, C. P. 1982. A general theory for analyzing catch at age data. Canadian Journal of Fisheries and Aquatic Sciences, 39: 1195–1207.

Fournier, D. A., Hampton, J., and Sibert, J. R. 1998. MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. Canadian Journal of Fisheries and Aquatic Sciences, 55: 2105–2116.

Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M. N., Nielsen, A., *et al.* 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optimization Methods and Software, 27: 233–249.

Francis, R. I. C. C. 2011. Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences, 68: 124–1138.

Gilbert, D. J. 1997. Towards a new recruitment paradigm for fish stocks. Canadian Journal of Fisheries and Aquatic Sciences, 54: 969–977.

Goethel, D. R., Quinn II, T. J., and Cadrin, S. X. 2011. Incorporating spatial structure in stock assessment: movement modeling in marine fish population dynamics. Reviews in Fisheries Science, 19: 119–136.

Goodyear, P. C. 1996. Variability of fishing mortality by age: consequences for maximum sustainable yield. North American Journal of Fisheries Management, 16: 8–13.

Gulland, J. A. 1965. Estimation of mortality rates. Annex to Arctic Fisheries Working Group Report (meeting in Ham burg, January 1965). International Council for the Exploitation of the Sea, Conference and Meeting Documents, Document 3 (mimeo), Copenhagen.

Gunderson, D. R. 1997. Trade-off between reproductive effort and adult survival in oviparous and viviparous fishes. Canadian Journal of Fisheries and Aquatic Sciences, 54: 990–998.

Haltuch, M. A., Punt, A. E., and Dorn, M. W. 2008. Evaluating alternative estimators of fishery management reference points. Fisheries Research, 94: 290–303.

Haltuch, M. A., Punt, A. E., and Dorn, M. W. 2009. Evaluating the estimation of fishery management reference points in a variable environment. Fisheries Research, 100: 42–56.

Hampton, J., and Fournier, D. A. 2001. A spatially disaggregated, length-based, age-structured population model of yellowfin tuna (*Thunnus albacares*) in the western and central Pacific Ocean. Marine and Freshwater Research, 52: 937–963.

Hampton, J., Sibert, J. R., Kleiber, P., Maunder, M. N., and Harley, S. J. 2005. Decline of Pacific tuna populations exaggerated? Nature, 434: E1–E2.

Harley, S. J., and Myers, R. A. 2001. Hierarchical Bayesian models of length-specific catchability of research trawl surveys. Canadian Journal of Fisheries and Aquatic Sciences, 58: 1569–1584.

Harley, S. J., Myers, R. A., and Dunn, A. 2001. Is catch-per-unit-effort proportional to abundance? Canadian Journal of Fisheries and Aquatic Sciences, 58: 1760–1772.

He, X., Mangel, M., and MacCall, A. 2006. A prior for steepness in stock–recruitment relationships based on an evolutionary persistence principle. Fishery Bulletin US, 104: 428–433.

Hilborn, R. 2002. Marine reserves and fisheries management. Science, 295: 1233–1234.

Hilborn, R. 2006. Faith-based Fisheries. Fisheries, 31: 554–555.

Hilborn, R., and Liermann, M. 1998. Standing on the shoulders of giants: learning from experience in fisheries. Reviews in Fish Biology and Fisheries, 8: 273–283.

Hilborn, R., Maunder, M. N., Parma, A., Ernst, B., Payne, J., and Starr, P. J. 2000. Documentation for a general age-structured Bayesian stock assessment model: code named Coleraine. Fisheries Research Institute, University of Washington, FRI/UW 00/01. Software available at http://fish.washington.edu/research/coleraine/.

Hilborn, R., and Walters, C. J. 1992. Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty. Chapman and Hall, New York.

Hoenig, J. M. 1983. Empirical use of longevity data to estimate total mortality rates. Fishery Bulletin US, 82: 898–903.

Jensen, A. L. 1996. Beverton and Holt life history invariants result from optimal tradeoff of reproduction and survival. Canadian Journal of Fisheries and Aquatic Sciences, 53: 820–822.

