Reviews

Review of integrated size-structured models for stock assessment of hard-to-age crustacean and mollusc species

Andre E. Punt[1,2]*, TzuChuan Huang[1], and Mark N. Maunder[3,4]

[1]School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA 98195-5020, USA
[2]CSIRO Wealth from Oceans Flagship, GPO Box 1538, Hobart, TAS 7001, Australia
[3]Quantitative Resource Assessment LLC, San Diego, CA 92129, USA
[4]Alternative address: Inter-American Tropical Tuna Commission, 8604 La Jolla Shores Drive, La Jolla, CA 92037-1508, USA

*Corresponding Author: tel: +1 206 221 6319; fax: +1 206 685 7471; e-mail: aepunt@uw.edu

Crustaceans and molluscs such as crabs, rock lobsters, prawns, abalone, and oysters constitute large and valuable fisheries. However, assessments of these species are hampered because they cannot be production aged, in contrast to many teleosts. The major data sources for these species, in addition to catch and abundance index data, are the size compositions of the catches and of any fishery-independent indices. Assessments of such species have been conducted using age-based methods of stock assessment, as well as surplus production models. However, size-structured methods are now preferred because they can make full use of size-composition data, are able to integrate multiple sources of data, and produce the types of outputs which are needed for management purposes. An advantage of size-based models over age-based models is that all processes can be size-based, and these processes can modify the (unmodelled) size-at-age distribution. We review these methods, highlighting the choices that need to be made when developing integrated size-structured stock assessments, the data sources which are typically available and how they are used for parameter estimation, and contrast a number of such assessments worldwide.

Keywords: Abalone, assessment, Bayesian, crab, estimation, maximum likelihood, oyster, prawn, rock lobster, size-structure.

Introduction

Single-species stock assessments of fish, invertebrates, and other species are conducted for a variety of purposes. Primary amongst these are to estimate current and historical biomass, to determine trends in biomass and recruitment, to evaluate stock status relative to reference points, to form the basis for forecasts to evaluate the implications of different management actions, and to serve as the basis for the application of harvest control rules. Single-species stock assessments can be classified into three broad classes: (i) surplus production (or biomass dynamics) models, (ii) age-structured models, and (iii) size- (or length-/ stage-) structured models. Yield-per-recruit models are often referred to as a class of stock assessment method, but these models do not provide estimates of biomass; rather they quantify the effects of choices for selectivity and fishing mortality on yield- (and spawning biomass-) per-recruit. The vast majority of

contemporary stock assessments that attempt to reconstruct population biomass for marine species are based on age-structured models. However, such methods cannot easily be applied to species that are hard to age, such as crustaceans and molluscs (in particular crabs, rock lobsters, prawns, abalone, and oysters), and teleosts such as tunas and billfish, all of which form the basis for very valuable fisheries.

Assessment methods should be selected to produce the types of outputs needed for management purposes, and to utilize the available data to the maximum extent possible. The data available for hard-to-age species typically include catch time-series (often from a variety of "fleets"), and indices of relative abundance, including those based on sampling of commercial catches and form fishery-independent sources. Data on tagging are available for some species. However, the most common sources of additional data available for hard-to-age species are the size-composition

of catches, and perhaps also size-compositions from fishery-independent sampling (size is often length in actual applications, but may be carapace width or another measure of the "size").

A variety of methods have been developed to conduct assessments for hard-to-age marine species (see, for example, the review by Smith and Addison (2003), which focuses on crustaceans). Surplus production models (e.g. Schaefer, 1954, 1957; Pella and Tomlinson, 1969; Polacheck *et al.*, 1993) are perhaps the simplest method of assessment because they require only a catch time-series and an index of abundance. The performance of these methods depends on how the values for the parameters of the model are estimated (e.g. Polacheck *et al.*, 1993; Ono *et al.*, 2012), with methods now available to allow for both observation and process error (e.g. Meyer and Millar, 1999). The results from these methods are sensitive to the assumed production function (Maunder, 2003), and their performance is also related to whether or not the impacts of transient age-structure are substantial (Punt and Szuwalski, 2012). Moreover, Maunder and Starr (1995) show how some key management quantities (e.g. current biomass as a proportion of the biomass corresponding to maximum sustainable yield) for New Zealand rock lobster stocks are highly sensitive to the Pella-Tomlinson shape parameter, and that estimating this parameter is problematic. Although biomass dynamics models have been extensively applied to shellfish (Smith and Addison, 2003), they cannot use a primary data source for most hard-to-age species—the catch size-compositions.

Delay-difference models (e.g. Deriso, 1980; Schnute, 1985; Horbowy, 1992) extend surplus production models by explicitly including the biological processes of natural mortality, growth and recruitment. While more biologically realistic, these methods cannot easily use catch size-composition data. Smith and Addison (2003) highlight the value of DeLury-type depletion and change-in-ratio methods as the basis for estimating stock abundance in data-poor situations. However, the assumptions on which these methods are based, such as that the population is closed with respect to recruitment and immigration, are often violated, and these methods have only rarely been applied to provide management advice for major fished stocks.

The focus of this review is on methods that can utilize size-composition data to estimate biomass and fishing mortality. The next section describes three major methods that have been applied in the past. However, the bulk of the review is focused on integrated size-structured methods of stock assessment. Conceptually, these methods arose from age-structured methods of stock assessment, in particular, the so-called "statistical" catch-at-age methods, primarily the method of Fournier and Archibald (1982), who formulated the first likelihood-based catch-at-age analysis, and that of Deriso *et al.* (1985), who outlined how an age-structured model could be fitted using weighted least squares. The earliest integrated size-structured assessment method was developed by Sullivan *et al.* (1990). That method considered a single population and sex, allowed selectivity to be gamma- or logistic-shaped for a single fishery, and assumed that the expected growth-increment was governed by the von Bertalanffy growth curve, while the distribution for the growth increment was gamma-distributed. The basic approach of Sullivan *et al.* (1990) has been extended substantially since. The theory of size-structured models is well developed (e.g. Caswell, 1989). The key developments required for using these models for stock assessment purposes are tailoring them to the specifics of the species concerned, as well as parameter estimation and representation of

uncertainty. These aspects are considered below. Table 1 lists recent applications of integrated size-structured stock assessment methods, which, along with studies that provided key historical developments in the field, will be used to illustrate available modelling options.

The focus for this review is on methods that involve estimating parameters to determine stock status. We do not consider models, the parameters for which are either guestimated or taken from other assessments, such as the individual-based models developed by Beard and Essington (2000) and Shin and Cury (2001), and the size-structured model developed by Bergh and Johnston (1992), although the methods outlined below have a similar structure to some of these models.

Historical approaches to the use of size data for assessment purposes

Although methods for estimating gear selectivity and growth curves using size-composition data have been developed, the focus for this section is on methods that have been developed for estimating fishing mortality and population biomass and abundance from such data (also see reviews by Pauly and Morgan, 1987; Gulland and Rosenberg, 1992; Sparre and Venema, 1998).

Methods for estimating total and fishing mortality using size-frequency data

There are two general historical approaches to estimating total mortality: (i) length-converted catch curves, and (ii) Beverton and Holt's average length method. Many variants of these approaches exist.

Length-converted catch curves

Catch curves have historically been used to estimate total mortality rates by regressing the logarithm of catch-at-age in numbers on age. Similarly, length-converted catch curves (Pauly 1983, 1984a and b) are based on the observation that in equilibrium (constant recruitment and total mortality), if length is related deterministically to age, the catch over some range of lengths ($L_1 : L_2$), C_{1-2} is given by:

$$C_{1-2} = \Omega e^{-Zt_{[L_1+L_2]/2}}(t_2 - t_1) \qquad (1)$$

where Z is total mortality, Ω is a constant, and t_i is the age corresponding to the length L_i, and $t_{[L_1+L_2]/2}$ denotes the age corresponding the mean of lengths L_1 and L_2. t_i can be computed using Equation 2 under the assumption that length-at-age follows a von Bertlanffy growth curve:

$$t_i = t_0 - \frac{1}{K}\ell n(1 - L_i/L_\infty) \qquad (2)$$

where L_∞, t_0 and K are the parameters of the von Bertalanffy growth curve. Consequently, the time to grow from length L_1 to length L_2 is:

$$t_2 - t_1 = \frac{1}{K}\ell n\left(\frac{L_\infty - L_1}{L_\infty - L_2}\right) \qquad (3)$$

Total mortality is estimated using this method by regressing $C_{i-(i+1)}/(t_{i+1} - t_i)$ on $t_{[L_i+L_{i+1}]/2}$ where $C_{i-(i+1)}$ is the catch between lengths L_i and L_{i+1}.

The Jones and van Zalinge extension to length-converted catch curves (Sparre and Venema, 1998) involves regressing the

Table 1. Examples of integrated size-structured assessments.

Stock	Reference
Rock lobster	
American lobster*	Chen *et al.,* 2005; ASMFC, 2009
Southern rock lobster	
Australia	
Tasmania*	Punt and Kennedy, 1997; Hartmann *et al.,* 2011
Victoria	Hobday and Punt, 2001; Walker *et al.,* 2012
South Australia	Punt *et al.,* 2012
Western Australia*	De Lestang *et al.,* 2011
New Zealand	
Northland (CRA 1)	Starr *et al.,* 2003
Bay of Plenty (CRA 2)	Starr *et al.,* 2003
Gisborne (CRA 3)*	Breen *et al.,* 2009; Starr *et al.,* 2009
Wairarapa – Hawke Bay (CRA 4)	Breen *et al.,* 2012
Marlborough Sounds to Cape Jackson (CRA 5)	Haist *et al.,* 2011
Southland (CRA 7 & CRA 8)	Breen *et al.,* 2006[+]
South Africa	
West coast*	S. Johnston (Univ. of Cape Town, pers. comm.)
South coast	S. Johnston (Univ. of Cape Town, pers. comm.)
Abalone	
Australia	
New South Wales	Worthington, 1997
New Zealand	
Chalky Inlet & South Coast (PAU 5A)	Fu and McKenzie, 2010a
Milford Sound, George Sound, Central Coast & Dusky Inlet (PAU 5A)*	Bentley *et al.,* 2001; Fu and McKenzie, 2010b
Stewart Island (PAU 5B)	Breen and Smith, 2008
Otago & Southland (PAU 5D)	Breen and Kim, 2007
Marlborough Sounds, Tasman & Golden Bays (PAU 7)	Fu, 2012
Crab	
Alaska	
red king crab, Bristol Bay*	Zheng and Siddeek, 2011
red king crab, Kodiak Island	Zheng *et al.,* 1996
red king crab, Norton Sound	Hamazaki and Zheng, 2011
snow crab, Eastern Bering Sea*	Turnock and Rugulo, 2011
Tanner crab, Eastern Bering Sea	Rugulo and Turnock, 2011
blue king crab, St Matthews Island	Gaeuman, 2011
Dungeness crab	Zhang *et al.,* 2004
Fraser delta, Canada	
Scampi	
Bay of Plenty, Wairarapa/Hawke Bay, and Auckland Islands (SCI 1)	Tuck and Dunn, 2012
Sea Scallops	
New England	NEFSC, 2004, 2010
Oyster	Fu and Dunn, 2009
Foveaux Strait, New Zealand	
Sea urchin	
Maine, USA	Chen and Hunter, 2003; Kanaiwa *et al.,* 2005
Prawns	
Australia's northern prawn*	Punt *et al.,* 2010
Pink Shrimp	
Kachemak Bay, Alaska	Fu and Quinn, 2000

The examples listed above are the most-recent published assessments of the stocks concerned (two references are given when the most recent assessment report does not summarize the methodology). The assessments indicated by asterisks are considered in more detail in Tables 2 and 3. + No longer considered reliable.

cumulative catch from various sizes, L, (C_{L+}) on $\ell n(L_\infty - L)$ using the equation:

$$\ell n C_{L+} = d + Z[t_0 - \tfrac{1}{K}\ell n(1 - L/L_\infty)] \qquad (4)$$

Pauly (1990) introduced a form of length-converted catch curve that allows for seasonal growth.

Care needs to be taken when selecting the length-classes to include in Equations 1 and 4 to avoid sizes that are poorly selected by the fishery (and hence for which total mortality will not be independent of size), and those for which the sample size is low. The choice of the largest length is also important because the denominator in Equation 3 can get to be very large if the final length is close to L_∞.

Beverton and Holt's method

Beverton and Holt (1956) derived the following estimator for total mortality, Z [see Ehrhardt and Ault (1992) for a recent derivation]:

$$Z = K \frac{L_\infty - \bar{L}_{L'}}{\bar{L}_{L'} - L'} \qquad (5)$$

where $\bar{L}_{L'}$ is the mean length of fish of length L' and larger where L' is a length such that all fish of that length and larger are fully selected by the fishery. Sparre and Venema (1998) introduce an estimator for Z that uses the average length of the total catch:

$$Z = K \frac{L_\infty - \bar{L}}{\bar{L} - L_{50}} \qquad (6)$$

where $\bar{L}$ is the average length of the catch, and L_{50} is the length-at-50% selectivity (assuming that selection increases monotonically with length). A number of other variants of this basic approach have been developed [see, for example, Sainsbury (1982) and the examples in Pauly and Morgan (1987)], several of which estimate the parameters of the growth curve as well as Z, and provide estimators for the variance for the estimates of Z. All of the methods in this class assume that the population is in equilibrium and that recruitment is constant, and will provide biased estimates when these assumptions are violated.

Length-based Virtual Population Analysis

Jones (1981, 1984) introduced a length-based analogy to the traditional Virtual Population Analysis (VPA) method. The method calculates the numbers by length-class using the formula:

$$N_{L_1} = N_{L_2} \left(\frac{L_\infty - L_1}{L_\infty - L_2} \right)^{M/K} + C_{1-2} \left(\frac{L_\infty - L_1}{L_\infty - L_2} \right)^{M/K}, \qquad (7)$$

where N_{L_i} is the number of animals at length L_i, C_{1-2} is the catch (in numbers) between lengths L_1 and L_2 and M is the rate of natural mortality. Pope and Jiming (1987) outline how this approach can be extended to apply to multiple species that interact through predation. In common with VPA, application of Equation 7 requires an estimate of the number of animals in the final length-class, in addition to estimates of M and K. In common with most of the historical approaches to using catch-at-length data, and unlike a traditional VPA, Equation 7 is based on the assumption of time-invariant fishing mortality and constant recruitment. This constraint can be overcome to some extent by averaging the catch-at-length data over several years.

Statistical-structured catch-at-size age methods

An alternative to length-based VPA is to assign size-classes to age-classes by "slicing" the size-composition data to create age-composition data that can then be used in age-based stock assessment methods. Applications of "slicing" have, however, tendeds to be somewhat *ad hoc*. MULTIFAN (Fournier *et al.*, 1990) provides a more formal way to create age-composition data from size-composition data. It predicts the observed size-composition data as the sum of normal distributions—one size distribution for each age-class. Each normal distribution is calculated using three parameters: the mean size, the standard deviation of the sizes about the mean size, and the proportion a given age comprises of the total size-composition data set. The mean sizes-at-age are

often represented by the von Bertalanffy growth curve, and the standard deviation of size-at-age by a linear function of mean size. The proportions for each age-class are estimated as free parameters in MULTIFAN. MULTIFAN-CL (Fournier *et al.*, 1998) integrated the MULTIFAN method with an age-structured model. The proportions for each age-class are therefore calculated in MULTIFAN-CL using an age-structured model, rather than being estimated as free parameters. The likelihood function includes the size-composition data as well as catch and effort data (and any other data). This approach is common to most contemporary integrated age-structured stock assessment programs (e.g. Stock Synthesis) (Methot and Wetzel, in press).

The expected size-composition can be computed as follows, given an age-structured population dynamics model that produces a matrix of catches-at-age, $C_{y,a}$:

$$C_{y,L} = \sum_a \int_{L_{low}}^{L_{hi}} C_{y,a} \, \phi(a, l) dl, \qquad (8)$$

where $C_{y,L}$ is the expected catch (in numbers) in size-class L during year y, $\phi(a, l)$ is the probability distribution for size (l) given age a, and L_{low} and L_{hi} are respectively the upper and lower bounds for size-class L.

Contemporary stock assessment packages based on age-structured population dynamics models (e.g. Stock Synthesis; Methot and Wetzel, in press) can use size-composition data as well as other data for parameter estimation (Maunder and Punt, in press). It is common to include both age- and size-composition data, and even data on age conditioned on size, in a stock assessment so that the data are included in the form they were collected, rather than manipulating the data to fit the limitations of the assessment method.

The disadvantage of using age-structured catch-at-size methods is that modelled processes such as selectivity are generally age-based so that the modelled size-at-age distributions do not change over time. This is particularly important for stocks that are managed using a minimum legal size and experience very high fishing mortalities for legal individuals so that size-specific fishing mortality will change the size-at-age distributions. Figure 1 illustrates the difference between the predicted age- and size-structure for a population with knife-edged selectivity (mimicking a minimum legal size) and high fishing intensity, modelled using age- and size-structured models. The predicted length- and age-distributions from the age- and size-structured models are similar for the smallest (poorly-selected) and largest (fully-selected) animals, but there are marked differences in predicted size-structure for animals near the minimum legal size (Figures 1a and b). The age-structured model, which assumes that the length-at-age distribution does not change over time, hence reduces the predicted number of animals below the minimum legal size (because the sizes of some cohorts would span the minimum legal size), and does not predict a marked reduction in numbers immediately above the minimum legal size (dashed line in Figure 1c). (McGarvey *et al.* (2007) propose a class of age-structured models that do not have this problem.) In contrast, the size-structured model only removes animals above the minimum legal size. Removing the individuals from the model at the correct length is particularly important when fitting to length-composition data, otherwise the model may
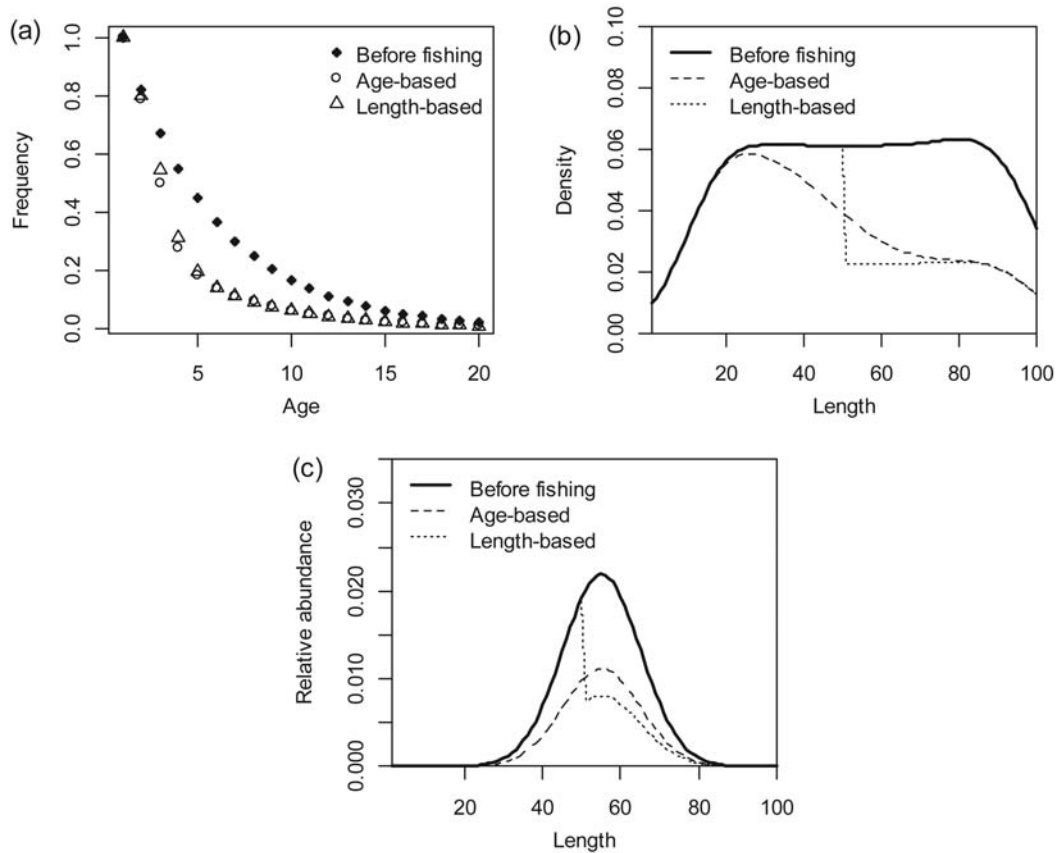
Figure 1. Initial (unfished) age-structure of a simulated population (solid circles), and the corresponding size-structure when growth follows a von Bertalanffy growth curve with $l_\infty = 100$ cm and $K = 0.2$ year^{-1} with a standard deviation of length-at-age of 10 mm ['solid line, (b)]. Panel (a) also shows the age-structure of the simulated population when it is fished if selectivity is knife-edged at 50 mm, fully-selected fishing mortality is 1 year^{-1}, and the population dynamics are modelled using an age-structured (open circles) or a size-structured (triangles) model. The dashed and dotted lines in panel (b) show the corresponding size-structures. Panel (c) shows the predicted length-at-age distribution for age 4, which is an age close to the simulated minimum legal size.

compensate for the misfitting by changing another model parameter, causing a bias in the parameter estimates.

Integrated size-structured assessment methods

The basic dynamics equation governing integrated size-structured assessment methods is in matrix algebra form:

$$N_t = \mathbf{X}_{t-1}\mathbf{S}_{t-1}N_{t-1} + R_t \qquad (9)$$

where N_t is a column vector of length H (number of size classes) containing the numbers-at-length at the start of time-step t, $\mathbf{X}_t$ is the transition matrix (dimension HxH) for time-step t (in the simplest models this matrix determines growth among size-classes), $\mathbf{S}_t$ is a HxH diagonal square matrix, with the diagonal terms equal to the proportion surviving (fishing and natural) mortality at size during time-step t, and R_t is a column vector of length H containing the recruitment to each size-class during time-step t. Figure 2 illustrates each of the processes in Equation 9 when selectivity is a logistic function of size, growth is governed by a von Bertalanffy growth curve, fully-selected fishing mortality is 1 year^{-1}, and recruitment to the modelled population occurs as a normal distribution centred at 40 mm. To accentuate the effect on the length distribution, the population is initially unfished
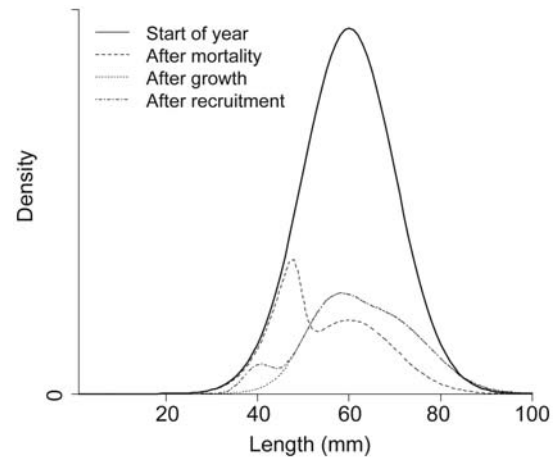

Figure 2. Changes over time in the length-distribution of a population.