Jurado-Molina, J., Livingston, P. A., and Ianelli, J. N. 2005. Incorporating predation interactions in a statistical catch-at-age model for a predator-prey system in the eastern. Bering Sea Canadian Journal of Fisheries and Aquatic Sciences, 62: 1865–1873.

Kenchington, T. J. 2013. Natural mortality estimators for information-limited fisheries. Fish and Fisheries, doi:10.1111/faf.12027.

Lee, H-H., Maunder, M. N., Piner, K. R., and Methot, R. D. 2011. Estimating natural mortality within a fisheries stock assessment model. An evaluation using simulation analysis based on twelve stock assessments. Fisheries Research, 109: 89–94.

Lee, H-H., Maunder, M. N., Piner, K. R., and Methot, R. D. 2012. Can steepness of the stock-recruitment relationship be estimated in fishery stock assessment models? Fisheries Research, 125–126: 254–261.

Lee, H-H., Piner, K. R., Methot, R. D., and Maunder, M. N. in press. Use of likelihood profiling over a global scaling parameter to structure the population dynamics model: an example using blue marlin in the Pacific Ocean. Fisheries Research, DOI: 10.1016/j.fishres.2013.12.017.

Magnusson, A., and Hilborn, R. 2007. What makes fisheries data informative? Fish and Fisheries, 8: 337–358.

Mangel, M., Brodziak, J., and DiNardo, G. 2010. Reproductive ecology and scientific inference of steepness: a fundamental metric of population dynamics and strategic fisheries management. Fish and Fisheries, 11: 89–104.

Maunder, M. N. 1998. Integration of Tagging and Population Dynamics Models in Fisheries Stock Assessment. PhD thesis, University of Washington.

Maunder, M. N. 2001. Integrated Tagging and Catch-at-Age ANalysis (ITCAAN). *In* Spatial Processes and Management of Fish Populations, pp. 123–146. Ed. by G. H. Kruse, N. Bez, A. Booth, M. W. Dorn, S. Hills, R. N. Lipcius, D. Pelletier, *et al.* Alaska Sea Grant College Program Report, AK-SG-01-02, University of Alaska Fairbanks.

Maunder, M. N. 2002. The relationship between fishing methods, fisheries management and the estimation of MSY. Fish and Fisheries, 3: 251–260.

Maunder, M. N. 2003. Is it time to discard the Schaefer model from the stock assessment scientist's toolbox? Fisheries Research, 61: 145–149.

Maunder, M. N. 2011. Review and evaluation of likelihood functions for composition data in stock-assessment models: estimating the effective sample size. Fisheries Research, 109: 311–319.

Maunder, M. N. 2012. Evaluating the stock-recruitment relationship and management reference points: application to summer flounder (*Paralichthys dentatus*) in the U.S. mid-Atlantic. Fisheries Research, 125–126: 20–26.

Maunder, M. N., and Deriso, R. B. 2003. Estimation of recruitment in catch-at-age models. Canadian Journal of Fisheries and Aquatic Sciences, 60: 1204–1216.

Maunder, M. N., and Harley, S. J. 2011. Using cross validation model selection to determine the shape of nonparametric selectivity curves in fisheries stock assessment models. Fisheries Research, 110: 283–288.

Maunder, M. N., and Punt, A. E. 2004. Standardizing catch and effort data: a review of recent approaches. Fisheries Research, 70: 141–159.

Maunder, M. N., and Punt, A. E. 2013. A review of integrated analysis in fisheries stock assessment. Fisheries Research, 142: 61–74.

Maunder, M. N., Sibert, J. R., Fonteneau, A., Hampton, J., Kleiber, P., and Harley, S. 2006. Interpreting catch-per-unit-of-effort data to asses the status of individual stocks and communities. ICES Journal of Marine Science, 63: 1373–1385.

Maunder, M. N., and Wong, R. A. 2011. Approaches for estimating natural mortality: application to summer flounder (*Paralichthys dentatus*) in the U.S. mid-Atlantic. Fisheries Research, 111: 92–99.

McAllister, M. K., and Ianelli, J. N. 1997. Bayesian stock assessment using catch-age data and the sampling/ importance resampling algorithm. Canadian Journal of Fisheries and Aquatic Sciences, 54: 284–300.