(solid line) and then subjected to high fishing mortality. The Effects of growth and recruitment are also illustrated. The length distribution is initially truncated by fishing (dashed line), growth then increases the sizes of the survivors (dotted line),

and recruitment replaces the animals that grew from the smallest size-classes (dash-dot line).

Depending on how it is implemented, Equation 9 can be generalized to include a wide variety of size-structured methods. For example, the simplest size-structured models represent the population using a small number of size-classes. Stock assessments based on Equation 9 are said to be "integrated" if they make use of multiple data types, and they generally also distinguish noise in the population dynamics from that in the observation process (Maunder and Punt, in press). The selections which are needed to fully specify this model include: (i) how the state variable (N) is structured and the model time-step, (ii) how the numbers-at-size at the start of the first time period in the model are specified, (iii) how growth is modelled, (iv) how recruitment is modelled, and (v) how fishing and natural mortality are modelled, including how fishery selectivity and discards are modelled. Each of the following sub-sections outlines the options considered in actual stock assessments based on the integrated size-structured assessment paradigm.

Specification of the state variable

Increasing complexity/realism occurs as the dimension of N is increased (allowing for sex generally doubles the length of the vector N, and allowing for m areas increases its length by a factor of m). The most basic formulation of N is that all animals by size are considered exchangeable, irrespective of sex, maturity state, etc. The simplest extension to this is to allow for sex (Hobday and Punt, 2001), but it is not uncommon for size-structured population dynamics models to consider spatial structure with movement among spatial areas (e.g. Haist *et al.*, 2009; McGarvey *et al.*, 2010; de Lestang *et al.* 2011), and to represent different shell conditions (e.g. Zheng *et al.* 1995, 1996; Siddeek and Zheng, 2007; Zheng and Siddeek, 2011), as well as maturity stages (e.g. Rugulo and Turnock, 2011; Turnock and Rugulo, 2011). The assessment model for western Australian rock lobster, *Panulirus cygnus*, models the number of "whites" and "reds" separately because they have different dynamics (de Lestang *et al.*, 2011). Modelling the mature and immature components of the population separately is important for cases (such as Eastern Bering Sea snow crab, *Chionoecetes opilio*, and the Tanner crab, *Chionoecetes bairdi*), for which there is a terminal moult to maturity so that the matrix X is a diagonal matrix for mature animals, but has off-diagonal elements for immature animals. The dimension of the matrix X increases to match that of N when allowance is made for sex, area, etc.

The number of size-classes considered in a model is a balance between realism, computational speed, and availability of data. For example, Breen and Kendrick (1997) developed a 3-size-class model and considered only males, while Schnute (1987) showed that Equation 9 could be represented as a delay-difference model under some restrictive assumptions. However, most size-structured stock assessments include > 20 size-classes.

The time-step is often annual (e.g. Hobday and Punt, 2001), but has been half-year season (e.g. Chen *et al.*, 2005) and month (e.g. Punt and Kennedy, 1997), depending on data availability, the need to model closed seasons, etc.

Spatial structure has been included in integrated size-structured models for a variety of reasons. For example, Punt and Kennedy (1997) developed an eight-region model for southern rock lobster, *Jasus edwardsii*, off Tasmania, Australia, owing primarily to latitudinal differences in growth rate, and to regional variation

in fishing pressure. The assessment for *Panulirus cygnus* is based on 11 regions along the Western Australian coast, with the results pooled into three zones for reporting purposes (de Lestang *et al.*, 2011). This model is based on a simulation model for *Panulirus cygnus* developed by Walters *et al.* (1993). Breen *et al.* (2006) developed a two-stock population dynamics model for *Jasus edwardsii* off southern New Zealand, where the primary data source for estimating movement was the length–frequency distribution for the different areas. There were similar biological differences among areas for *Jasus lalandi* off South Africa (S. J. Johnston, University of Cape Town, pers. Comm.), and *Jasus edwardsii* off southern New Zealand, but a defining reason for including spatial structure in those assessments was the need to provide management advice for areas that are subject to different total allowable catch limits. Except for two management regions off southern New Zealand (CRA 7, CRA 8) (Haist *et al.*, 2009), the populations of rock lobster in the management regions off New Zealand are modelled as separate independent populations.

Initial conditions

The numbers-at-size at the start of the first year considered in the model can be treated as estimable parameters (e.g. Turnock and Rugulo, 2011), perhaps subject to "smoothness penalties". However, it is common to adopt a more parsimonious formulation. For example, Punt and Kennedy (1997) set the numbers-at-size 20 years before data are available, based on the assumption that the population was in equilibrium, and projected the population forward with estimated annual recruitments so that the numbers-at-size at the start of the first year with data was non-equilibrium. In contrast, Sullivan *et al.* (1990) estimated the total numbers at the start of the first year, and calculated the initial numbers-at-size by assuming that the catches-at-size for the first year were measured without error, with estimated selectivity relating catch numbers to starting population numbers.

Natural and fishing mortality

The matrix S in Equation 9 represents the combined effects of natural and fishing mortality. Natural mortality is usually assumed to be constant over size [but see Turnock and Rugolo (2011) for an exception to this] and time [see Zheng *et al.* (1995) for an assessment for red king crab, *Paralithodes camtschaticus*, in Bristol Bay Alaska, in which natural mortality is time-varying], while fishing mortality is usually the sum of contributions from a variety of sources. Most assessments include multiple fleets. For example assessments of *Jasus edwardsii* off Australia and New Zealand include removals by commercial and recreational fishers, mortality owing to discarded animals, and the impacts of illegal fishing. In contrast, assessments for Alaska crab include catches by the directed (pot) fishery, bycatch in the (trawl and pot) fisheries for groundfish, as well as discards of males and females in the directed fishery. See the section on catch and discard data below for the methods used to implement fishing mortality.

In common with age-based integrated frameworks (e.g. Bull *et al.*, 2005; Methot and Wetzel, in press), there are many ways to model selectivity [which captures the combined effects of gear selectivity and availability in most integrated size-structured models; see, for example, NEFSC (2004)], including making allowance for time-dependence in selectivity. A feature found in several species to which integrated size-structured models have been applied (such as abalone and rock lobster) is a minimum legal

landing size that can be implemented accurately by fishers, given the nature of the fisheries, leading to a discrete change in selectivity at the minimum legal size.

Several of the species to which size-structured methods of stock assessment have been applied are highly selective and there is little discard and bycatch (e.g. abalone). In contrast, bycatch and discard can be very high in some fisheries. For example, the discard of *Chionoecetes bairdi* in the eastern Bering Sea has regularly exceeded the retained component of the catch, primarily because of discard in pot fisheries other than the directed pot fishery. Discard and bycatch can be included in assessments either be pre-specifying the (dead) discard as a proportion of the catch below a minimum legal size (e.g. Gaeuman, 2011) or by treating bycatch in non-directed fisheries as a separate fleet (e.g. Rugulo and Turnock, 2011; Turnock and Rugulo, 2011; Zheng and Siddeek, 2011). In these latter cases, observer estimates of discard and bycatch (along with the associated size-composition data) can be used when fitting the model. Crustaceans will not necessary die after being discarded, and a key parameter is the mortality rate for discards. The value for this parameter will depend on the gear that caught the animal, conditions on deck at the time of capture, etc.

Growth

Size-based models require a matrix that determines how many animals move from each size-class into the other size-classes (or, for example, transit among maturity classes or areas) during a time-step. When considered as a growth (or size-transition) matrix, $\mathbf{X}$ is often specified to be lower triangular to represent the assumption that animals (such as rock lobsters and abalone) do not shrink when they grow. The matrix $\mathbf{X}$ is usually modelled assuming that the expected growth increment follows the von Bertalanffy growth curve (i.e. the growth increment is a decreasing linear function of current size), Schnute's (1981) general curve, or a quadratic or exponential function of size. $\mathbf{X}$ is either pre-specified based on auxiliary information, or estimated along with the other parameters of the model. The distribution for growth increment has variously been assumed to be gamma (Sullivan *et al.*, 1990), or normal (Punt *et al.*, 1997), but other options such as log-normal are feasible. A variety of methods (e.g. Punt *et al.*, 1997; McGarvey and Feenstra, 2002; Chen *et al.* 2003; NEFSC, 2004) have been developed to estimate size-transition matrices using tagging data. It is arguably better to model the probability of moving from one size-class to the other size-classes as a multinomial process (Punt *et al.*, 2010) because this explicitly captures the discrete nature of growth rather than approximating growth using, for example, a von Bertalanffy growth equation. Punt *et al.* (2009) compare the performance of a subset of these methods using simulation.

A disadvantage of pre-specifying the size-transition matrix rather than estimating it simultaneously with the other parameters is that the resulting matrix may be inconsistent with other information included in the assessment (such as the catch size-composition data). This can arise if the range of sizes for recaptured animals is narrow compared to that in the catches. This can occur for abalone and rock lobsters because it is common to tag predominantly sub-legal animals (Punt *et al.*, 1997). Tagging analyses to estimate the size-transition matrix conducted outside of an assessment may also include some assumptions that are violated (e.g. that selectivity is uniform over sizes), but that can be

accounted for better within a stock assessment model. The assessment will be overly pessimistic if a growth matrix is pre-specified which implies that animals grow to larger size than they really do because the interpretation of few animals observed near the inferred asymptotic size is that fishing mortality must be high. Furthermore, pre-specifying the size-transition matrix ignores its uncertainty, as well as other data (e.g. size composition), which may provide information on growth and can be included in an integrated stock assessment.

The parameters defining the matrix $\mathbf{X}$, which include both those which define the mean and uncertainty of the growth increment as a function of size, can be estimated by including the growth increment data from tagging experiments in the assessment's objective function (e.g. Bentley *et al.*, 2001; Punt *et al.*, 2010). Many species (e.g. lobsters and crabs) assessed using size-structured models grow by moulting, and growth is therefore conditional on the probability of moulting. Zheng *et al.* (1995, 1996, 1998) therefore model the process of moulting using an inverse logistic function and the growth increment given that a crab moults, as separate processes, and allow for time-dependence in the probability of moulting.

Growth can be assumed to occur once each year (e.g. Hobday and Punt, 2001). However, models with multiple within-year time-steps can allow growth to occur multiple times during the year (e.g. twice – de Lestang *et al.*, 2011; four times – Punt and Kennedy, 1997; weekly – Punt *et al.*, 2010). Allowing growth to occur a number of times during the year increases biological realism, but at computational cost as well as requiring the estimation of more parameters. Punt *et al.* (2010) overcomes the latter problem by assuming that the parameters determining the matrix $\mathbf{X}$ are the same for all weeks during the year.

Recruitment

Recruitment is usually modelled as the product of annual recruitment and the proportion of the annual recruitment that recruits to each size-class. The annual recruitments are either estimated as free parameters, assumed to be temporally auto-correlated (e.g. Chen *et al.*, 2005) or related to spawning stock biomass according to a stock-recruitment relationship (e.g. Breen *et al.*, 2000b; Chen *et al.*, 2000; Chen and Hunter, 2003). The proportion of the recruitment in each size-class can either be pre-specified (e.g. Punt and Kennedy, 1997; Chen *et al.*, 2005) or estimated along with the other parameters of the model (e.g. Zheng and Siddeek, 2011). Recruits have been assigned to the first size class (Punt and Kennedy 1997), distributed uniformly over a pre-specified number of size classes (Bentley *et al.*, 2001), or distributed using a distribution such as normal (Starr *et al.*, 2009) or gamma (Turnock and Rugulo, 2011). Punt *et al.* (2010) treats the annual recruitments as parameters of the model, but fit a Ricker stock-recruitment relationship to outputs from the model to form the basis for projections. Most assessments assume an equal male-female split of the annual recruitment, but this need not be so (e.g. Zheng and Siddeek, 2011).

Incorporation of stock-recruitment relationships in size-structured models can be problematic, because without age there is no obvious link between spawning stock size and recruitment at some size. Assumptions consequently have to be made about the age of the recruits. It is generally assumed that all animals entering the model are of the same age, although this is unlikely to be correct.

Maturity

Maturity can be modelled using a function that represents the proportion mature at size, or mature and immature animals can be modelled separately. The advantage of modelling mature and immature animals separately is that selectivity to the fisheries may be a function of maturity, which may modify the probability of being mature as a function of size over time. Maturity may only occur after growth, and may therefore be coupled with the probability of moulting and the growth transition matrix. Models that represent mature and immature animals separately require a function which defines the probability of maturing as a function of size. This function has been estimated externally to the assessment (e.g. Turnock and Rugulo, 2009) or as set of estimated parameters which are constrained using a penalty function to be smooth and monotonically increasing as a function of size (e.g. Rugulo and Turnock, 2011; Turnock and Rugulo, 2011). Zheng et al. (1998) allow for the possibility that the probability of maturing is time-varying.

Parameter estimation

The parameter estimation procedure used for integrated size-structured assessment models is essentially the same as for their age-structured counterparts. The parameters of a size-structured stock assessment are essentially the same as for their age-structured counterparts, and typically include those which define the size-structure and abundance at the start of the first year, the annual deviations in recruitment about a time-series of expected values, those which determine annual (fully-selected) fishing mortality by fleet and sex, natural mortality, those which determine growth and maturity, and those which determine selectivity. Some of these parameters are often pre-specified (e.g. the relationship between size and weight, natural mortality), while the remaining parameters are estimated by fitting the population dynamics model to available data. Many assessments, particularly those that are spatially-structured or include many fleets, make the assumption that some parameters are equal between areas or fleets to reduce the number of estimable parameters. For example, it is common to assume that the selectivity patterns for recreational and commercial fishers are the same when conducting assessments of rock lobsters (e.g. Punt and Kennedy, 1997), primarily because the same gear types are used, due to paucity of data for sectors other than the commercial fishery.

A variety of data sources have been used to fit integrated size-structured models. The data typically include fishery catches, fishery catch-rate data [perhaps with a non-linear relationship between catch-rate and abundance; e.g. Chen et al. (2005)], fishery-independent relative and absolute indices of abundance, tagging data, size-composition data for fisheries (including samples of non-retained animals), and any fishery-independent monitoring data.

The earliest (and some current) size-structured assessment methods were based on weighted least squares, where the weights were selected semi-arbitrarily, and sensitivity was explored to choices of the weights (e.g. Sullivan et al., 1990; Zheng et al., 1996). However, the most recent assessments based on size-structured population dynamics models have generally been based on an objective function that approximates a likelihood function. Reasons for the change in strategy regarding the objective function may be that weighting of data becomes somewhat more straightforward (but see Discussion below), and because it becomes possible to represent

uncertainty using methods such as profile likelihood and Bayesian posteriors. For assessments in which the objective function is a likelihood, parameter estimation is either based on maximum likelihood, penalized maximum likelihood or Bayesian methods.

Catch and discard data

The catches are either assumed to be measured without error (e.g. the rock lobster and abalone assessments off Australia and New Zealand) (Punt and Kennedy, 1997; Breen et al., 2009), or assumed to be normally or log-normally distributed (e.g. Alaskan crab assessments-Turnock and Rugulo, 2011; Siddeek and Zheng, 2007). A particular problem associated with removals due to bycatch and discard (and recreational fisheries and illegal harvest) is that catch data are often not available for these fleets, or they are sparse and may be imprecise if available. Approaches identified to deal with this problem range from postulating scenarios for these catches and then assuming the catch time-series to be measured without error and that the fishing mortality by these fleets is related to observed effort, or using a relative measure of removals and estimating a scaling factor to relate the relative measure to actual removals (e.g. Plagányi et al., 2011).

Abundance index data

In common with most age-structured stock assessments, indices of relative abundance are usually assumed to be log-normally distributed about the corresponding model predictions. However, some assessments use "robust" likelihood functions. For example, Chen et al. (2000) explore the robust normal, t and log-normal distributions for index data. In general, abundance indices are assumed to be linearly related to population biomass. However, for fishery catch-rate indices in particular, this assumption may be violated. Consequently, some assessments have estimated a non-linear relationship between catch-rate and abundance (e.g. Breen et al., 2009; Fu and McKenzie, 2010a and b). Catchability can also be assumed to depend on time of year (e.g. Punt and Kennedy, 1997; Punt et al., 2010) as well as on size and temperature (e.g. de Lestang et al., 2011).

Compositional data

Size-composition data have been included in size-structured stock assessments under the assumption that they are multinomially distributed, or distributed according to the Fournier et al. (1990) robust normal distribution for proportions (e.g. Breen et al., 2000a and b). Punt and Kennedy (1997) used a lognormal likelihood function for the proportions at length, and assumed that the variance was inversely proportional to the proportion. In principle, the effective sample sizes for the length-frequencies collected during surveys can be inferred from the survey data themselves (e.g. Pennington and Vølstad, 1994). However, the estimates of sampling error for survey length–frequency data may still under-estimate the true extent of variability between the model estimates and survey-based size-compositions because, for example, the model does not consider process error such as temporal changes in selectivity or growth. Punt and Kennedy (1997) used a lognormal likelihood function for the proportions at length and assumed that the variance was inversely proportional to the proportion. Compositional data are usually size-frequencies by sex (because species such as crabs and lobsters are easy to sex externally, and many fisheries only allow the retention of males so the fishing mortality differs between the genders), but size-frequencies for some fisheries are also available by shell condition and maturity stage. Shell

condition provides a rough measure of age and may help with the estimation of some parameters (e.g. new shell animals will be animals which recently moulted) or retention may differ among shell conditions. Some assessments (e.g. Haist *et al.*, 2011; Breen *et al.*, 2012) have restricted the range of sizes included in the likelihood function by implementing minus and plus groups to avoid fitting the model to small proportions, which may be unduly influential, while the length-based assessment method CASA (NEFSC, 2004) allows for errors when measuring the sizes of animals.

Tagging data
Tagging data can be used to inform growth rates (e.g. Bentley *et al.*, 2001; Punt *et al.*, 2010), movement patterns (McGarvey *et al.*, 2010), total mortality (which may be split into fishing and natural mortalities), and selectivity. To date, focus has been placed on the first two of these uses, primarily because of concerns regarding tag loss and, in particular, tag-reporting rates. Use of tagging data to estimate growth rates involves integrating the approaches for estimating size-transition matrices into the stock assessment. Breen *et al.* (2003) (and associated assessments) assume that the growth increment is normally distributed about its expected value and account for the measurement error that arises when measuring recaptured lobsters (not usually recaptured by scientists), while Punt *et al.* (2010) assume that the probability of recapturing a tagged animal in a given size-class is the outcome of a multinomial experiment. McGarvey *et al.* (2010) include tagging data to estimate movement rates conditional on animals being recaptured, which ensures that the results are unbiased in the face of any constant rate of tag non-reporting (but depend on the assumption that reporting rates do not vary spatially during any time period, a much weaker assumption).

Priors and penalties
Almost all integrated size-structured stock assessments include some form of prior information (i.e., they are not strictly maximum likelihood estimators). Most size-structured assessments place a penalty (or prior) on the deviations in recruitment, about mean recruitment, or about the stock-recruitment relationship (the deviations are usually assumed to be normally-distributed around the natural logarithm of median recruitment so recruitment is log-normal). Priors must be imposed on all of the parameters of the model to construct a Bayesian posterior distribution for the parameters and model outputs. However, the priors for parameters are usually chosen to be "non-informative" (often taken to be uniform over some range) (e.g. Breen *et al.*, 2003). Exceptions are the placement of informative priors on natural mortality (e.g. Breen *et al.*, 2003), on the steepness of the stock-recruitment relationship (e.g. Chen *et al.*, 2000), on the parameters that determine the size-transition matrix and the probability of maturing (e.g. Turnock and Rugulo, 2011), and on survey catchability (e.g. Zheng and Siddeek, 2011; Tuck and Dunn, 2012).

Representation of uncertainty
Representing uncertainty has become a key consideration in fisheries management (Patterson *et al.*, 2001), and some fisheries jurisdictions (e.g. the USA and the International Whaling Commission) explicitly "discount" catch limits given the extent of scientific uncertainty. A variety of methods exist to evaluate whether the specifications for an assessment are reasonable, and to quantify uncertainty.

Model fit diagnostics
Standard model fit diagnostics (such as plotting observed data versus model-predictions, Q-Q plots, residual plots. etc.) are commonly used to identify model mis-specification. These diagnostics (e.g. patterns in residuals) often reveal potential conflicts among data sources (e.g. the indices may indicate declining abundance, but, in contrast, the presence of large animals in catches and surveys implies that the reduction in population abundance cannot be too great). The correlations among the parameter estimates can be examined to identify whether any parameters are obviously confounded.

A second set of ways to evaluate model fit relates to whether the residuals correspond to the assumptions regarding the extent of uncertainty associated with the data (e.g. due to pre-specifying the variance about an index or the effective sample size for data which are assumed to be multinomially distributed). In New Zealand, the standard deviation of the normalized (Pearson) residuals, which should be 1 if the assumptions regarding standard deviations and effective sample sizes are correct, is commonly used to evaluate these assumptions. A variety of plots and diagnostic statistics, such the plot of the pre-specified effective sample sizes against the effective sample sizes computed using the method of McAllister and Ianelli (1997), have been developed. A summary of the statistics and plots used for Alaskan crab stock assessment is available at http://www.alaskafisheries.noaa.gov/npfmc/PDFdocuments/resources/SAFE/Appendix_CrabWKSHPReport909.pdf.

Retrospective analyses (e.g. Mohn, 1999; Ralston *et al.*, 2011) provide another tool to diagnose model problems. These involve removing the most recent year of data and re-running the assessment, removing the two most recent years of data, etc., and plotting key model outputs (such as time-trajectories of biomass). Consistent higher or lower estimates of biomass as more years are removed indicate that the model is mis-specified. Similarly, prospective analyses (leaving out the first year of data and re-running the assessment, removing the first two years of data, etc.) can be informative. However, caution should be taken before adjusting the assessment results to remove any consistent bias because the model using all the data may not necessarily be the least biased.

Quantifying uncertainty
Closed form solutions for standard deviations of quantities of interest (such as maximum sustainable yield, MSY, and current depletion level) do not exist for integrated size-structured models. Four basic approaches have therefore been applied to estimate standard deviations and confidence intervals (credibility intervals) of parameters and model outputs: (i) approximate asymptotic methods, (ii) likelihood profile, (iii) bootstrapping, and (iv) Bayesian methods. The first three of these approaches are frequentist, while Bayesian methods, usually based on Markov chain Monte Carlo, sampling, MCMC, lead to Bayesian measures of uncertainty (Punt and Hilborn, 1997). In many cases, these four methods all lead to essentially the same results. However, it is well known (e.g. Ralston *et al.*, 2011) that all of these methods tend to underestimate the true extent of uncertainty because they are all based on the assumption that the population dynamics (including the pre-specified parameter values) and observation models (including the pre-specified residual variances and effective sample sizes) are correct.