McAllister, M. K., Pikitch, E. K., Punt, A. E., and Hilborn, R. 1994. A Bayesian approach to stock assessment and harvest decisions using the sampling/importance resampling algorithm. Canadian Journal of Fisheries and Aquatic Sciences, 52: 2673–2687.

Methot, R. D. 1990. Synthesis model: an adaptable framework for analysis of diverse stock assessment data. International North Pacific Fishery Commission Bulletin, 50: 259–277.

Methot, R. D., and Wetzel, C. 2013. Stock Synthesis: a biological and statistical framework for fish stock assessment 557 and fishery management. Fisheries Research, 142: 86–99.

Myers, R. A., Bowen, K. G, and Barrowman, N. J. 1999. Maximum reproductive rate of fish at low population sizes. Canadian Journal of Fisheries and Aquatic Sciences, 56: 2404–2419.

Myers, R. A., and Worm, B. 2003. Rapid worldwide depletion of predatory fish communities. Nature, 423: 280–283.

Nielsen, A., and Berg, C. in press. Estimation of time-varying selectivity in stock assessments using state-space models. Fisheries Research.

Pascual, M. A., and Iribarne, O. O. 1993. How good are empirical predictions of natural mortality? Fisheries Research, 16: 17–24.

Pauly, D. 1980. On the interrelationships between natural mortality, growth parameters, and mean environmental temperature in 175 stocks. Journal du Conseil International pour l'Exploration de la Mer, 39: 175–192.

Pauly, D., Christensen, V., Dalsgaard, J., Froese, R., and Torres, F. 1998. Fishing down marine food webs. Science, 279: 860–863.

Pella, J. J., and Tomlinson, P. K. 1969. A generalized stock production model. Inter-American Tropical Tuna Commission Bulletin, 13: 421–458.

Piner, K. R., Hamel, O. S., Menkel, J. L., Wallace, J. R., and Huthchinson, C. E. 2005. Age validation of canary rockfish (*Sebastes pinniger*) from off the Oregon coast (USA) using the bomb radiocarbon method. Canadian Journal of Fisheries and Aquatic Sciences, 62: 1060–1066.

Polacheck, T. 2006. Tuna longline catch rates in the Indian Ocean: did industrial fishing result in a 90% rapid decline in the abundance of large predatory species? Marine Policy, 30: 470–482.

Pope, J. G. 1972. An investigation of the accuracy of virtual population analysis using Cohort analysis. International Commission for Northwest Atlantic Fisheries Research Bulletin, 9: 65–74.

Punt, A. E., and Dorn, M. 2013. Comparisons of Meta-analytic methods for deriving a probability distribution for the steepness of the stock-recruitment relationship. Fisheries Research, 149: 43–54.

Punt, A. E., and Hilborn, R. 1997. Fisheries stock assessment and decision analysis: the Bayesian approach. Reviews in Fish Biology and Fisheries, 7: 35–63.

Punt, A. E., Huang, T-C., and Maunder, M. N. 2013. Review of integrated size-structured models for stock assessment of hard-to-age crustacean and mollusc species. ICES Journal of Marine Science, 70: 16–33.

Punt, A. E., and Kennedy, R. B. 1997. Population modelling of Tasmanian rock lobster, *Jasus edwardsii*, resources. Marine and Freshwater Research, 48: 967–980.

Quinn, T. J., II. 2003. Ruminations on the development and future of population dynamics models in fisheries. Natural Resource Modeling, 16: 341–392.

Quinn, T. J., II, and Deriso, R. B. 1999. Quantitative Fish Dynamics. Oxford University Press, New York, NY.

Ricker, W. E. 1954. Stock and recruitment. Journal of the Fisheries Research Board of Canada, 11: 559–623.

Roberts, C. M., Bohnsack, J. A., Gell, F., Hawkins, J. P., and Goodridge, R. 2001. Effects of marine reserves on adjacent fisheries. Science, 294: 1920–1923.

Sampson, D. B., and Scott, R. D. 2011. A spatial model for fishery age-selection at the population level. Canadian Journal of Fisheries and Aquatic Sciences, 68: 1077–1086.