In recent years, Bayesian methods, along with penalized likelihood methods, have been the most commonly applied approaches for quantifying uncertainty. The primary advantage of the Bayesian approach is the ease with which the results of Bayesian methods can form the basis for forecasts and decision analyses. However, sampling from Bayesian posteriors is still computationally very challenging, and failure of the MCMC algorithm to converge is common even after this algorithm has been applied for millions or tens of millions of cycles. In addition, the specification of uninformative (or uninfluential) priors, which is often an aim when selecting priors, may be difficult, particularly in data-limited situations.

Given the fact that variance estimates are conditional on model assumptions, most assessments explore the sensitivity of the results by modifying some of the assumptions of the assessment, such as including or excluding data sources, changing values for pre-specified parameter values (such as natural mortality, M), and the weights assigned to the different data sources. In principle, Bayesian model averaging (Hoeting *et al.*, 1999) can be used to combine results across alternative models (assuming that they use the same data) [see, for example, Brandon and Wade (2006) for whale assessments]. However, this has yet to be done for size-structured integrated models.

Summaries of applications

Table 2 summarizes the structure and key modelling assumptions for nine actual stock assessments based on the integrated size-structured approach, illustrating the variety of assumptions made when conducting assessments. All of the assessments are sex-structured, with two of them including multiple areas linked by movement. The treatment of selectivity (survey and fishery) varies considerably among assessments. Reasons for this include whether there is a legal minimum size, whether the fisheries are male-only, and the types of data available for assessment purposes. Recruitment was log-normal in all assessments, with the distribution of recruitment sizes ranging from the first size-class only to a continuous distribution over all size-classes. Growth was estimated in three of the assessments, and was pre-specified based on auxiliary information in remaining assessments. Moulting was explicitly modelled in three of the nine assessments. Only two of the assessments (southern New Zealand abalone and eastern Bering Sea snow crab) estimated the maturity ogive; the remaining pre-specified it. Natural mortality was generally pre-specified, but was estimated in three of the assessments. Only the assessment of Bristol Bay red king crab allowed for time-varying natural mortality. Most of the assessments calculated the initial size-structure by projecting from an equilibrium state, but three assessments treated the initial size-structure as parameters to estimate (subject to constraints).

Table 3 summarizes the data included in the objective function and the approach for parameter estimation. Two of the assessments (rock lobster off Tasmania and New Zealand) were based on the Bayesian paradigm while the remaining assessments involved penalized maximum likelihood. Several of the assessments imposed priors on key parameters (such as natural mortality and catchability). All but the two Alaskan crab assessments used commercial CPUE data, and all except southern New Zealand lobster incorporated fishery-independent index data. Discards were included in four of the nine assessments, but only two assessments estimated parameters related to discards. All of the assessments made use of tagging data, but only four assessments

included the tagging data in the likelihood function. Three of these assessments used the tagging data to estimate growth, and the fourth assessment used these data to estimate movement rates. Recruitment indices were available for three of the stocks, but were only formally included in the likelihood function for one assessment (Western Australian rock lobster).

Discussion
Advantages and disadvantages of size-based integrated methods

Sullivan *et al.* (1990) identify the benefits of integrated size-structured approaches over, for example, length-cohort analysis. These approaches require no assumption of deterministic growth, and no assumption that the population is in steady state. Additionally, integrated methods can use multiple sources of data to recreate the population state, and can represent the uncertainty associated with the model outputs in several ways.

Size-structured stock assessments have been applied for a variety of purposes, primary amongst these are to estimate biomass, and biomass relative to reference points such as the biomasses at which maximum sustainable yield (e.g. Turnock and Rugolo, 2011; Zheng and Sideek, 2011) and maximum economic yield (e.g. Punt *et al.*, 2010) are achieved. They have also formed the basis for projections to evaluate changes in, for example, catch levels (e.g. Chen and Hunter, 2003; Siddeek and Zheng, 2007), minimum legal sizes (e.g. Green *et al.*, 2012), and definitions for the proxy level of fishing mortality at which maximum sustainable yield is achieved (e.g. Siddeek and Zheng, 2007). Fitted size-structured models have also formed the basis for evaluating management procedures (Butterworth, 2007) for rock lobster fisheries in New Zealand, South Africa and Australia (e.g. Starr *et al.*, 1997; Johnston and Butterworth, 2005; Punt *et al.*, 2012). The evaluated management procedures have tended to be "empirical", e.g. based on trends in commercial catch-rates, primarily because it is computationally easier to evaluate such management procedures compared to management procedures which involve fitting population dynamics models, and because they tend to be more transparent to stakeholders and decision-makers. However, fitted size-structured models have also been used to evaluate model-based management procedures (e.g. Punt *et al.*, 2012).

Parameter confounding can be more serious in size-structured than age-structured models, particularly when there is considerable variation in size-at-age making it hard to discern cohorts in the size-composition data. This type of confounding can be alleviated, but not eliminated, by including tagging data in the assessment. Natural mortality can be confounded with selectivity in both age- and sex-structured models. The availability of survey data, for a survey which selects all available animals can potentially remove this confounding. This type of data is used to estimate selectivity for snow crab in the Eastern Bering Sea.

Maunder and Punt (in press) discuss the major disadvantages of integrated models (size- and age-structured), e.g. they tend to be very complicated and hence not transparent to biologists and decision makers (Hilborn, 1997), they can require a considerable amount of data [and perform poorly if the size-transition matrix is uncertain; Punt (2003)], and, in common with all integrated methods of stock assessment, they can be subject to model-mis-specifications and the implications of contradictory data. Weighting of multiple data sources is often a major challenge

Table 2. Summary of the model structure assumptions on which the nine stock assessments identified by asterisks in Table 2 are based.

Stock	Assessment structure	Selectivity/catchability	Growth	Recruitment
American lobster	Sex (except for GBK); 5-mm CL classes	Commercial selectivity separated into legal size, gear characteristics (escape gap), conservation measures (discarding berried females), and other. Some vary over time, sexes, and quarters. Survey selectivity double logistic, shared for quarter and sex, but catchability differs	Sex-specific; varies by season; variation in growth normal; calculated outside model	Annual autocorrelated lognormal deviates; proportion at length fixed for first three size groups; sex ratio of recruits estimated
Tasmanian rock lobster	Sex; 2-mm CL classes; 11 regions, some linked by movement	Selectivity logistic; same for commercial and recreational fleets	Sex-specific; von Bertalanffy; varies by season; variation in growth normal; calculated outside model	Annual lognormal deviates; all recruitment to first size-class; 1:1 sex-ratio
New Zealand rock lobster (CRA 3)	Sex, maturity; 2-mm CL classes;	Selectivity double normal	Sex-specific; Schnute (1981) growth model; variation in growth normal with s.d. proportional to growth increment; estimated in the assessment	Annual lognormal deviates; distributed over sizes using a truncated normal; 1:1 sex-ratio
South African west coast rock lobster	Sex; 1-mm CL classes	Selectivity three parameter logistic, but modified for different fisheries and sexes (e.g. broken stick for small, linear with time for females); sex-specific; female selectivity scaled so maximum is not one (vulnerability)	Adults: linear growth increment with time-varying intercept; juvenile growth (equal for both sexes): quadratic function of length with time blocks. variation in growth normally distributed truncated at ± 2 s.d.; some lobsters can shrink; calculated outside the assessment	Beverton-Holt stock-recruitment curve used to penalize recruitment deviations; uniformly distributed over first 15 size classes; 1:1 sex ratio
Western Australian rock lobster	Sex, whites/red; 2-mm CL classes; 11 regions linked by movement	Selectivity pre-specified, depends on escape gap size; catchability depends on red/white, regions, maturity state; temperature	Logistic function of size; varies by season; variation in growth normal with cv 0.05; calculated outside the assessment	Annual lognormal deviates by region and year; pre-specified length distribution
Southern New Zealand abalone (PAU 5A)	Single-sex; 2-mm SL classes	Selectivity logistic with time change due to change in harvest size	Schnute (1981) growth model; variation in growth normal; estimated in the assessment	Annual lognormal deviates by region and year; equal over the first five size-classes
Eastern Bering Sea Snow Crab	Sex, maturity, shell condition; 5-mm CW classes,	Survey selectivity and catchability time blocks; sex-specific and logistic; retention selectivity by shell condition; availability shared between surveys.	Sex-specific; von Bertalanffy; equal intercept for males and females; terminal moult; variation in growth gamma; estimated in the assessment.	Annual lognormal deviates; truncated gamma length distribution
Bristol Bay red king crab	Sex, shell condition; 5-mm CL classes	Selectivity and retention logistic; some sex-specific; special selectivity for plus group; male pot bycatch broken stick; trawl survey selectivity logistic, separate for males and females, time blocked	Sex-specific; von Bertalanffy; time dependent for females; variation in growth gamma; calculated outside the assessment	Annual lognormal deviates; gamma length distribution
Northern Australian prawns	Sex; 1-mm CW classes	Catch and survey selectivity logistic	Sex-specific; von Bertalanffy; variation in growth normal; estimated in the assessment	Annual lognormal deviates; all recruitment to first size-class; 1:1 sex-ratio; estimated proportions by week

Continued

Table 2. *Continued*

Stock	Natural Mortality	Molting	Initial conditions	Seasons	Maturity
American lobster	0.15 year^{-1}, sensitivity analyses to a change over time	Not modelled explicitly	Scaling parameter for males and females; initial proportions based on equilibrium conditions with F similar to that in first model year	Quarterly, growth and recruitment in some quarters, some parameters can vary by quarter	Pre-specified proportion mature at size
Tasmanian rock lobster	0.1 year^{-1} for all classes	Not modelled explicitly	Estimated by projecting from fished equilibrium	Monthly; growth and recruitment in some months	Pre-specified proportion mature at size
New Zealand rock lobster (CRA 3)	Estimated	Not modelled explicitly	Estimated by projecting from fished equilibrium	Seasonal; growth occurs each season; recruitment in first season	Pre-specified probability of maturing
South African west coast rock lobster	Small lobster have different M, with linear increase to larger sizes; sex specific	Adults once a year; juveniles multiple times a year, inter-molt period is a linear function of size	Starts from unfished equilibrium	Annual model	Pre-specified proportion mature at size
Western Australian rock lobster	0.3 year^{-1} for whites; 0.22 year^{-1} for reds	Not modelled explicitly	Estimated by projecting from fished equilibrium	11 time steps; growth and recruitment in a subset of months	Pre-specified proportion mature at size
Southern New Zealand abalone	Estimated	Not modelled explicitly	Starts from unfished equilibrium	Annual model	Logistic proportion mature at size estimated in the assessment
Eastern Bering Sea Snow Crab	Estimated for immatures, males and females	All immature moult annually; mature don't moult; terminal moult	Estimated by length, sex, and shell condition; smoothed with first differences	Annual model; accounts for fishery timing	Size-specific parameters for the probability of maturing smoothed with second difference; estimated in the assessment
Bristol Bay red king crab	Estimated, sex-specific additional mortality; base M constant over length and shell condition	Estimated for males negative logistic; equal to one for females	Smoothed survey size and shell composition from first survey with estimated scaling parameters by sex	Annual model; accounts for fishery timing	Logistic proportion mature at size with 50% maturity time blocks or annual estimates
Northern Australian prawns	Pre-specified by species; constant over size and time	Not modelled explicitly	Estimated by projecting from fished equilibrium	Weekly growth and recruitment	Pre-specified proportion mature at size

CL = carapace length, CW = carapace width, SL = shell length.

Table 3. Data included in the nine stock assessments identified by asterisks in Table 1.

Stock	Fishery data	Discards	Survey	Tagging	Recruitment indices	Inference method, data weighting and priors
American lobster	Catch lognormal; robust normal for compositions, dynamic binning to ensure proportions in tails greater than or equal to 0.01	Assumed nil	Index lognormal; compositions robust normal	Used to estimate parameters outside model	None	Penalized maximum likelihood; iterative reweighting to estimate s.d.'s and sample sizes.
Tasmanian rock lobster	Commercial and recreational fleets; catches assumed known; catch-in-numbers and CPUE lognormal; weighted normal for compositions	Assumed nil	Change-in-ratio indices normal	Used to estimate growth outside model	Pre-recruit indices available, but not used	Bayesian; weighting factors estimated
New Zealand rock lobster (CRA 3)	Commercial, recreational, customary and illegal fleets; catches assumed known; CPUE lognormal; length composition multinomial by maturity state summed to one over males and females	Retention curve	None	Used in model to estimate growth	Pre-recruit indices available, but not used	Bayesian; residual variances estimated; informative priors on recruitment deviations
South African west coast rock lobster	Single fleet; CPUE lognormal; weighted normal for compositions (sex-specific); percent females binomial	Applied to lobster less than a certain size	Index lognormal; weighted normal for compositions (sex-specific); percent females binomial	Used to estimate growth outside model	None	Penalized maximum likelihood; residual variances estimated
Western Australian rock lobster	Square root of catch normal; sex-specific catch size-composition data mutltinomial	Predicted using selectivity ogive; assumed discard rate 0.03; 10% high-grading	Index lognormal; compositions multinomial	Used to estimate some movement parameters	Related to recruitment using a power function	Penalized maximum likelihood; weights estimated for indices; pre-specified for compositions, tagging and recruitment data
Southern New Zealand abalone (PAU5A)	Commercial, recreational, customary and illegal fleets; catches assumed known; CPUE lognormal; weighted normal for compositions	Assumed nil	Index lognormal; weighted normal for compositions	Used in model to estimate growth	None	Bayesian; weighting factors estimated
Eastern Bering Sea Snow Crab	Catch normal, separate for retained and total. Shell condition, sex, size-composition robust multinomial	Discard sex and size composition robust multinomial; assumed handling mortality rate	Total biomass by sex and maturity separately lognormal; compositions by sex, size, and shell condition robust multinomial	Used to develop priors for growth parameters	None	Penalized maximum likelihood; weights pre-specified, priors on M and mean growth parameters
Bristol Bay red king crab	Catch lognormal. Size and shell condition composition normal	Size, shell condition, sex composition normal; assumed handling mortality rate; effort used to predict historical discards	Index lognormal; compositions by sex and shell condition lognormal	Used to estimate growth outside model	None	Penalized maximum likelihood; weights pre-specified, priors on q
Northern Australian prawns	Three species linked through bycatch; square root of catch by week normal; sex-specific catch size-composition mutltinomial	Assumed nil	Two surveys; lognormal additional variance estimated; sex-specific survey compositions mutltinomial	Used in model to estimate growth	None	Penalized maximum likelihood; priors on recruitment deviations

A. E. Punt et al.

for integrated methods of stock assessment (Maunder and Punt, in press), and this is clearly the case for size-structured methods. To these general concerns with integrated methods must be added that all size-structured methods rely on the assumption that biological and fishery processes are size- and not age-based. This assumption could be violated if animals mature or migrate ontogenetically so, for example, selectivity is age- as well as size-based. Furthermore, age-structured models fitted to reliable data on fishery or survey age-composition should be able to estimate year-class strength. In contrast, the ability to estimate year-class strength from size-composition data depends critically on being able to adequately characterize growth. Consequently, year-class strength may be incorrectly assigned if growth rates change over time, but growth is assumed to be time-invariant.

Evaluating the performance of integrated size-structured methods

Considerable work has been undertaken to evaluate the estimation performance of integrated size-structured methods of stock assessment. For example, Fu and Quinn (2000) evaluated whether it was feasible to estimate time-trends in natural mortality using integrated size-based methods, Punt (2003) found that the ability to estimate the virgin biomass depends critically on having catch-rate or size-composition data for earliest years of exploitation, that estimates can be highly biased and imprecise in the absence of such data, and highlighted the importance of the size-transition matrix on the performance of integrated size-structured methods. Several assessment models (e.g. Rugulo and Turnock, 2011; Turnock and Rugulo, 2011) impose penalties on the extent to which fishing mortality can vary from one year to the next. Szuwalski and Punt (2012) identified the key role that these penalties can play in (i) stabilizing estimation, and (ii) biasing estimates of stock status; and Punt et al. (2012) quantify the extent of bias in estimation which results due to violation of the assumption that population dynamics processes are not stationary.

Chen et al. (2000) used simulation to explore the consequences of violation of the form of the likelihood function on the performance of size-structured assessment methods, and recommended using "fat-tailed" distributions for likelihoods and priors, while Chen et al. (2005) evaluated the stock assessment for *Homarus americanus* and found it to be robust to errors in fishery selectivity, growth and landings.

Other integrated models that fit to size-composition data

A further extension of integrated models is to represent both age- and size-structure explicitly. While straightforward in simulation models, this adds a substantial computational burden to the calculations. De Leo and Gatto (1995) outline an age- and size-structured model and fit it to the data for European eel, *Anguilla anguilla*. However, that application considers only a single year of data. Deriso and Parma (1988) outline a full age-size-structured population dynamics model and describe the likelihood function which could be used to estimate its parameters for Pacific halibut, *Hippoglossus stenolepis*. Quinn et al. (1998) extend this approach by discretizing the size distributions. Gilbert et al. (2006) develop an age- and size-structured model for New Zealand snapper, *Pagrus auratus*, that allows growth to be a function of both age and length and to vary over time, and

McGarvey et al. (2007) modify a standard age-structured model to allow for a legal minimum size. A few general stock assessment models are based on age- and size-structured models (e.g. Fleksibest – Froysa et al., 2002, Gadget – Begley, 2005 [http://www.hafro.is/gadget/]). Taylor and Methot (in press) outline an extension to Stock Synthesis that explicitly accounts for size-specific mortality by dividing each cohort into a number "platoons", each of which has its own growth curves and which, in principle, unifies age- and size-structured models within a parsimonious framework. However, this approach has not yet been fully evaluated or applied for management purposes.

Future directions

There is a trend towards the development of general software packages for conducting stock assessments based on the integrated paradigm. Several such packages have been developed for age-structured models (e.g. Stock Synthesis – Methot and Wetzel, in press; CASAL – Bull et al., 2005; MULTIFAN-CL – Fournier et al., 1998). These platforms lead to an enhanced opportunity to evaluate the sensitivity of results to model assumptions in a straightforward manner, and to view the fits of the model to the data [e.g. Taylor et al. (2011) for Stock Synthesis]. In addition, the availability of general software packages increases the number of scientists who are familiar with the way assessments are conducted, which simplifies the process of reviewing assessments as well as increasing the likelihood that major programming errors will be detected and corrected. Given its general availability, Stock Synthesis has already been the focus of several studies using simulation testing to evaluate performance (e.g. Garrison et al., 2011; Yin and Sampson, 2004; Wetzel and Punt, 2011; see the review in Punt and Maunder, in press). Disadvantages of general packages are that (i) it can be difficult to add new options unless the software is written in a modular manner and a version control system is in place to ensure that multiple changes are not made simultaneously, (ii) the code will not be as efficient as it could be if it was developed for a specific application, and (iii) the management-related outputs may not be suitable for all management systems. In relation to this last disadvantage, while all packages estimate quantities such as MSY, the ability to use generalized packages as the basis for management strategy evaluation is limited because, in general, the management arrangements for any given fishery are very specific to that fishery. However, adoption of a modular approach to programming can overcome this potential problem (A. E. Punt, pers. comm.). A key needed development is to construct a general software package for applying size-structured models. The stock assessment program CASAL (Bull et al., 2005), which is generally used for age-based assessments, has a length-based option and has been used for the assessment of scampi (Tuck and Dunn, 2012) and oysters (Fu and Dunn, 2009). CASAL may not be able to model moulting appropriately for species such as crab. Such an effort is currently underway (A. E. Punt, pers. comm.).

Most size-structured stock assessment methods are single-species. Punt et al. (2010) conduct projections for three species in Australia's northern prawn fishery in which account is taken for technical interactions that arise due to bycatch when targeting particular prawn species.

Model selection remains very complicated for integrated size-structured models. A key need is the development of methods for model selection and identification of model mis-specification.

Conclusion

Integrated size-structured methods of stock assessment, while complicated and occasionally difficult to apply, provide analysts with a way to relatively seamlessly integrate a wide variety of data sources for hard-to-age species within a single modelling framework. These approaches are now being adopted widely for assessment and management purposes in Australia, South Africa, New Zealand and the USA. Their broader adoption will require the development and testing of generic software platforms, as has been done for integrated age-structured methods of assessment.

Acknowledgements

Funding

References

CORPASMFC (Atlantic States Marine Fisheries Commission). 2009. American Lobster Stock Assessment Report for Peer Review. Stock Assessment Report No. 09-01 (Supplement), Washington DC.

Beard, T. D., and Essington, T. E. 2000. Effects of angling and life history processes on bluegill size structure: insights from an individual-based model. Transactions of the American Fisheries Society, 129: 561–568.

Begley, J. 2005. Gadget User Manual. Technical Report 120, Marine Research Institute, Reykjavik.

Bentley, N., Breen, P. A., Starr, P. J., and Kendrick, T. H. 2001. Assessment of the CRA 3 and NSS substocks of red rock lobster (*Jasus edwardsii)* for 2000. New Zealand Fisheries Assessment Report 2001/69.

Bergh, M. O., and Johnston, S. J. 1992. A size-structured model for renewable resource management, with application to resources of rock lobster in the south-east Atlantic. South African Journal of Marine Science, 12: 1005–1016.

Beverton, R. J. H., and Holt, S. J. 1956. A review of methods for estimating mortality rates in exploited fish populations, with special reference to sources of bias in catch sampling. Rapp. P.-V. Reun CIEM., 140: 67–83.

Brandon, J. R., and Wade, P. R. 2006. Assessment of the Bering-Chukchi-Beaufort Seas stock of bowhead whales using Bayesian model averaging. Journal of Cetacean Research and Management, 8: 225–239.

Breen, P. A., Andrew, N. L., and Kendrick, T. H. 2000a. The 2000 stock assessment of paua *(Haliotis iris)* in PAU 5B using an improved Bayesian length-based model. New Zealand Fisheries Assessment Report 2000/28.

Breen, P. A., Andrew, N. L., and Kendrick, T. H. 2000b. The 2000 stock assessment of paua *(Haliotis iris)* in PAU 5B and PAU 5D using a new length-based model. New Zealand Fisheries Assessment Report 2000/33.

Breen, P. A., Haist, V., and Starr, P. J. 2006. CRA2006 Objective 4 Stock assessment of red rock lobsters (*Jasus edwardsii)* in CRA 7 and CRA 8 in 2006, using a new multi–stock length–based model (MSLM). Unpublished Final Research Report to Ministry of Fisheries, 19 December 2006.