Schaefer, M. B. 1954. Some aspects of the dynamics of populations important to the management of commercial marine fisheries. Inter-American Tropical Tuna Commission Bulletin, 1: 25–56.

Shepherd, J. G. 1999. Extended survivors analysis: an improved method for the analysis of catch-at-age data and abundance indices. ICES Journal of Marine Science, 56: 584–591.

Sibert, J., Hampton, J., Kleiber, P., and Maunder, M. N. 2007. Fishing for Good News: response. Science, 316: 200–201.

Smith, A. D. M., Sainsbury, K. J., and Stevens, R. A. 1999. Implementing effective fisheries-management systems: management strategy evaluation and the Australian partnership approach. ICES Journal of Marine Science, 56: 967–979.

Somerton, D., Ianelli, J., Walsh, S., Smith, S., Godø, O. R., and Ramm, D. 1999. Incorporating experimentally derived estimates of survey trawl efficiency into the stock assessment process: a discussion. ICES Journal of Marine Science, 56: 299–302.

Stewart, I. J., and Piner, K. R. 2007. Simulation of the estimation of the functional form of ageing bias in an integrated stock assessment of canary rockfish using bomb radiocarbon based ages. Marine and Freshwater Research, 58: 905–913.

Stramma, L., Johnson, G. C., Sprintall, J., and Mohrholz, V. 2008. Expanding oxygen-minimum zones in the tropical oceans. Science, 320: 655–658.

Then, A. Y., Hoenig, J. M., Hall, N. G., and Hewitt, D. A. 2015. Evaluating the predictive performance of empirical estimators of natural mortality rate using information on 201 fish species. ICES Journal of Marine Science, 72: 82–92.

Thompson, G. G., and Lauth, R. R. 2011. Assessment of the Pacific Cod Stock in the Eastern Bering Sea and Aleutian Islands Area. NPFMC Bering Sea and Aleutian Islands SAFE, pp. 269–476.

Thompson, G. G., and Lauth, R. R. 2012. Assessment of the Pacific cod stock in the Eastern Bering Sea and Aleutian Islands Area. In: Plan Team for Groundfish Fisheries of the Bering Sea/Aleutian Islands (Compiler), Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Bering Sea/Aleutian Islands regions. North Pacific Fishery Management Council, Anchorage, AK, pp. 245–544.

Vetter, E. F. 1988. Estimation of natural mortality in fish stocks: a review. Fishery Bulletin US, 86: 25–43.

Walters, C. J. 2003. Folly and fantasy in the analysis of spatial catch rate data. Canadian Journal of Fisheries and Aquatic Sciences, 60: 1433–1436.

Wang, S-P., Maunder, M. N., Piner, K., Aires-da-Silva, A., and Lee, H-H. in press. Evaluation of R0 profiling as a diagnostic for selectivity curve structure in integrated stock assessment models. Fisheries Research, DOI: 10.1016/j.fishres.2013.12.009.

Wild, A. 1986. Growth of yellowfin tuna, *Thunnus albacares*, in the eastern Pacific Ocean based on otolith increments. Inter-American Tropical Tuna Commission Bulletin, 18: 421–482.

Williams, E. H., and Shertzer, K. W. 2003. Implications of life-history invariants for biological reference points used in fishery management. Canadian Journal of Fisheries and Aquatic Sciences, 60: 710–720.

Worm, B., Barbier, E. B., Beaumont, N., Duffy, J. E., Folke, C., Halpern, B. S., Jackson, J. B. C., *et al.* 2006. Impacts of biodiversity loss on ocean ecosystem services. Science, 314: 787–790.

Zhu, J.-F., Chen, Y., Dai, X. J., Harley, S. J., Hoyle, S. D., Maunder, M. N., and Aires-da-Silva, A. 2012. Implications of uncertainty in the spawner-recruitment relationship for fisheries management: an illustration using bigeye tuna (*Thunnus obesus*) in the eastern Pacific Ocean. Fisheries Research, 119–120: 89–93.

Handling editor: Ernesto Jardim