Breen, P. A., Haist, V., Starr, P. J., and Kendrick, T. H. 2009. The 2008 stock assessment of rock lobsters (*Jasus edwardsii*) in CRA 3. New Zealand Fisheries Assessment Report 2009/23.

Breen, P. A., Haist, V., Starr, P. J., and Pomarede, M. 2012. The 2011 stock assessment and management procedure development for red rock lobsters (*Jasus edwardsii*) in CRA 4. New Zealand Fisheries Assessment Report 2012/09.

Breen, P. A., and Kendrick, T. H. 1997. A fisheries management success story: the Gisborne, New Zealand, fishery for red rock lobsters (*Jasus edwardsii*). Marine and Freshwater Research, 48: 1103–1110.

Breen, P. A., and Kim, S. W. 2007. The 2006 stock assessment of paua (*Haliotis iris*) stocks PAU 5A (Fiordland) and PAU 5D (Otago). New Zealand Fisheries Assessment Report 2007/09.

Breen, P. A., Kim, S. W., and Andrew, N. L. 2003. A length-based Bayesian stock assessment for the New Zealand abalone *Haliotis iris*. Marine and Freshwater Research, 54: 619–634.

Breen, P. A., and Smith, A. N. H. 2008. The 2007 assessment for paua (*Haliotis iris*) stock PAU 5B (Stewart Island). New Zealand Fisheries Assessment Report 2008/05.

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.

Butterworth, D. S. 2007. Why a management procedure approach? Some positives and negatives. ICES Journal of Marine Science, 64: 613–617.

Caswell, H. 1989. Matrix Population Models: Construction, Analysis, and Interpretation. Sinauer Associates, Sunderland, MA.

Chen, Y., Breen, P. A., and Andrew, N. L. 2000. Impacts of outliers and mis-specification of priors on Bayesian fisheries-stock assessment. Canadian Journal of Fisheries and Aquatic Sciences, 57: 2293–2305.

Chen, Y., and Hunter, M. 2003. Assessing the green sea urchin (*Strongylocentrotus droebachiensis*) stock in Maine. USA. Fisheries Research, 60: 527–537.

Chen, Y., Hunter, M., Vada, R., and Beal, B. 2003. Developing a growth-transition matrix for the stock assessment of the green sea urchin (*Strongylocentrotus droebachiensis*) off Maine. Fishery Bulletin, 101: 737–744.

Chen, Y., Kanaiwa, M., and Wilson, C. 2005. Developing and evaluating a size-structured stock assessment model for the American lobster, *Homarus americanus*, fishery. New Zealand Journal of Marine & Freshwater Research, 39: 645–660.

De Leo, G. A., and Gatto, M. 1995. A size and age-structured model of the European eel (*Anguilla anguilla* L.). Canadian Journal of Fisheries and Aquatic Sciences, 52: 1351–1367.

De Lestang, S., Caputi, N., How, J., Melville-Smith, R., Thomson, A., and Stephenson, P. 2011. Draft stock assessment for the West Coast rock lobster fishery. Fisheries Research Report No. 217. Department of Fisheries, Western Australia.

Deriso, R. B. 1980. Harvesting strategies and parameter estimation for an age structured model. Canadian Journal of Fisheries and Aquatic Sciences, 37: 268–282.

Deriso, R. B., and Parma, A. M. 1988. Dynamics of age and size for a stochastic population model. Canadian Journal of Fisheries and Aquatic Sciences, 45: 1054–1068.

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

Ehrhardt, N. M., and Ault, J. S. 1992. Analysis of two length-based mortality models applied to bounded catch length frequencies. Transactions of the American Fisheries Society, 121: 115–122.

Fournier, D. A., and Archibald, C. P. 1982. A general theory for analysing 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., Sibert, J. R., Majkowski, J., and Hampton, J. 1990. MULTIFAN a likelihood-based method for estimating growth parameters and age-composition from multiple length frequency data sets illustrated using data for southern bluefin tuna (*Thunnus maccoyii*). Canadian Journal of Fisheries and Aquatic Sciences, 47: 301–317.

Froysa, K. G., Bogstad, B., and Skagen, D. W. 2002. Fleksibest – an age-length structured fish stock assessment model. Fisheries Research, 55: 87–101.

Fu, D. 2012. The 2011 stock assessment of paua (*Haliotis iris*) for PAU 7. New Zealand Fisheries Assessment Report 2012/27.

Fu, D., and Dunn, A. 2009. An updated stock assessment of Foveaux Strait dredge oysters (*Ostrea chilensis*) for the 2008–09 fishing year. New Zealand Fisheries Assessment Report 2009/53.

Fu, D., and McKenzie, A. 2010a. The 2010 stock assessment of paua (*Haliotis iris*) for Chalky and South Coast in PAU 5A. New Zealand Fisheries Assessment Report 2010/36.

Fu, D., and McKenzie, A. 2010b. The 2010 stock assessment of paua (*Haliotis iris*) for Milford, George, Central, and Dusky in PAU 5A. New Zealand Fisheries Assessment Report 2010/46.

Fu, C., and Quinn, T. J. 2000. Estimability of natural mortality and other population parameters in a length-based model: *Pandalus borealis* in Kachemak Bay, Alaska. Canadian Journal of Fisheries and Aquatic Sciences, 57: 2420–2432.

Gaeuman, W. 2011. 2011 Saint Matthew Island blue king crab stock assessment. Pp. 417–468. *In* Stock Assessment and Fishery Evaluation Report for the King and Tanner Crab Fisheries of the Bering Sea and Aleutian Islands Regions. North Pacific Fishery Management Council, Anchorage, AK, USA.

Garrison, T. M., Hamel, O. S., and Punt, A. E. 2011. Can data collected from marine protected areas improve estimates of life-history parameters? Canadian Journal of Fisheries and Aquatic Sciences, 68: 1761–1777.

Gilbert, D. J., Davies, N. M., and McKenzie, J. R. 2006. Development of an age-length structured model of the Hauraki Gulf-Bay of plenty snapper (*Pagrus auratus*) population. Marine and Freshwater Research, 57: 553–568.

Green, B. S., Gardner, C., Linnane, A. D, Hobday, D., Chandrapavan, A., Punt, A., McGarvey, R., *et al.* 2012. Spatial management of southern rock lobster fisheries to improve yield, value and sustainability. FRDC Report 2006/220.

Gulland, J. A., and Rosenberg, A. A. 1992. A review of length-based approaches to assessing fish stocks. FAO Fisheries Technical Paper 323.

Haist, V., Breen, P. A., and Starr, P. J. 2009. A multi-stock, length-based assessment model for New Zealand rock lobster (*Jasus edwardii*). New Zealand Journal of Marine & Freshwater Research, 43: 355–371.

Haist, V., Breen, P. A., Starr, P. J., and Kendrick, T. H. 2011. The 2010 stock assessment of red rock lobsters (*Jasus edwardsii*) in CRA 5, and development of an operational management procedure. New Zealand Fisheries Assessment Report 2011/12.

Hamazaki, T., and Zheng, J. 2011. Norton Sound red king crab stock assessment for the fishing year 2011/12. Pp. 469–512. *In* Stock Assessment and Fishery Evaluation Report for the King and Tanner Crab Fisheries of the Bering Sea and Aleutian Islands Regions. North Pacific Fishery Management Council, Anchorage, AK, USA.

Hartmann, K., Gardner, C., and Hobday, D. 2011. Fishery Assessment Report: Tasmanian Rock Lobster Fishery 2009/10. Institute of Marine and Antarctic Studies, University of Tasmania. http://www.imas.utas.edu.au/__data/assets/pdf_file/0008/164951/RL_Assess_0910_FINAL.pdf. (last accessed September 2012).

Hilborn, R. 1997. Lobster stock assessment: report from a workshop, II. Marine and Freshwater Research 48: 945–947.

Hobday, D., and Punt, A. E. 2001. Size-structured population modelling and risk assessment of the Victorian southern rock lobster, *Jasus edwardsii*, fishery. Marine and Freshwater Research, 52: 1495–1507.

Hoeting, J. A., Madigan, D., Raftery, A.E., and Volinsky, C. T. 1999. Bayesian Model Averaging: a tutorial. Statistical Science, 14: 382–417.

Horbowy, J. 1992. The differential alternative to the Deriso difference production model. ICES Journal of Marine Science, 49: 167–174.

Johnston, S. J., and Butterworth, D. S. 2005. Evolution of operational management procedures for the South African West Coast rock lobster (*Jasus Ialandii*) fishery. New Zealand Journal of Marine and Freshwater Research, 39: 687–702.

Jones, R., 1981. The use of length composition data in fish stock assessments (with notes on VPA and cohort analysis). FAO Fisheries Circular 734.

Jones, R. 1984. Assessing the effects of changes in exploitation pattern using length composition data. FAO Fisheries Technical Paper 256.

Kanaiwa, M., Chen, Y., and Hunter, M. 2005. Assessing a stock assessment framework for the green sea urchin *Strongylocentrotus drobachiensis* fishery in Maine, USA. Fisheries Research, 74: 96–115.

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., and Punt, A. E. in press. A review of integrated analysis in fisheries stock assessment. Fisheries Research.

Maunder, M. N., and Starr, P. J. 1995. Sensitivity of management reference points to the ratio of BMSY/B0 determined by the Pella–Tomlinson shape parameter fitted to New Zealand rock lobster data. New Zealand Fisheries Assessment Document 95/10.

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

McGarvey, R., and Feenstra, J. 2002. Estimating length-transition probabilities as polynomial functions of premoult length. Marine and Freshwater Research, 52: 1517–1526.

McGarvey, R., Feenstra, J. E., and Ye, Q. 2007. Modeling fish numbers dynamically by age and length: partitioning cohorts into 'slices'. Canadian Journal of Fisheries and Aquatic Sciences, 64: 1157–1173.

McGarvey, R., Linnane, A., Feenstra, J. E., Punt, A. E., and Matthews, J. M. 2010. Integrating recapture-conditioned movement estimation into spatial stock assessment: a South Australian lobster fishery application. Fisheries Research, 105: 80–90.

Methot, R. D., and Wetzel, C. R. in press. Stock Synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fisheries Research.

Meyer, R., and Millar, R. B. 1999. BUGS in Bayesian stock assessments. Canadian Journal of Fisheries and Aquatic Sciences, 56: 1078–1087.

Mohn, R. K. 1999. The retrospective problem in sequential population analysis: an investigation using cod fishery and simulated data. ICES Journal of Marine Science, 56: 473–488.

NEFSC (Northeast Fisheries Science Center). 2004. 39th Northeast Regional Stock Assessment Workshop (39th SAW) assessment summary report. US Dept Commer., Northeast Fish. Sci. Cent. Ref. Doc. 04-10a; Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026.

NEFSC (Northeast Fisheries Science Center). 2010. 50th Northeast Regional Stock Assessment Workshop (50th SAW) Assessment Report. US Dept Commer., Northeast Fish Sci Cent Ref Doc. 10-09; http://www.nefsc.noaa.gov/publications/crd/crd1017/. (last accessed September 2012).

Ono, K., Rivot, E., and Punt, A. E. 2012. Model performance analysis using bias, precision and reliability estimators in a Bayesian framework. Fisheries Research, 125–6: 173–183.

Patterson, K., Cook, R., Darby, C., Gavaris, S., Kell, L., Lewy, P., Mesnil, B., *et al.* 2001. Estimating uncertainty in fish stock assessment and forecasting. Fish and Fisheries, 2: 125–157.

Pauly, D. 1983. Some simple methods for the assessment of tropical fish stocks. FAO Fisheries Technical Paper 234.

Pauly, D. 1984a. Length-converted catch curves. A powerful tool for fisheries in the tropics (Part II). ICLARM Fishbyte, 2: 17–19.

Pauly, D. 1984b. Length-converted catch curves. A powerful tool for fisheries in the tropics (III: Conclusion). ICLARM Fishbyte, 2: 9–10.

Pauly, D. 1990. Can we use traditional length-based fish stock assessment when growth is seasonal? ICLARM Fishbyte, 8: 29–32.

Pauly, D., and Morgan, G. R. (Eds). 1987. Length-based methods in fisheries research. ICLARM Conference Proceedings 13. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait. 468 pp.

Pella, J. J., and Tomlinson, P. K., 1969. A generalized stock production model. IATTC Bulletin, 13: 421–458.

Pennington, M., and Vølstad, J. H. 1994. Assessing the effect of intra-haul correlation and variable density on estimates of population characteristics from marine surveys Biometrics 50: 725–732.

Plagányi, É., Butterworth, D., and Burgener, M. 2011. Illegal and un-reported fishing on abalone—Quantifying the extent using a fully integrated assessment model. Fisheries Research, 107: 221–232.

Polacheck, T., Hilborn, R., and Punt, A. E. 1993. Fitting surplus production models: comparing methods and measuring uncertainty. Canadian Journal of Fisheries and Aquatic Sciences, 50: 2597–2607.

Pope, J. G., and Jiming, Y. 1987. Phalanx analysis: an extension of Jones' length-cohort analysis to multiple species cohort analysis. *In* Length-based Methods in Fisheries Research, pp. 177–191. Ed. by D. Pauly, and G. R. Morgan. ICLARM Conference Proceedings 13. International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Punt, A. E. 2003. The performance of a size-structured stock assessment method in the face of spatial heterogeneity in growth. Fisheries Research, 65: 391–407.

Punt, A. E., Buckworth, R. C., Dichmont, C. M., and Ye, Y. 2009. Performance of methods for estimating size–transition matrices using tag–recapture data. Marine and Freshwater Research 60: 168–182.

Punt, A. E., Deng, R. A., Dichmont, C. M., Kompas, T., Venables, W. N., Zhou, S., Pascoe, S., *et al.* 2010. Integrating size-structured assessment and bioeconomic management advice in Australia's northern prawn fishery. ICES Journal of Marine Science, 67: 1785–1801.

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

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

Punt, A. E., Kennedy, R. B., and Frusher, S. D. 1997. Estimating the size-transition matrix for Tasmanian rock lobster, Jasus edwardsii. Marine and Freshwater Research, 48: 981–992.

Punt, A. E., and Maunder, M. N. in press. Stock synthesis: advancing stock assessment and stock assessment research through the use of a general model. Fisheries Research.

Punt, A. E., McGarvey, R., Linnane, A., Phillips, J., Triantafillos, L., and Feenstra, J. 2012. Evaluating empirical decision rules for southern rock lobster fisheries: A South Australian example. Freshwater Research, 115–116: 6–71.

Punt, A. E., and Szuwalski, C. 2012. How well can F_{MSY} and B_{MSY} for Alaska crab stocks be estimated using empirical measures of surplus production? Fisheries Research, 134–136: 113–124.

Punt, A. E., Trinnie, F., Walker, T. I., McGarvey, R., Feenstra, J., Linnane, A., and Hartmann, K. 2012. The performance of a management procedure for rock lobsters, *Jasus edwardsii*, off western Victoria, Australia in the face on non-stationary dynamics. Freshwater Research, 137: 116–128.

Quinn, T. J., II, Turnbull, C. T., and Fu, C. 1998. A length-based population model for hard-to-age invertebrate populations. *In* Fishery stock assessment models. Ed. by F. Funk, T.J. Quinn, II, J. N. Ianelli, J. E. Powers, J. F. Schweigert, P. J. Sullivan, and C-I Zhang. Rep. No. 98-01 of the Alaska Sea Grant College Program, University of Alaska–Fairbanks, Juneau, pp. 531–556.

Ralston, S., Punt, A. E., Hamel, O. S., DeVore, J., and Conser, R. J. 2011. An approach to quantifying scientific uncertainty in stock assessment. Fishery Bulletin 109: 217–231.

Rugolo, L. J., and Turnoch, B. J. 2011. 2011 Stock Assessment and Fishery Evaluation Report for the Tanner Crab Fisheries of the Bering Sea and Aleutian Islands Regions, pp. 285–354. *In* Stock Assessment and Fishery Evaluation Report for the King and Tanner Crab Fisheries of the Bering Sea and Aleutian Islands Regions. North Pacific Fishery Management Council, Anchorage, AK, USA.

Sainsbury, K. J. 1982. Population dynamics and fishery management of the paua, *Haliotis iris* I. Population structure, growth, reproduction, and mortality. New Zealand Journal of Marine and Freshwater Research, 16: 147–161.

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

Schaefer, M. B. 1957. A study of the dynamics of the fishery for yellowfin tuna in the eastern tropical Pacific Ocean. Bulletin of the Inter-American Tuna Commission, 2: 247–268.

Schnute, J. T. 1981. A versatile growth model with statistically stable parameters. Canadian Journal of Fisheries and Aquatic Sciences, 38: 1128–1140

Schnute, J. T. 1985. A general theory for analysis of catch and effort data. Canadian Journal of Fisheries and Aquatic Sciences, 42: 414–429.

Schnute, J. T. 1987. A general fishery model for a size-structured fish population. Canadian Journal of Fisheries and Aquatic Sciences, 44: 924–940.

Shin, Y-J., Cury, P., and , P. 2001. Exploring fish community dynamics through size-dependent trophic interactions using a spatialized individual-based model. Aquatic Living Resources, 14: 65–80.

Siddeek, M. S. M., and Zheng, J. 2007. Evaluating the parameters of a MSY control rule for the Bristol Bay, Alaska, stock of red king crabs. ICES Journal of Marine Science 64: 995–1005.

Smith, M. T., and Addison, J. T. 2003. Methods for stock assessment of crustacean fisheries. Fisheries Research, 65: 231–256.

Sparre, P., and Venema, S. C. 1998. Introduction to tropical fish stock assessment. Part 1. Manual. FAO Fisheries Technical Paper 306.1 Rev. 2. Rome, FAO.

Starr, P. J., Bentley, N., Breen, P. A., and Kim, S. W. 2003. Assessment of red rock lobsters (*Jasus edwardsii*) in CRA 1 and CRA 2 in 2002. New Zealand Fisheries Assessment Report 2003/41.

Starr, P. J., Breen, P. A., Hilborn, P. A., and Kendrick, T. H. 1997. Evaluation of a management decision rule for a New Zealand rock lobster substock. Marine and Freshwater Research, 48: 1093–1101.

Starr, P. J., Breen, P. A., Kendrick, T. H., and Haist, V. 2009. Model and data used for the 2008 stock assessment of rock lobsters (*Jasus edwardsii*) in CRA 3. New Zealand Fisheries Assessment Report 2009/22.

Sullivan, P. J., Lai, H-L., and Gallucci, V. F. 1990. A catch-at-length analysis that incorporates a stochastic model of growth. Canadian Journal of Fisheries and Aquatic Sciences, 47: 184–198.

Szuwalski, C., and Punt, A. E. 2012. Identifying research priorities for management under uncertainty: The estimation ability of the stock assessment method used for eastern Bering Sea snow crab (*Chionoecetes opilio*). Freshwater Research, 134–136: 82–94.

Taylor, I. G., and Methot, R. D. in press. Hiding or dead? 1 A computationally efficient model 2 of selective fisheries mortality. Freshwater Research.

Taylor, I. G., Stewart, I. J., Hicks, A., Garrison, T. M., Punt, A. E., Wallace, J. R., and Wetzel, C. R. 2011. r4ss: R code for Stock Synthesis. R package version 1.16. http://R-Forge.R-project.org/projects/r4ss/. (last accessed September 2012).

Tuck, I. D., and Dunn, A. 2012. Length-based population model for scampi (*Metanephrops challengeri*) in the Bay of Plenty (SCI 1), Wairarapa / Hawke Bay (SCI 2) and Auckland Islands (SCI 6A). New Zealand Fisheries Assessment Report 2012/01.

Turnock, B. J., and Rugolo, L. J. 2009. Stock assessment of eastern Bering Sea snow crab. Pg. 29–123. *In* Stock Assessment and Fishery Evaluation Report for the King and Tanner Crab Fisheries of the Bering Sea and Aleutian Islands Regions. North Pacific Fishery Management Council, Anchorage, AK, USA.

Turnock, B. J., and Rugolo, L. J. 2011. Stock assessment of eastern Bering Sea snow crab. Pg. 37–168. *In* Stock Assessment and Fishery Evaluation Report for the King and Tanner Crab Fisheries of the Bering Sea and Aleutian Islands Regions. North Pacific Fishery Management Council, Anchorage, AK, USA.

Walker, T. I., Trinnie, F. I., and Reilly, D. J. 2012. Victorian rock lobster fishery stock assessment report 2012. Report to Rock Lobster Resource Assessment Group. Meeting 7. 8 March 2012. Draft 3 (14 March 2012). Fisheries Victoria Branch, Department of Primary Industries: Queenscliff, Victoria, Australia.

Walters, C. J., Hall, N., Brown, R., and Chubb, C. 1993. Spatial model for the population dynamics and exploitation of the Western Australian rock lobster, *Panulirus cygnus*. Canadian Journal of Fisheries and Aquatic Sciences, 50: 1650–1662.

Wetzel, C. R., and Punt, A. E. 2011. Performance of a fisheries catch-at-age model (Stock Synthesis) in data-limited situations. Marine and Freshwater Research, 62: 927–936.

Worthington, D. G. 1997: Demography and dynamics of the population of blacklip abalone, *Haliotis rubra*, with implications for management of the fishery in NSW, Australia. PhD thesis, Macquarie University, Sydney, Australia. 203 pp.

Yin, Y., and Sampson, D. B. 2004. Bias and precision of estimates from an age-structured stock assessment program in relation to stock and data characteristics. North American Journal of Fisheries Management, 24: 865–879.

Zhang, Z., Hajas, W., Phillips, A., and Boutillier, J. A. 2004. Use of length-based models to estimate biological parameters and conduct yield analyses for male Dungeness crab (*Cancer magister*). Canadian Journal of Fisheries and Aquatic Sciences, 61: 2126–2134.

Zheng, J., Kruse, G. H., and Murphy, M. C. 1998. A length-based approach to estimate population abundance of Tanner crab, *Chionoecetes bairdi*, in Bristol Bay, Alaska. Canadian Special Publication in Aquatic and Fishery Sciences, 125: 97–105.

Zheng, J., Murphy, M. C., and Kruse, G. H. 1995. A length-based population model and stock-recruitment relationships for red king crab, *Paralithodes camtschaticus*, in Bristol Bay, Alaska. Canadian Journal of Fisheries and Aquatic Sciences, 52: 1229–1246.

Zheng, J., Murphy, M. C., and Kruse, G. H. 1996. A catch-length analysis for crab populations. Fishery Bulletin, 94: 576–588.

Zheng, J., and Siddeek, M. S. M. 2011. Bristol Bay red king crab stock assessment in fall 2011. *In* Stock Assessment and Fishery Evaluation Report for the King and Tanner Crab Fisheries of the Bering Sea and Aleutian Islands Regions, pp. p169–283. North Pacific Fishery Management Council, Anchorage, AK, USA.

Handling editor: Howard Browman